Language, Cognition and Theory of Mind: The Emergence of Mental State Language and Mental State Understanding in the Third Year of Life

Dissertation zum Erwerb des Doctor of Philosophy (Ph.D)

am Munich Center of the Learning Sciences

der Ludwig Maximilians-Universität

München

vorgelegt von

Irina Jarvers

18.01.2021

1st Supervisor/Erstgutachterin: PD Dr. Susanne Kristen-Antonow

2nd Supervisor/Zweitgutachter: Prof. Dr. Markus Paulus

Datum der mündlichen Prüfung: 09.06.2021

ACKNOWLEDGEMENTS

First of all, I would like to express my sincere gratitude to my first supervisor PD Dr. Susanne Kristen-Antonow, who despite not being my official supervisor until the last stage of this dissertation, has been giving me continuous support and extensive feedback throughout the research and writing process. Not only have I felt supported in my scientific career, but also in my emotional well-being. I thank Prof. Dr. Beate Sodian for providing me with opportunities to broaden my research experience and knowledge through countless conference participations and lab visits. I would also like to thank the Deutsche Forschungsgesellschaft (DFG) for financing the project and for enabling me to learn from more than one area of developmental psychology through the Crossing Project.

Besides my advisor, I would like to thank the rest of my supervisory committee: Prof. Dr. Markus Paulus and Prof. Dr. Paul L. Harris, for their insightful comments and guidance to consider several theoretical perspectives. I am also grateful to Prof. Dr. Paul L. Harris for welcoming me in Harvard during my stay and not only providing research input, but also aiding me in my personal development.

My sincere thanks also goes out to PD Dr. Tobias Schuwerk who has supported my research, provided valuable insight and taught me additional methods to employ in my research. I am grateful to my fellow lab mates, especially Stella S. Grosso, for stimulating discussions and help no matter the topic.

Last but not least, I am thankful to my family. I thank my sister Maria, for supporting me to make it this far in my career, for establishing creative writing as a second passion to my love for research and for sharing the things that make me genuinely happy. Finally, I thank my husband Christian for being there for me along the way, for accepting my peculiarities and for always enabling me to pursue my personal and professional goals.

SUMMARY

Theory of mind (ToM), the ability to attribute mental states to ourselves and others, is crucial for human social interaction and has been argued to fully develop around the age of 4. However, recent research suggests that children can perform rudimentary, preverbal ToM inferences at an earlier age, indicating a discrepancy between this early, implicit ToM and a later mastery of explicit ToM tasks. Already in the second year of life children show competence in grasping what an agent knows and does not know in preverbal communication (Moll, Carpenter, & Tomasello, 2007; O'Neill, 1996), whereas explicit performance on ToM tasks can only be observed around the age of 4 (Wellman & Liu, 2004; Wimmer & Perner, 1983). One possible explanation is that children require an additional representational resource in order to be able to associate their preverbal understanding with explicit terms that can be used to denote it in standard ToM tasks. The purpose of this thesis is to extend research on ToM in the third year of life by investigating one possible representational resource, namely mental state language.

Mental state language is considered crucial for ToM development as it is the first explicit means by which children can verbalize mental states and it plays an important role for both precursors (Kristen, Sodian, Thoermer, & Perst, 2011) and subsequent mastery of explicit ToM tasks (Brooks & Meltzoff, 2015; Olineck & Poulin-Dubois, 2007). However, how exactly mental state language production reflects children's awareness and comprehension of mental states remains understudied. Harris, Yang, and Cui (2017) have argued that children use mental state terms early on and possess basic representations of their own knowledge and ignorance. Since both mental state terms and basic representations appear to be present at the same time, one might assume

that children already possess associations between them. However, experimental tasks that investigate mental state verbs like ‘know’ and ‘think’ are only able to find competence to differentiate between them around the age of 5 (Kristen-Antonow, Jarvers, & Sodian, 2019). Thus, the extent to which the appropriate use of mental state terms in conversation reflects comprehension of the denoted mental states remains unclear.

Given that mental state language develops around the second and third year of life, one may gain insight into the role it plays for a transition from implicit to explicit competence by investigating abilities which, according to our current understanding, develop around the same time and show both preverbal and verbal aspects. Perspective-taking and implicit false-belief understanding are such abilities. Children show early preverbal perspective-taking competence at the age of 24 months (Moll & Tomasello, 2006), but only develop verbal perspective-taking competence around the age of 3 (Gonzales, Fabricius, & Kupfer, 2018). Previous work has identified a relationship between mental state language and verbal perspective-taking at the age of 30 months (Chiarella, Kristen, Poulin-Dubois, & Sodian, 2013), but no study has investigated this relationship in connection to preverbal perspective-taking. Also in the area of false-belief understanding studies have mostly investigated the relationship between mental state language and explicit false-belief tasks, but no study has examined how mental state language may relate to an implicit false-belief task.

Finally, mental state language has been found to correlate with a number of different competencies, among them inhibitory control and general language, but no study has systematically investigated which of these competencies serve as precursors that predict children’s later mental state language production. Determining precursors is relevant for understanding crucial aspects of mental state language development and how it might relate to children’s preverbal representations of mental states.

The aims of the present thesis were thus to investigate the relationship between children’s mental state language production and their comprehension of mental state terms and mental states, to determine the role that mental state language production may play in preverbal and verbal perspective-taking and implicit false-belief understanding and finally, to identify the basic linguistic and cognitive skills that contribute to the development of mental state language. These aims were pursued through 2 studies.

In study 1, children were assessed at two time-points. Once they were tested at the age of 24 months with measures of general language, cognitive and motoric development and inhibition skills. An additional time they were tested at 27 months in order to assess children's mental state language production, their preverbal and verbal perspective-taking skills, meta-cognitive awareness of ignorance, implicit false-belief understanding and finally a task that measured their ability to infer a speaker's need for information from a statement about knowing or not knowing the location of an item and a speaker's desire for an object from statements about wanting or not wanting a particular object. The aim was to determine the concurrent relationship of children's mental state language production and their ability to comprehend and use these terms on experimental tasks, to identify the concurrent relationship between mental state language and preverbal and verbal perspective-taking and implicit false-belief understanding and finally to determine which developmental skills at the age of 24 months are significant predictors of children's later mental state language production. Study 1 showed that 27-month-olds were already able to produce mental state terms and showed a basic understanding of their own knowledge and ignorance, but failed to show comprehension of epistemic mental state terms in experimental tasks. Thus, despite their early competence in producing epistemic mental state terms, children's comprehension of these terms appeared to be limited. The results of the first study also identified that mental state language production was related to children's own verbal level-1 visual perspective-taking, in particular to usage of the verb 'know/don't know' in several naturalistic contexts, independent of general language. There was also continuity from children's preverbal understanding of perspectives to their ability to report their own perspective in a verbal task. However, there was no relationship between preverbal understanding of perspectives and mental state language production according to a parental questionnaire, suggesting that early preverbal understanding of perspectives

is not related to the mere production of mental state terms unless children already produce these terms in appropriate contexts. In relation to implicit false-belief understanding, children's performance suggested that they followed a simple strategy of looking at the last object location instead of demonstrating false-belief understanding. Finally, study 1 also identified sentence production, inhibition and fine motor skills at 24 months as significant predictors of children's later production of mental state language at 27 months.

In study 2, children's ability to infer a speaker's need for information was investigated further by administering a pragmatic inference task from study 1 to 2-, 3- and 5-year-olds and determining at which age children showed competence. Study 2 showed that despite being able to draw the right pragmatic inferences for the mental state verb 'want' at the age of 2, only at the age of 5 children were able to draw the right pragmatic inferences for the mental state verb 'know/don't know'.

The overall results of this thesis implied that the crucial aspect of mental state language is not mere production of the verbs, but children's experience with a variety of contexts in which mental state verbs can be used, corresponding with a socio-constructivist approach to cognitive development. The number of naturalistic contexts related to children's metarepresentations of their own ignorance, to their ability of inferring a speaker's desire and their ability to verbalize their own perspective, thus suggesting that the conversational input children receive is one of the main ways in which mental representations become explicit. Furthermore, the theoretical view that children's preverbal competence may be better explained by a production rule approach instead of early conceptual understanding is discussed.

CONTENTS

TABLES

FIGURES

ABBREVIATIONS

AOI	Area of Interest
APT	Action-based Perspective-Taking
DLS	Differential Looking Score
EPT	Epistemic Perspective-Taking
FAM	Familiarization
FB	False-Belief
FBU	False-Belief Understanding
FDR	False-Discovery Rate
FF	First Fixation
GEE	Generalized Estimating Equations
LL	Longest Look
MCI	Metacognition of Ignorance
MSL	Mental State Language
MSLQ	Mental State Language Questionnaire
MSU	Mental State Understanding
ST	Simulation Theory
ToM	Theory of Mind
TT	Theory Theory
VPT	Visual Perspective-Taking

1. INTRODUCTION

Humans are remarkably social and interact with several people a day, be it at home, at work, during grocery shopping or even with a stranger on the street. Interacting with others is so crucial that lack of it can result in negative health outcomes such as mental illness or worse recovery after a heart attack (Ruberman, Weinblatt, Goldberg, & Chaudhary, 1984; Umberson & Karas Montez, 2010). As interactions with others and keeping deep social relationships is important for our daily lives and our health, one might wonder what exactly is necessary to foster social relationships and to understand our conversational partners or interlocutors. One important aspect of communication is the ability to understand that our interlocutors have certain beliefs about things and intentions that they wish to fulfill. These beliefs and intentions can agree with ours, but they can also diverge. The awareness that mental states (e.g. beliefs, desires) exist and that these are present for both ourselves and others, is an ability that has been termed 'theory of mind' (ToM, Premack & Woodruff, 1978). Theory of mind is not only crucial for our general social interactions, but also has importance for later life outcomes, be it popularity amongst peers (Slaughter, Imuta, Peterson, & Henry, 2015) and ability to create mutual friendships in primary school children (Fink, Begeer, Peterson, Slaughter, & de Rosnay, 2015), or quality of social relationships in the elderly (Lecce et al., 2017).

Since the ability to assign mental states to ourselves and others is so important for future relationships, the question remains when children are capable of ToM. Research in infants and 2-year-old children shows that they already have a preverbal grasp of other people's goal-directed behavior (Wellman, Phillips, Dunphy-Lelii, & LaLonde, 2004), and are able to act in accordance with what their interlocutor knows or does not know (O'Neill, 1996). This suggests that either children at the age of two have conceptual understanding of mental states, in this case knowledge, that they are able to act on or that children are able to infer production rules from behavior that inable them to draw inferences similar to a working concept of mental states. Both possibilities of either preverbal production rules or preverbal concepts can be called *implicit* ToM. Implicit ToM is thus preverbal and does not require children to know language in order to be aware of another person's knowledge and to draw correct inferences or to develop accurate production rules[1]. Similarly to knowledge states, several studies have found evidence for preverbal, implicit false-belief understanding (FBU) in the second year of life (Baillargeon, Scott, & He, 2010; Knudsen & Liszkowski, 2012; Southgate, Senju, & Csibra, 2007), suggesting that children have some rudimentary conceptual representation to predict an agent's behavior according to his false belief[2]. However, despite this preverbal competence, traditional ToM tasks only find verbal and thus *explicit* competence around the age of 4 (Wellman & Liu, 2004; Wimmer & Perner, 1983). When younger children are required to explicitly differentiate between an agent's belief that is false and the state of reality (Wellman, Cross, & Watson, 2001), or have to differentiate between knowing and not knowing (Gonzales et al., 2018) they consistently fail in a variety of tasks.

In the realm of FBU several longitudinal studies have identified predictive relationships from early implicit FBU to later explicit performance on ToM tasks, which hint at a developmental relationship between these two types of understanding (Sodian, 2016; Thoermer, Sodian, Vuori, Perst, & Kristen, 2012). Still, the more general question remains how exactly children move from an implicit, preverbal understanding of beliefs and intentions to a subsequent mastery of explicit, verbal ToM tasks.

[1] Please note that several views of implicit ToM are possible. In this thesis implicit is used as a term for both, implicit ToM in the sense of children possessing early, preverbal concepts, and a term for children's production rules which enable them to solve preverbal tasks for ToM correctly.

[2] Discrepancies in implicit false belief research will be addressed in the literature review. However, as investigating the transition between children's preverbal competence and later verbal competence on mental states in general (especially knowledge states) is the objective of this thesis, it remains important independent of whether one argues for or against children's early mastery of implicit false-belief tasks.

One way to address the transition between implicit and explicit, is to investigate the third year of life, which happens to be the time period during which children show mental state language (MSL) and thereby the means to express their thoughts on mental states. MSL might provide the structures necessary to associate preverbal mental presentations with verbal terms, which might then result in a verbal representation of mental states. Around the age of 2, children already use several terms that denote mental states and have no physical referent in the real world such as 'know', 'think' and 'believe'. The way children use these terms accurately in daily conversation has been considered to be a possible sign for their early awareness of mental states i.e. their own ignorance (Harris, Ronfard, & Bartz, 2017), but results on the association between children using mental state terms and having an understanding of the underlying inferences remains inconclusive (Kristen-Antonow et al., 2019; Rubio-Fernández, 2019). By investigating the third year of life and MSL during this time period, it is possible to address the relationship between children's usage of mental state terms and their understanding of what these terms denote. Thus, it might be possible to address whether possessing the verbal terms to denote mental states results in an association between these terms and the mental representations, or whether implicit representations and verbal terms remain independent and only later become associated, if at all. Furthermore, if children already show basic understanding of MSL early on it is important to determine how MSL relates to other abilities during the third year of life that are known to be relevant for early ToM development. If children show understanding of MSL and their MSL relates to early preverbal and subsequent verbal ToM precursors, then that may speak for a continuous development from early preverbal and implicit ToM competence to subsequent competence in standard ToM tasks.

Mental state language is not the only major ability that is considered to be relevant for the emergence of explicit mastery of ToM tasks. Two basic developmental abilities have been argued to be important, namely general language and executive functions (Devine & Hughes, 2014; Milligan, Astington, & Dack, 2007). However, despite research implying these two abilities as crucial parts in ToM development, it is not quite clear how they relate to MSL. Previous research has only identified the impact of inhibitory control for MSL production concurrently (Bellagamba, Laghi, Lonigro, Pace, & Longobardi, 2014), but no large longitudinal studies that include several basic developmental abilities have been conducted as of today. Knowing which abilities are predictive of early MSL enables us to construe interventions for young children at risk for language delay and thereby foster their MSL development even before they begin to use the terms.

In summary, ToM is a skill that appears to be present implicitly in preverbal social interaction, but only manifests itself in explicit ToM tasks around the age of 4. This discrepancy in performance needs to be addressed and one way to do so is by looking at the third year of life, which is the time at which children develop MSL and begin to talk about others' and their own mental states. Investigating the relationship between MSL production and children's understanding of mental state terms allows us to address the question whether children already understand mental state terms in the third year of life, or whether they are only able to infer mental states from behavior without having the respective terms for it i.e. whether they only possess implicit understanding or whether association with mental state terms results in explicit understanding. Furthermore, investigating the concurrent relationship between MSL and additional preverbal and verbal abilities that develop around the same time allows us to address the role that MSL plays for a possible transition between implicit, preverbal ToM and subsequent mastery of explicit, verbal ToM tasks. Finally, precursors of MSL development have not been investigated longitudinally and identifying relevant precursors may play a major role in developing early interventions in order to foster children's MSL development. Overall, the above mentioned aspects allow us to identify which early abilities are necessary for MSL development and what role MSL may play for a transition from implicit, preverbal skills to subsequent mastery of explicit, verbal ToM tasks.

In the following literature review, ToM as a concept and theoretical accounts on its development will be briefly listed and elaborated. Next, the discrepancy between early and late ToM is introduced by addressing the various stages of preverbal, implicit and verbal, explicit ToM and theories on how they might relate to each other. As as discrepancy between implicit and explicit ToM remains, possibilities to bridge the discrepancy will be elaborated, among them a focus on competences that develop in the third year of life, in particular MSL which will be introduced in detail. Finally, the main research questions that remain open will be presented and the literature review is concluded with the research questions and hypotheses of this thesis.

2. LITERATURE REVIEW

The third year of life is a crucial time-period for the development of Theory of Mind (ToM) and its emergence as it constitutes the time in between a preverbal, implicit understanding of others' intentions and knowledge states and an explicit, verbal understanding of false belief and the ability to solve standard ToM tasks. Several theoretical accounts have attempted to explain the emergence of mental state understanding (MSU) and will be the start of this literature review. In order to understand the relationship between preverbal ToM and explicit performance on ToM tasks, it is important to present ToM and its known developmental trajectory as a whole. Thus, ToM and its precursors at different ages will be presented subsequently, followed by current inconsistencies in our understanding of the transition between preverbal and verbal ToM. One way to investigate this transition is by focusing on abilities that develop during this period of time and are important for MSU. Perspective-taking will be briefly introduced before the remaining part of the literature review focuses on the second developing ability, namely mental state language. Mental state language and its relevance for MSU, including experimental evidence will be presented extensively. Finally, the question whether mental state language can be considered sufficient for MSU will be addressed, before the literature review is concluded with an overview of questions that remain to be answered by the present literature and the research questions and hypotheses of this thesis.

2.1 *Theory of Mind and Theories of Theory of Mind*

Imagine a typical exam situation, sitting at a desk with the exam paper in front of you and reading out an open biology exam question that you have to answer, namely

'Is a platypus a mammal?'. If you had been asked the same question outside of the exam situation, let's assume by a friend during a coffe break, your answer might be a different one. Most likely you would have replied with a simple 'yes', as that was what your friend was mostly interested in, without the necessity to elaborate that a platypus lays eggs. During an exam, however, one needs to be aware that the examiner is not necessarily interested in the simple answer to the question, but rather in how much one knows about the topic and is able to show in one's answer. The examiner would most likely expect an elaboration on why a platypus is a mammal and how it differs from a typical mammal, namely by laying eggs. This assessment of what a person might expect and want to hear is very much dependent on what beliefs and intentions we assign to that person. The situation described in this example is not the only time that we infer what another person knows, thinks or believes. Almost every social interaction requires this skill, be it ordering a coffee and spelling out our name in order to help, or reminding our partner to shop for groceries because we assume that they have forgotten.

Premack and Woodruff (1978) were the first ones to name the ability to attribute mental states to another person *Theory of Mind*. In their article named 'Does the chimpanzee have a theory of mind?' they did not believe the chimpanzee to have an actual theory on how the mind works, but rather that it was inferring beliefs and intentions in order to predict the behavior of others. Dennett (1989) has famously termed this process 'taking the intentional stance' as humans tend to not only assign beliefs and intentions to other humans, but sometimes also to animals or even inanimate objects e.g. 'the computer hates me and always turns off'. Nowadays, the label 'theory of mind[1]' encompasses the understanding of mental states in general and revolves around three main categories: beliefs, desires and intentional actions (Wellman, 2018). The main thought is that most behaviors can be explained by people performing particular actions because they believe that performing that action will help them gain something

[1] Throughout this thesis the terms theory of mind and mental state understanding will be used interchangeably.

that they desire. However, when we attribute beliefs and intentions to another person, we also need to be able to attribute a belief that is false and does not correspond to our own knowledge. For example, I need to be able to infer that my partner falsely believes a drink to be inside the fridge, even though I have the knowledge that I drank it the day before and in reality the fridge is empty. It is particularly crucial to be able to understand a false belief, because a true belief can be easily inferred from the state of reality. If there is a drink in the fridge, it is easy to infer that my partner may believe that there is a drink in the fridge. A false belief on the other hand, requires that another person's mental state (i.e. belief about something) is represented independent of the state of reality e.g. in reality the fridge is empty whereas my partner believes a drink to be there. This skill to differentiate another person's false beliefs from reality and one's own beliefs, is considered a crucial step towards the understanding of mental states and a critical test of ToM (Dennett, 1989; Wellman, 2018). If we want to understand the development of ToM, we need to make sure that a child really infers the beliefs and intentions of another person instead of their own and for that comprehension of false-belief (FB) is crucial. Another reason for FBU being considered this important for MSU is that children's ability to understand false beliefs has been related to many positive outcomes later in life, among them social maturity (C. Peterson, Slaughter, Moore, & Wellman, 2016) and popularity among peers (Slaughter et al., 2015). But despite FBU being a crucial step of MSU, it is not the only way in which children gain insight into the their own minds and those of others.

Not only beliefs, desires and intentional actions are important for human behavior, but also perceptions and emotions play a role and one action may result in a subsequent reaction (Wellman, 2018) i.e. our behavior is influenced by what we have seen and how we respond to it. Thus, one might wonder which abilities are encompassed in a 'full-fledged' understanding of mental states. For this purpose, Wellman and Liu (2004) have developed the Theory-of-Mind Scale. The scale includes tasks for the understanding of desires, then beliefs and finally hidden emotions. First of all, a child needs to understand that people can have different desires and these desires can even apply to the same things (Diverse Desires), i.e. 'I enjoy eating a banana, but somebody else detests bananas'. Second, a child needs to understand that people can have different beliefs, even about the same situation (Diverse Beliefs) i.e. 'I believe the fridge to be opened, but somebody else believes it to be closed'. Third, a child needs to understand that something might be true, but another person does not know that (Knowledge-Access) i.e. 'there is a cake in the fridge that I made, but my partner is not aware of that'. Fourth, a child learns that something can be true, but another person falsely believes something different i.e. the example with the empty fridge even though the partner assumes the drink to be present (False-Belief). Lastly, a child learns that someone can feel a particular way, but display a different emotion i.e. smile despite being sad (Hidden Emotions). Several studies on preschoolers have shown that even across country borders, children appear to acquire these skills in the above mentioned order, beginning with an understanding for desires, followed by an understanding for various beliefs and finally hidden emotions (Kristen, Thoermer, Hofer, Aschersleben, & Sodian, 2006; C. C. Peterson, Wellman, & Liu, 2005). An exception to this rule appear to be collectivist societies in which steps two and three are switched during development. Chinese and Iranian preschoolers acquire knowledge-access before they acquire the ability to assign different beliefs to others (Shahaeian, Peterson, Slaughter, & Wellman, 2011; Wellman, Fang, & Peterson, 2011). This is further supported by a study that showed that Chinese parents more commonly use the word 'know', whereas in the U.S. parents predominantly use 'think' (Bartsch & Wellman, 1995; Tardif & Wellman, 2000). When looking at the age at which children reach the various stages of ToM proficiency, it becomes apparent that desires are understood by 3 years of age, diverse beliefs by around 4 years of age and finally that around 4-5 years of age children show the most development in their FBU (Wellman et al., 2001; Wellman & Liu, 2004). Since FBU is considered one of the gold standards to measure ToM, it has therefore been concluded that children are only able to represent others' mental states at the age of four (Wimmer & Perner, 1983). However, there have been several attempts to challenge this view, among them FB tasks with reduced task demands (Scott &

Roby, 2015; Setoh, Scott, & Baillargeon, 2016), but also nonverbal 'implicit' ToM tasks (Grosse Wiesmann, Friederici, Singer, & Steinbeis, 2017; Onishi & Baillargeon, 2005; Thoermer et al., 2012)[1]. Independent of whether one considers these types of FB tasks accurate representations of ToM, theoretical accounts need to be able to explain why children pass some, but not others.

When we think of the challenges of understanding one's own mind in comparison to the minds of others, it becomes apparent that we have an easier access to our own thoughts and desires than to those of another person (German & Leslie, 2001). Thus, it is likely that comprehension of mental states, for as long as they are applied to ourselves, might be easier than those of another person. Several theories have attempted to explain how we are able to understand the mental states of another person, even though they are not easily accessible to us and why children only seem to do so consistently around the age of 4. In the following, I will briefly introduce them before presenting them in detail in the upcoming sections.

As the term 'theory of mind' suggests, one might assume children to act like scientists and develop an actual theory of how the mind works, an account that has been greatly supported and is one of the reasons the term ToM stayed throughout the years (Wellman, 2018). When we think about other people's behavior in terms of mental states, it becomes apparent that all states we usually talk about, namely 'wanting', 'thinking' and 'seeing', are somehow related to each other. People can want something because they have seen it, or they think about something because they want it. All these states appear to have a causal relation to each other and this relation enables us to predict behavior based on them. This observation led some researchers to conceive of the development of ToM as analogue to the development of a scientific theory (see Gopnik & Meltzoff, 1998; Perner, 1991; Wellman, Phillips, & Rodriguez, 2000). This view is commonly known as *Theory Theory* (see 2.1.1), the idea that children explore minds like they would a theory, accumulating evidence and abandoning hypotheses if

[1] Both types of false belief task have received their fair share of criticism which I will elaborate on in 2.2.3.

the evidence speaks against them.

An alternative view to this does not focus on mental states per se, but more specifically on what we would think or do if we were in the situation of another person. The idea is that humans think and behave similarly and that this similarity enables us to predict accurately how a person would act based on our own judgment. Thus, there is no need to establish a ToM as the same processes that enable us to grasp our own mental states and processes can be recycled in order to understand those of another person. This account is called *Simulation Theory* (see 2.1.2) and has been supported by several researchers over the years (Gordon, 1986; Harris, 1992; Heal, 1986). Overall, it argues that there is no need for a real theory as putting ourselves into the shoes of another person enables us to predict their behavior.

The interrelated relationship of mental states and the observation that children seem to acquire them independent of other abilities led some researchers to propose a third view, namely the existence of something called the 'mental module', a structure in the brain that is particularly in charge of processing mental states (see Scholl & Leslie, 1999, for a detailed account on the benefits of modularity theories). In general, modularity theories propose that ToM lacks the structure of a proper theory as it is built up in innate modules that mature throughout infant development. Several different accounts have been proposed including the "Theory-Of-Mind Mechanism" by Leslie, Friedman, and German (2004) and some extreme positions (see Fodor, 1983). Modularity theories are the only ones that have very specific claims on how exactly ToM is represented in the brain.

All of the above theories are considered to be cognitive theories of ToM. Recently there has been criticism from a view called 'socio-constructivism' which postulates that instead of focusing on the individual only, it is important to consider the individual's position in society and the experiences he or she makes during conversation with other individuals (Carpendale & Lewis, 2004). The main idea is that instead of passively acquiring a theory, children are active and engage with the world through social interactions and talk about mental states within the family. For an extensive critique of cognitive theories and the important role that experience seems to play see Fenici (2017).

In summary, ToM is a crucial ability that is relevant for a large number of positive outcomes later in life. It consists of several stages of understanding, ranging from understanding diverse desires, diverse belief and knowledge-access to false-belief, which is considered to be the gold standard and finally hidden emotions. Depending on which tasks are used, even for FBU, children pass tasks at different ages. Theoretical accounts are necessary in order to explain differences in task performance and to address how children are able to understand the minds of others without direct access to them. The theoretical approaches briefly introduced in this section were the Theory Theory, the Simulation Theory, Modularity Theories and the socio-constructivist approach.

In the following I will mostly focus on the Theory Theory and the Simulation Theory while briefly outlining the advantages of a hybrid account that considers social aspects as part of the development of ToM.

2.1.1 Theory Theory

The term 'Theory Theory' (TT) was first used by Morton (1980) when he argued that all of commonsense psychology is build up like a theory. With this view, he did not necessarily argue that commonsense psychology is indeed a scientific theory, but he rather considered it to have some of the main elements that enable us to make inferences. When it comes to ToM, there have been proponents that considered it to be just like a scientific theory (Gopnik & Wellman, 1992, pp.31-33), but mostly ToM is considered to be a theory in the loose sense as described above. In his version of the TT, Perner (1991, 1995) argues that children have a 'representational understanding' of the mind which develops through the ability to metarepresent from a 'situational understanding'. The idea is that during infancy children are unable to represent their immediate environment as they fail several simple tasks like searching for an item that had been hidden and making the A-not-B error (searching at the previous location where the object had been found). Then, around the first year of life, children are capable of building their first representation, namely representing an object that is not visible. Following their first representation they acquire their second one during the second year of life which is concerned with the past, the present, the future and pretend situations. At this stage children know which of their representations are true and which aren't, namely that the ones in the future and pretend representations are only hypothetical. Children are able to do pretend play and understand that the current situation does not correspond to the state of reality, but they do not grasp that their inner representations are also able to differ from reality i.e. one can believe something false similarly to how one pretends that something is true. The crucial point is that even around the second year of life, children do not yet know that their inner representations are *models* of reality. According to Perner (1991), only at the age of 4 children begin to understand that they have mental models and these models can be either true or false. This skill is then termed a metarepresentation and through the knowledge that a model can misrepresent reality, children finally understand that an agent can have a FB. Thus, according to Perners's version of the TT, children do not understand mental states at the age of 2 as they do not have the metarepresentational ability to view them as a model of reality. Only around the age of 4 children acquire a representation of belief that enables them to understand that, similarly to pretend situations, also inner representations are able to misrepresent reality and thus be false.

An updated account of the TT by Wellman (2016) postulates a constructivist view in the sense that learning takes place through integrating what a child already knows and new experiences he or she makes along the way. According to Wellman (2016), there are several empirical findings that support this account. First of all, children appear to learn in orderly conceptual sequences as can be seen in the ToM Scale (Wellman & Liu, 2004). Second, the learning process depends on experiences that children make along the way which is also emphasized by training studies that show improvement with additional input (e.g. Arslan, Verbrugge, Taatgen, & Hollebrandse, 2018). Third, children's learning success depends on their prior knowledge as has been found in a training study by Amsterlaw and Wellman (2006), where children improved on FBU if they had already learned knowledge access. For this type of TT, the crucial aspect is that certain concepts are acquired before others i.e. desire before belief. However, it does not assume that children understand the self before they understand others, as they acquire a concept as a whole through experiences without necessarily requiring introspection.

Most studies that show children's explicit FB reasoning around the age of 4 (Wellman et al., 2001), support the TT account as children should not be able to meta-represent at that stage and therefore not be able to conceive of somebody having a false model of the world. However, the account struggles with findings that show 3-year-old children passing FB tasks with lower processing demands and the view that executive functions are what hinders younger children from showing their mental state understanding (Setoh et al., 2016). One way to argue against the evidence is to postulate that these tasks do not measure genuine FBU as children can use simple behavioral rules to solve them (Rubio-Fernández, Jara-Ettinger, & Gibson, 2017). At this point it is important to note though that currently no tasks are able to differentiate between children having mentalist reasoning processes or simple production rules according to which they solve tasks. Perner et al. (2010) has argued that even if all mentalist vocabulary would be taken out of children's reasoning and it is merely based on production rules, even then these production rules reveal the causalities of the mind to us and therefore have explanatory value. Thus, if production rules show the same flexibility that a representational concept of mental states would, then these production rules may just as well be considered competence. Overall, relations to previously established FB tasks would be necessary to ensure that tasks with reduced demands measure the same underlying concept and that children's abilities to solve them are reproducible.

Another issue that might speak against the conclusions drawn by Wellman (2016) is that children show obvious cultural differences in their order of acquisition of ToM concepts. However, culture does not necessarily seem to be the main factor that brings about these differences (C. C. Peterson & Slaughter, 2016). Rather social situations and conversations are believed to play a crucial role (Ruffman, Slade, & Crowe, 2002), which can be seen as part of the theory-acquisition process that children go through during development. Finally, according to Wellman (2016), children do not need to acquire knowledge of the self before acquiring knowledge of the other, which does not correspond to recent work by Gonzales et al. (2018) who have shown that children seem to solve tasks for their own perspective before they are able to solve it for that of another person.

In summary, two variants of the TT were introduced in detail, the metarepesentational account by Perner (1991, 1995) and the updated, constructivist approach by Wellman (2016). Acquiring explicit ToM around the age of 4 is in line with both TT accounts as they (1) assume metarepresentations to develop and (2) children to acquire the necessary experience to solve explicit task around that age. However, evidence on implicit ToM and children's preverbal understanding of mental states provide objections that can only be answered by either rejecting early competence as a different construct, as an overestimation of children's performance, or by accepting production rules as means of solving these tasks that may or may not need metarepresentations. Possibly children are capable of showing competence through these production rules without acquiring metarepresentations, but the application of these production rules is limited and needs to be explored further.

2.1.2 Simulation Theory

Simulation Theory (ST), opposed to the TT, does not see a necessity in creating a theory of other minds, as there is already a present model that we can use, namely our own mind. Several versions have been proposed, the earliest ones being by Gordon (1986) and Heal (1986). Here I will focus on the version that has been introduced by Harris (1992). The main idea behind this account of ST is that children have a great capacity for pretense and by the age of 2 they are able to engage in pretend play and imagine non-existent objects (Harris, Kavanaugh, Wellman, & Hickling, 1993). By the age of 3 they are even able to imagine having particular mental states, for example being hungry when they are not or being wet because imaginary tea has been spilled over them (Harris, German, & Mills, 1996). Children's ability to imagine mental states is therefore thought to give them the ability to infer those of others. In order to explain how children are able to solve certain tasks, but not others, Harris (1992) proposes that children have something called 'default settings', which are always set to their own mental states, but can be adjusted to understand the mental state of another. Two types of default settings are present, the first one is the current state of reality and the other is the child's current mental state, depending on that reality. The idea is that with an accurate assessment of reality, let's assume the presence of two cookies in a jar, and the child's judgment of his or her desires (e.g. liking the strawberry cookie, but not the ginger cookie), an accurate behavior is produced which could also be generated for another person. Most tasks that children solve before the age of 4 only require changing one default setting. For pretense play, children only need to adjust reality and in order to understand that somebody else dislikes a food they enjoy, they only need to change their mental state default setting. On the other hand, FB tasks require the child to change two default settings as 1) the reality of the agent is different (believing their beloved toy to be somewhere else) and 2) they need to change their own mental state towards the situation. This fits well with research on children's passing age for different FB tasks. It is possible that the reduced executive demands task by Setoh et al. (2016), also requires children to change less default settings and is therefore easier to solve at the age of 3.

Despite explaining current evidence on FB tasks, ST does not specify how exactly children acquire a concept of desires and belief, which are necessary to simulate the desires and beliefs of others. Furthermore, it seems difficult to believe that children are able to solve FB tasks without any need for a theory or simple production rules. When solving an unexpected contents task in which children are shown a smarties tube that harbors pencils instead of smarties, young children struggle to understand that another person that has not looked inside the tube would still expect smarties. In order to solve this task a production rule appears to be necessary, namely that not seeing something equals not knowing about it (Apperly, 2010, pp.187). Furthermore, there seem to be cases where simulation does not work to solve a task. Perner (1991) argues that visual perspective-taking is such a case, as no simulation skills will enable the child to figure out how an object looks from a different side. Rather, it is experience and exploring that makes such an inference possible. Another problematic finding for ST is a study by Ruffman (1996) which showed that children failed to make simple inferences to infer the true belief of a puppet, which corresponded to their own knowledge while correctly inferring the puppet's false belief. Children observed how a sweet was placed inside a box whereas the puppet was only informed about the location. Children did not show difficulties to incorporate the puppet's knowledge and what the puppet had been told. This pattern is opposed to what would be expected according to the ST and Ruffman (1996) argues the findings can be better explained by over-application of rules than simulation.

In summary, the ST approach proposed by Harris (1992) was introduced which proposes that children possess two default settings, one for the state of reality and another for their mental state. Overall, children fail to solve FB tasks, because they need to adjust two default settings which they are not capable of doing before the age of 4. However, the ST struggles with explaining cases in which children struggle to infer the truth belief of a person despite having the resources ot adjust their default settings.

2.1.3 Social-Constructivist View

Carpendale and Lewis (2004) argue that both TT and ST consider introspection and theory construction to be something that happens within the individual, mostly without much interaction with the outside[2]. However, socio-constructivists argue that children's concepts are strongly influenced by their communication with others and the input they receive from their caregivers. Triadic interactions with an object and a caregiver are crucial for children to learn about intentions and dyadic interactions

[2] Note that the updated TT approach by Wellman (2016) considers experience a major contributing factor in the acquisition of ToM.

with the object alone do not have the same effect (Brandone, Stout, & Moty, 2019). Furthermore, Harris, de Rosnay, and Pons (2005) have argued that language and conversation in general appears to play a big role in children's acquisition of ToM. For example children with advanced language skills tend to be better at understanding mental states (Pons, Lawson, Harris, & De Rosnay, 2003) and exposure to maternal language that has an abundance of mental state terms also improves children's abilities (Ruffman et al., 2002). But, opposed to arguing that achieving a better vocabulary or acquiring syntactic tools is crucial for ToM, Harris et al. (2005) argues that it is conversations about different minds and perspectives that aid children in understanding ToM the most.

Fenici (2017) proposes that social interaction is particularly relevant for the acquisition of mental state understanding and especially for passing a traditional FB task, because children and adults that are deprived of it never succeed in FB reasoning. The main thought is that despite children having a basic capacity to reason about an agent's intentions early on and predict behavior or attribute knowledge states, the ability to talk about another agent's beliefs and to reason about his or her behavior is something that only develops in conversation with another person. Only through conversation with their caregivers children learn about another person's explicit beliefs and the actions those beliefs result in.

In summary, it appears that both ST and TT need to take children's social environments and the advances they make through interactions with their caregivers into account.

2.1.4 Theory Theory vs. Simulation Theory

As current evidence appears to point both towards ST in some cases and TT in other cases, there are several researchers that adopt an extended hybrid view between the two (e.g. Nichols & Stich, 2003). This is not particularly surprising as both theories have areas in which one provides predictions and the other does not. In the following I will briefly address main commonalities and differences while pointing out which questions would remain for a hybrid account to answer.

Two main aspects for which the TT and ST differ are how these theories explain children's performance on explicit ToM tasks and simplified or 'implicit' ToM tasks. While the TT explains mastery of explicit ToM tasks by the acquisition of metarepresentations around the age of 4, the ST explains it by children's ability to switch both of their default settings. Implicit ToM mastery on the other hand, the TT could explain through production rules or primary representations which are consistently applied, but do not require metarepresentations, whereas the ST assumes that implicit ToM tasks only require the adjustment of a single default setting, which children are already capable of as the concepts are present. The crucial difference here is that while the TT assumes children to lack a concept of mental states early on, the ST argues that children are merely unable to show it yet. Furthermore, while the TT predicts desire to be acquired before belief, the ST only predicts understanding of the self to be acquired before the other, as the self is used as a model for the perspective of the other (See Table 2.1 for an overview). Here it is important to note that since both theories only provide predictions in one area, but not the other, a synthesis would be possible. It is feasible that children acquire a verbal concept of desire before belief and acquire the verbal concept for themselves before they are capable of applying it to others. This is in line with recent research by Gonzales et al. (2018) who investigated verbal visual and epistemic perspective-taking in 2- to 4-year-old children and found evidence for a hybrid approach. Introspection matters as the self condition of both visual and epistemic perspective-taking was easier than the other condition, however, at the same time visual perspective-taking (VPT) was easier than epistemic perspective-taking (EPT). This suggest an interaction of the two accounts as perception was acquired before knowledge (showing an order in concepts) while the self was understood before the other.

Apart from predictions on the acquisition of self before other, the ST does not define how exactly children acquire an understanding of mental states and similarly to Wellman (2016), one might postulate that experience plays a crucial role. In his version

Tab. 2.1: Overview of predictions for the Theory Theory, the Simulation Theory and a Hybrid Approach

Phenomena	Theory Theory	Simulation Theory	Hybrid Theory
Explicit ToM at 4	Children only acquire metarepresentations around the age of 4	Explicit ToM tasks require children to change two default settings, which they are only capable of at the age of 4	Children are either capable of metarepresentations OR have acquired the ability to change both default settings
'Implicit' ToM before 4	Children use production rules to solve the tasks, but still lack the required metarepresentations	Children possess the concepts and only need to adjust one default setting, therefore they are able to pass the tasks	Either children possess the concepts and are able to show it OR children us production rules to solve the tasks
Acquisition of Desire and Belief	Desire is acquired before belief	Makes no predictions about the order of acquisition	Desire is acquired before belief
Acquisition of Self and Other	Makes no predictions about the order of acquisition	Self is acquired before other as the self serves as a model for the other	Self is acquired before the other

of the ST, Harris et al. (2005) argues that language and conversation are crucial for the acquisition of concepts. Thus, it is feasible that children construct their concepts of mental states through accumulating evidence and in conversation with their caregivers and that their experiences with desire before belief and the vividness of their own perspective results in the acquisition pattern observed by Gonzales et al. (2018).

If children reason about themselves before they do about others, one might assume that if children make a mistake in the application of a learned concept, then that mistake should be egocentric. However, Gonzales et al. (2018) have found that if children made a mistake in reporting another's perspective, they did not consistently report their own perspective instead. Thus, it appears that not a focus on the self keeps children from reasoning about the perspective of somebody else. According to the ST, two types of mistakes are possible. Either children show an egocentric bias in applying their own perspective to another person (which appears not to be the case) or children have an incomplete or false concept of their own mental states that they attempt to apply to others. The TT, however, would be able to explain such mistakes by the misapplication of production rules as children do not possess working concepts of mental states yet.

A main question that remains to be answered when creating a hybrid approach between the ST and the TT, is whether one assumes children to have concepts of mental states but be unable to show it (ST: default settings cannot be adjusted well enough) or whether one assumes children to only have production rules, but no concepts yet (TT: primary representations or productions rules, but no metarepresentations yet). Both, concepts and production rules/primary representations might be acquired through observing behavior and statistical learning (Ruffman & Taumoepeau, 2016), but which of the two it is remains to be investigated.

As I have mentioned previously, Perner et al. (2010) pointed out the problem of differentiating between a mentalist approach in which children reason according to a ToM and an approach that is based on production rules which might result in the same answers to a task. However, if it is indeed true that research is unable to differentiate between children solving tasks through production rules or children assigning mental states, then there is no advantage of making that distinction as both system are functionally equivalent and show the same performance. It appears that in the area of preverbal tasks a distinction between a mentalist approach and production rules is difficult as inferences that children are able to draw might be explained by production rules in the same way that they are explained by a concept of mental states. Even the failure to show children's competence in preverbal tasks might either mean that they did not establish the production rules yet or have not developed the concepts yet.

However, in the transition from preverbal competence to verbal competence it might be possible to tease production rules and conceptual understanding apart. If children's preverbal competence was based on a concept of mental states that only requires verbalization, one might assume that the acquisition of language should be sufficient to enable such verbalization. But if children's preverbal competence is mostly built up on production rules, it seems more feasible that these production rules would require to be verbalized separately as they are not part of an integrated concept that is merely associated with a word i.e. a concept of knowledge that is associated with the word 'know'. See Figure 2.1 for a visualization of this relationship.

In summary, it is possible to create a hybrid account between the TT and the ST which proposes that concepts of the self are acquired before concepts of the other and that the concept of desire is acquired before the concept of belief. However, a hybrid account would have to determine whether children possess concepts early on (ST), or whether children create production rules (TT). Preverbally the difference is difficult to determine, but during the transition from preverbal to verbal it should be possible to address as associations between language and a concept should be quicker and easier, than associations between several independent productions rules and language.

Thus, a study that examines the transitional period from preverbal to verbal competence and examines whether children are able to draw the right inferences from mental state terms, might be able to determine whether children draw inferences based on production rules or a conceptual understanding of mental states. Finally, it is very well possible that children develop a metarepresentational understanding of mental states through an accumulation of production rules which mirror the behavior of conceptual understanding.

2.2 *Theory Of Mind Throughout Early Development*

In the following sections, the earliest signs of ToM will be introduced, first for infancy in the form of precursors and during nonverbal communication and later for early childhood including the discrepancy between implicit and explicit FB tasks.

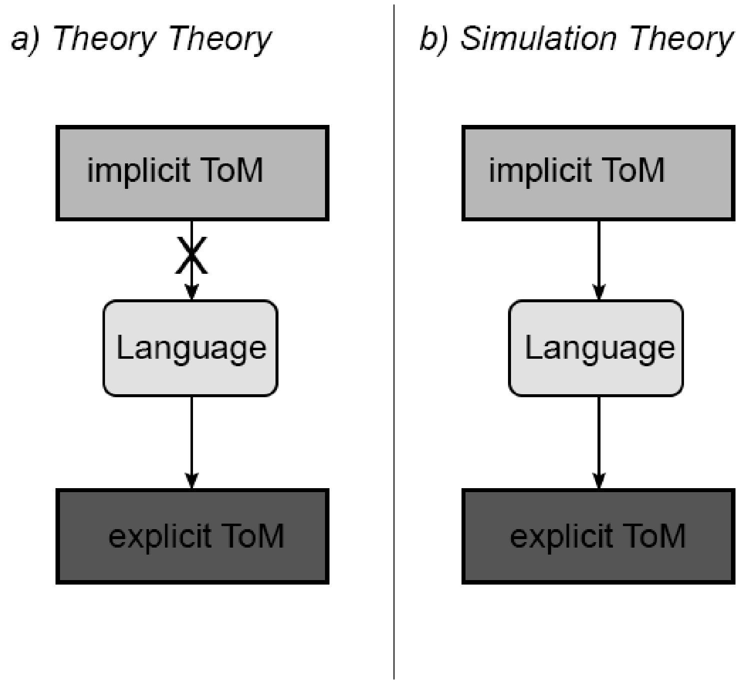

Fig. 2.1: Overview of predictions of the Theory Theory and the Simulation Theory in regard to the relationship between implicit and explicit ToM. *a)* As the TT considers implicit ToM to be mostly production rules (or primary representations), these production rules should not be able to structure language acquisition. However, upon the acquisition of language, language aids children in developing a metarepresentational understanding of mental states. *b)* As the ST assumes children to possess concepts of mental states early on, but not have the means to show it, these concepts should aid the acquisition of language and the association of language with these concepts, which then in turn influences the verbalization of these concepts in explicit tasks.

2.2.1 Theory of Mind in Infancy

Human infants show a preference for social stimuli very early on in their lives. Before being able to assign beliefs and intentions to another human, it is important to identify them as social partners, be it by recognizing biological motion at birth and also later (Sifre et al., 2018; Simion, Regolin, & Bulf, 2008) or showing a preference for human faces over that of a doll (Legerstee, Pomerleau, Malcuit, & Feider, 1987). Already 1-month-old infants show the ability to differentiate between the faces of human infants and monkey infants and at 3 months of age they even show a preference for human faces (Sanefuji, Wada, Yamamoto, Mohri, & Taniike, 2014). The same preference has been found with 3,5-month-old infants and both, human adult faces and bodies over non-human primate faces and bodies (Heron-Delaney, Wirth, & Pascalis, 2011).

Looking at faces is not sufficient to understand what is going on inside another person's head. For that it is important to be aware of what a person sees or does not see, what they have awareness of and what not. Meltzoff and Brooks (2007) argued that before 9 months of age, children mostly pay attention to the position of an adult's head and whether they are turned towards them or away from them, but do not have the ability to follow the adult's gaze towards a new object. However, a study by Senju and Csibra (2008) showed that infants at the age of 6 months were able to follow an adult's gaze, but only if it was preceded by direct gaze at the infant or infant-directed speech. *Gaze-following* is one of the earliest precursors considered to be important for later ToM abilities (Brooks & Meltzoff, 2015).

Another major ability that has often been termed early ToM is *joint attention.* Joint attention is based on the coordination of attention between oneself, another person and some external object or event (Tomasello, 1995). This ability follows from gaze-following because identifying where another person is looking enables the child to direct attention towards the same object. This is in line with a finding by Senju and Csibra (2008) that adults had to establish eye-contact first before gaze-following was possible. Joint attention is important for establishing a common ground for social interaction and social learning, because adults are able to focus the infant's attention on relevant objects and events. Brandone et al. (2019) showed that there is a connection between early interactions with a caregiver and an object and later understanding of goal-directed behavior. In a longitudinal study, the authors found that children who had more triadic engagements at the age of 6 months (i.e. interactions with a caregiver and an object) were better at reasoning about the goals of reaching actions at the age of 9 months. This emphasizes the relationship between joint attention and intention understanding. Two types of joint attention can be differentiated, declarative joint attention and imperative joint attention. In declarative joint attention, the goal is to convey information about and an opinion on a particular object or event, whereas in imperative joint attention the goal is to bring about a change in the communicative partner, be it a change in opinion or a change in behavior to reach a desired object. In a longitudinal study by Sodian and Kristen-Antonow (2015), only declarative joint attention at the age of 12 months predicted later FBU at the age of 50 months. One hypothesis is that declarative joint attention is particularly important for ToM development as the goal is to convey information and thereby change the knowledge state of another person (Camaioni, Perucchini, Bellagamba, & Colonnesi, 2004).

As has been mentioned above, joint attention might be the first step towards infants learning that actions are performed in order to reach a particular goal (Brandone et al., 2019). Understanding goal-directed behavior or *goal encoding* is a crucial step towards intention understanding, since connecting an agent's action with a goal makes it easier to identify the purpose of the action and thereby the intention of the agent. In the first year of life, infants begin to show a rudimentary understanding of goal-directed behavior. In a study by Cannon and Woodward (2012), 11-month-old infants were presented with an eye-tracking task during which a hand consistently grabbed one out of two objects. Following the familiarization for the goal object, the position of objects was swapped and infants only observed the incomplete grasping motion. Infants consistently looked at the object that was grasped before, independent of its location change. A control that repeated the experiment with a claw instead of a human hand, showed that infants looked consistently at the location instead of the object, indicating they only associated the human hand with a goal to reach the object and not an artificial claw. A study by Woodward (1998) presented evidence for goal-encoding earlier. In a visual habituation paradigm, 6-month-olds observed an agent grasp for one out of two toys. After infants were habituated to the grasping action, the toys were switched in place and an agent either reached out towards the previous toy or towards the previous location. Infants were shown to look longer when the agent grasped towards a new toy and did not show that preference if the agent was an inanimate claw. The same pattern but weaker was also identified in 5-month-olds in an additional experiment. Thus, children seem to recognize human actions as goal-directed early on in their first year of life.

Interpreting human actions as goal-directed has been found to be a precursor of ToM by several longitudinal studies (Aschersleben, Hofer, & Jovanovic, 2008; Wellman et al., 2004). Wellman et al. (2004) have investigated the relationship between 14-month-olds goal-encoding in a habituation paradigm similar to Woodward (1998) and their subsequent ToM understanding at 4 years of age. Results indicated a relationship between children's attention towards intentional actions during the habituation phase (with an agent reaching towards specific object) and their later ToM performance, but not during the test phase (with an agent reaching towards the specific toy or the location). A later study by Wellman, Lopez-Duran, LaBounty, and Hamilton (2008) looked at whether the relationship between early attention towards intentional actions and later ToM could be mediated by other factors like language, IQ and executive functions. However, the relationship remained even when controlling for these factors. Aschersleben et al. (2008) investigated this relation earlier and looked at the relationship between goal-encoding at 6 months in a habituation paradigm and ToM performance on the ToM Scale around 4 years of age. Similarly to Wellman et al. (2004), the authors found a correlation between infants' attention towards intentional grasping actions at 6 months and their subsequent FB performance at 4 years. This relationship was independent of general language abilities.

In summary, this section has introduced gaze-following, declarative joint attention and goal-encoding as important abilities for the subsequent development of ToM. The presence of longitudinal relationships between these early abilities and later ToM performance suggests a continuous development to an understanding of mental states, among them FB. However, before children understand FBs, they interact with their caregivers, initially in a non-verbal manner. In the following section, I will present current research that suggests children already show the ability to track desires and knowledge states in their nonverbal communication with their caregivers.

2.2.2 Theory of Mind in Preverbal Communication

In the second year of life, children begin to show sensitivity for desires and epistemic states based on situational cues, such as a person's facial expressions or their presence or absence during an event. These skills are relevant for later ToM development as they show that children track information that is important for mental state understanding. Repacholi and Gopnik (1997) found that 18-month-old infants respond non-egocentrically to another person's expression of desire for a certain kind of food that contrasts with their own desire, a finding that was replicated only in 24-month-olds in subsequent research (Carlson, Mandell, & Williams, 2004; Sodian et al., 2016). By the age of 2,5 years, children are able to predict another person's action based on information about their desire. Wellman and Woolley (1990) presented children with stories in which a character looked for a particular item and either found it, didn't find anything or found a similarly desirable object. Two and a half to 3-year-old children were able to correctly identify whether the character would continue searching or not, showing the ability to associate a person's desire with their action.

Similarly, two-year-olds appear to take others' knowledge states into account in communication. O'Neill (1996) showed children a new toy which was subsequently placed into one of two containers on a high shelf, which was out of reach. The mothers were either aware of the event (because they watched it) or not (because they left the room or closed their eyes). Children had to ask their mother for help in order to retrieve the toy. Two-year-olds named the toy, gestured towards its location or named its location significantly more often when the mother was ignorant of the location than when she was knowledgeable, thus indicating sensitivity for the mother's knowledge state.

More recent studies have suggested that even in the second year of life, infants know what others have been engaged with in the recent past and draw adequate inferences about others' need for information. In a study by Moll et al. (2007), an adult engaged in joint visual attention with 14-month-old infants around two novel toys in turn for one minute each. The adult then left the room. During their absence, the infant explored a third novel object together with an assistant. Finally, the adult returned, looked at a tray containing all three objects, expressed excitement, and then made an ambiguous request for the infant to hand 'it' to them. The result was that infants handed the adult the object that she did not know from past experience more often than the familiar object. Similarly, 12- to 18-month-olds have been shown to selectively point to the location of an object the experimenter did not know about (Behne, Liszkowski, Carpenter, & Tomasello, 2012; Liszkowski, Carpenter, & Tomasello, 2007, 2008). These studies indicate pragmatically adequate informing behaviors in communication in young children, taking another person's prior knowledge and likely intentions into account.

In summary, pointing and gesturing towards objects that an interlocuter does not see, is a first sign for an understanding of different perspectives and that infants are generally interested in the inner lives of other people. Mastering gaze-following, joint attention and goal-encoding in infancy enables children to draw correct inferences about other people's knowledge states, even before they are able to express themselves verbally. In the next section, current research on the transition from a preverbal understanding of mental states to a verbal, explicit understanding will be elaborated.

2.2.3 Theory of Mind in Early Childhood

In toddlerhood and early childhood, the emergence of language enables children to understand more complex intentions and beliefs (Bartsch & Wellman, 1995, pp.17-20). However, despite children drawing correct inferences in infancy, it is difficult to identify when exactly they move from their preverbal understanding to a verbal understanding. Even though 2-year-old children are already able to infer the knowledge state of another person during communication (O'Neill, 1996), they are not able to answer correctly when explicitly asked about that person's knowledge state (Wellman et al., 2001). One way to differentiate between early MSU and later MSU has been to distinguish between implicit and explicit ToM. Implicit ToM is defined as children having preverbal concepts of mental states that they are able to act on and which have not been associated with the respective words yet. Explicit ToM on the other hand, is when implicit concepts have been associated with the words that denote them (i.e. 'know' with the concept of knowledge or 'believe' with the concept of belief) and children are able to use them accordingly in verbal tasks. It is also possible to adopt a view of implicit ToM that is not build on the presence of a concept, but rather on a collection of production rules that helps children to draw pragmatically adequate inferences[3]. Studies in which 3-year-olds looked towards the correct location in a FB task, but still answered wrongly when asked about a character's mental state directly, suggest that young children might not yet be aware of their insight into human behavior (Clements & Perner, 1994; Low, 2010)[4]. Children's lack of awareness in early childhood when it comes to reporting their own knowledge states has been shown by several studies (Hembacher & Ghetti, 2014; Kim, Paulus, Sodian, & Proust, 2016). Kim et al. (2016) found that 3- to 4-year-olds tend to overestimate their own knowledge when asked to report on it verbally, but still choose not to inform an adult when their knowledge is limited. This supports the idea that there is a discrepancy between reporting answers verbally in FB tasks or reporting

[3] Consider the distinction introduced in 2.0.1 which pointed out the focus of ST on early concepts and TT on production rules or early representations.

[4] According to Perner (1991) one might claim that they lack a metarepresentational understanding of mental states.

one's own knowledge state, and being able to make correct inferences from others' and one's own knowledge states. This discrepancy has been attributed to the emergence of implicit FBU, suggesting that children before the age of 4 have a concept of FB that has not been verbalized yet and only later when children become aware of their own concepts and are able to express them verbally, they are also able to pass explicit FB tasks.

In the following section, I shall further elaborate on the discrepancy between implicit and explicit FBU and why research on the time period between the acquisition of the two is crucial to answer remaining questions.

Implicit vs. Explicit False-Belief Understanding and Theories on their Relation

FBU is considered to be the critical test for ToM and its traditional version is only acquired around the age of 4 (Wellman & Liu, 2004; Wimmer & Perner, 1983). What makes FBU much more complex than simply storing facts about what an agent knows or thinks, is that a mental state is not only something that concerns the agent, but rather the agent's relation to something else, be it the world, another object or a situation. Just because a person is able to predict another person's behavior, we cannot conclude that they are able to understand other people's mental states. Predicting behavior can also be done by observation and being aware of the state of reality alone without a need for recognition of mental states (Dennett, 1989, pp.16-29). Testing children on FBU enables us to make sure that they can represent beliefs as they are able to understand that another person can have a false one which does not correspond to reality. Simply keeping in mind a false fact about another person is not sufficient for FBU (Apperly, 2010, pp.77-80). For example, Leslie and Polizzi (1998) have argued that passing a FB task requires the child to overcome the default assumption that beliefs are always true. In order to accept a belief as false, the general assumption that they are true needs to be inhibited, which adds another skill that is necessary in order to master FBs. Perner, Leekam, and Wimmer (1987) argued that children need to understand that a single proposition can have conflicting truth values, i.e. be true for a person with a false belief, but false for the child themselves. Overall, there seem to be additional abilities that children need to master in order to be able to understand another person's false beliefs. The question is whether FBU can be tested in a way that does not require the above mentioned abilities in order to determine competence.

Harris et al. (2005) used the example of Little Red Riding Hood to explain the different stages at which children show an understanding of other's false beliefs. Around 3 years of age, children do not understand that the little red riding hood expects to see her grandmother when she knocks on the door of the cottage. However, 4- and 5-year-olds are able to understand her expectation and her false belief. Interestingly, they do not understand the associated emotions as they expect the little girl to be afraid while knocking as the wolf awaits her, even though they themselves said that she expects to see her grandmother. Only around 6 years of age children grasp the whole story. This example illustrates that even when children grasp FB, they are not yet capable of adult-like mental state understanding. This is in accordance with the assessment of the ToM Scale by Wellman and Liu (2004).

The standard test for FBU is easier than the fairy tale and usually involves two individuals, one of which (Maxi) hides a beloved item and leaves the scene (Wimmer & Perner, 1983). The second individual then places the beloved item into a different location before also leaving the scene. The critical test occurs upon Maxi's return when he has to look for his beloved item. Children are usually either asked about Maxi's mental state, e.g. 'What does Maxi think?' or about Maxi's subsequent action e.g. 'Where is Maxi going to look for his beloved item?' As children consistently fail these tasks before the age of 4 (Wellman et al., 2001), several authors have argued that children only achieve a fundamental understanding of mental states around that time (Astington & Gopnik, 1991; Perner, 1991; Wellman & Liu, 2004). However, the standard FB task requires children to do much more than only represent an agent's false belief. Additionally, children are required to follow the story presented to them with puppets, to remember the location of the beloved item, to keep track of individuals leaving the scene and finally to be able to reply verbally to the posed questions (Bloom

& German, 2000). Since there are many more skills than only FBU necessary to pass the standard FB task, the established view that children only acquire FBU around the age of 4 has been challenged by several studies conducted over the course of the past 15 years which examined implicit FBU (Baillargeon et al., 2010; Grosse Wiesmann et al., 2017; Knudsen & Liszkowski, 2012; Southgate et al., 2007).

In implicit FB studies, children show that during their second and third year of life they seem to be able to nonverbally predict an agent's action based on his or her FB. This earlier way to reason about FB has been termed 'implicit FBU' or 'implicit ToM'. There are several types of implicit FB tasks in which infants younger than 2 years of age are able to predict how an agent with a false belief will behave (Baillargeon et al., 2010; Onishi & Baillargeon, 2005; Sodian, 2016; Southgate et al., 2007), but in this literature review I will mainly focus on anticipatory looking paradigms. The main idea of anticipatory looking paradigms is that children observe an agent behave, but just before the resolution of an action the action pauses. Children are expected to anticipate what an agent will do next, which hints at their action predictions. If children anticipate the behavior of an agent with a false belief correctly, one might conclude that the child has understood FB and is able to infer actions based on it. There are critical views on this conclusion which will be introduced later in this section.

One of the first anticipatory looking paradigms for implicit FBU was introduced by Southgate et al. (2007). In this eye tracking study, 25-month-olds observed an agent sitting in front of a wall with two closed windows, positioned on the left and the right respectively. A box was placed in front of each window, similar to a traditional FB task. Children were familiarized with the pattern that the agent was interested in an object that was placed into one of the boxes by a puppet. The main action that was anticipated in this task was the grasping motion of the agent through one of the two windows. Before the agent reached through one of the windows towards a box, a bell sounded and children's anticipatory gaze was measured. Two types of FB were tested, which were identical to each other with the exception that FB1 allowed an alternative strategy to solve the task (e.g. looking at the last location of the object), whereas FB2 did not. FB1 allowed an additional strategy as the agent observed one transfer of the object before being distracted by the ringing of a telephone, thus observing the object in its last location before it disappeared from the scene, whereas in FB2 the agent was instantly distracted by the ringing of the telephone after the object was placed inside a box and did not observe the last location before the object disappeared from the scene. Children anticipated the grasping motion of the agent correctly for both of the FB conditions, suggesting that they did not use an alternative strategy to solve the task.

However, despite several recent studies suggesting the existence of implicit FBU before the age of 4, there are two ways in which this conclusion has been questioned over the past years. On the one hand, implicit FB may not be a sign for children acquiring a concept for FB, but rather applying behavioral rules that help them show the expected performance without actual understanding (Ruffman, 2014). But not only a theoretical concern has been voiced as a problem. There has also been a recent debate on the replicability of these findings (He, Bolz, & Baillargeon, 2012; Kulke, Johannsen, & Rakoczy, 2019; Kulke, Reiß, Krist, & Rakoczy, 2018; Kulke, von Duhn, Schneider, & Rakoczy, 2018; Phillips et al., 2015; Poulin-Dubois et al., 2018). Not many studies have investigated the stability of implicit FBU within and across paradigms and most have focused on children around the age of 2 years. Grosse Wiesmann, Friederici, Disla, Steinbeis, and Singer (2018) investigated children's implicit FBU in a longitudinal study from the age of 2 to 4. The study used the anticipatory looking paradigm by Southgate et al. (2007) with additional trials. In contrast to earlier findings with the same paradigm, children only performed significantly above chance around the age of 4, but not before. Furthermore, contrary to the original study showing children performing similarly on FB1 and FB2 conditions (Southgate et al., 2007), there was a discrepancy between the two types of FB with even 4-year-olds only performing above chance in FB1. This is partially in line with a recent study that has also found a discrepancy between the two types of FB (Kulke, von Duhn, et al., 2018), with children performing significantly above chance in FB1 and additionally significantly below chance on FB2. One possible explanation is that in FB2, children have to remember the agent's belief for a longer period of time while keeping in mind the real location of the object (which might have resulted in significant working memory load). Another alternative explanation is that children were mostly focusing on the last location of the object in FB1, which corresponded to the FB location of the agent. Grosse Wiesmann et al. (2018) argue that such a strategy should have resulted in below chance performance in FB2 which was not the case in their study, but applicable in the replication attempt by Kulke, von Duhn, et al. (2018). Thus, the findings on anticipatory looking tasks for implicit FB appear to be inconclusive when it comes to children's performance on the FB1 and FB2 conditions. It is important to note that the same concerns on replicability have been voiced for other types of implicit FB tasks, among them violation-of-expectation tasks (VOA) (Powell, Hobbs, Bardis, Carey, & Saxe, 2018), action-based tasks (Kammermeier & Paulus, 2018), interaction-based and helping tasks (Crivello & Poulin-Dubois, 2018; Dörrenberg, Rakoczy, & Liszkowski, 2018; Priewasser, Rafetseder, Gargitter, & Perner, 2018) and do not only apply to anticipatory looking paradigms.

Overall, results on implicit false belief appear inconsistent and it is difficult to conclude whether children are genuinely able to infer others' false beliefs implicitly before the age of 4. Early anticipatory looking paradigms hint at a sensitivity for FB (Southgate et al., 2007), however, replications for these experiments fail more often than not and later longitudinal work with the same task suggests competence to develop around 4 years, similarly to explicit FBU (Grosse Wiesmann et al., 2018). Thus, task performance alone makes it difficult to argue for the presence of implicit FBU, however longitudinal relations between implicit and explicit tasks paint a slightly different picture.

Older work by Clements and Perner (1994) identified that implicit FBU precedes explicit FBU by about half a year. It might therefore be necessary to investigate implicit and explicit performance more frequently throughout development to identify whether implicit does indeed precede explicit performance or whether both develop around the same time as suggested by Grosse Wiesmann et al. (2018). Especially the tasks chosen might play a role as in order to have reliable longitudinal associations children will need to show replicable competence on an implicit task[5].

At this point it is important to note that even without early understanding of FB, there are hints towards infants understanding mental states (e.g. knowledge states) in interaction which can be considered early signs of ToM as has been mentioned in previous sections. For example joint attention (a precursor) and children's preverbal understanding of seeing and knowing is correlated with later ToM performance, while children's ability to perform implicit FBU remains debated. Therefore, independent of the existence of implicit FBU, the question remains how preverbal mental state understanding relates to later explicit ToM performance.

Given the presence of early preverbal understanding of knowledge states, one might ask how and whether preverbal understanding and later explicit understanding are related. Is it the same mental state understanding that is measured with different methods? In the realm of FBU, there are several accounts that attempt to explain the relation between the two types of FBU which can also be applied to the relationship between preverbal understanding of knowledge states and subsequent explicit understanding. All approaches are mostly concerned with whether there is a continuous transition from implicit to explicit or whether the two abilities are utterly independent (Sodian, 2016). *Continuity accounts* argue that implicit FB tasks already measure infants' ability to represent mental states, but that certain task demands make showing these abilities in explicit tasks impossible (Baillargeon et al., 2010). The idea is that certain basic abilities are necessary to solve these tasks, among them general language abilities, inhibitory control and working memory and that implicit tasks do not require these prerequisites and are therefore able to show infants' true FB skills (Helming, Strickland, & Jacob, 2014). Opposed to continuity accounts, the *dual systems account* (or two systems account) argues for implicit and explicit ToM being independent sys-

[5] It is also important to note that longitudinal relations between implicit and explicit do not require a rich interpretation of implicit FB. Also implicit FB in form of behavioral rules might relate to children's explicit false belief.

tems for tracking beliefs (Apperly & Butterfill, 2009). Implicit tasks are solved by an early and quick system 1, which is in charge of tracking belief-like states and enables children to perform level-1 perspective-taking, but not level-2 perspective-taking (which only comes with the later development of system 2). System 2 is responsible for explicit FBU and is therefore cognitively more demanding, but also more flexible as it can adapt to situations better (see also Sodian, 2011). A similar alternative was introduced by Perner and Roessler (2012) who argued that children are able to use system 1 in order to form an 'experiential record' of what an agent has seen or experienced. The dual systems account postulates that since system 1 enables children to pass level-1 perspective-taking tasks, there should be a relation between early perspective-taking skills and implicit FB performance.

Another perspective that is often termed minimalism or *low-level accounts* was proposed by Ruffman (2014), who argues that instead of infants understanding mental states as has been claimed by Baillargeon et al. (2010), they instead only understand behavior. Children have some innate qualities that enable them to identify behavioral patterns and therefore to predict an agent's actions based on these behavioral patterns. According to Ruffman (2014) infants have innate abilities early on that enable them to identify these patterns. Some of these abilities have already been mentioned in the beginning of this literature review, namely infant's preference for faces and their perception of biological motion. An additional ability is statistical learning which enables infants to make rules about behavior and apply them. Another type of low-level account was proposed by Heyes (2014) who argued that children's performance on implicit FB tasks can be explained by domain-general processes, among them perceptual and imaginal novelty of the test stimuli.

One way to investigate how implicit FBU relates to ToM is to look at cross-sectional and longitudinal relations to explicit FBU. Identifying correlational relationships between early and later FBU suggests that both skills have a common source of variance. This does not necessarily allow us to differentiate between the continuity account or the dual systems account, but it helps to argue against extreme low-level accounts (e.g. Heyes, 2014), since domain-general processes should not predict children's later explicit performance. Longitudinal work showed a predictive relation between early implicit FBU and later explicit FBU (Sodian, 2016; Thoermer et al., 2012). However, cross-sectional studies have mixed results on the relationship between implicit and explicit understanding (Grosse Wiesmann et al., 2017; Low, 2010; Ruffman, Garnham, Import, & Connolly, 2001). Furthermore, not all implicit FB tasks seem to be measuring the same concept, as several studies that have tested for inter-task relations between different types of implicit FB tasks have failed to find significant correlations (Dörrenberg et al., 2018; Poulin-Dubois & Yott, 2018; Yott & Poulin-Dubois, 2016). Thus, it may be very well possible that relationships between implicit FBU and later explicit FBU are strongly dependent on the task chosen. The only study up to date that has investigated implicit and explicit FBU simultaneously in 3- and 4-year-olds is a study by Grosse Wiesmann et al. (2017). All children were tested on an anticipatory looking task for their implicit FBU and on standard explicit FB tasks developed by Wimmer and Perner (1983). The implicit FB task involved a mouse playing hide and seek with different animals and escaping through a tunnel with two exists. Each exit had an associated box and the mouse hid in one of the two boxes. Apart from the scenario, both FB1 and FB2 conditions were similar to the type described above with the exception that instead of turning away, the agent left the scene. The results showed that 3-year-olds performed significantly below chance in one of the explicit tasks whereas 4-year-olds performed significantly above chance, showing the expected development from 3 to 4 years of age. In contrast, there was no difference between 3- and 4-year-olds on the implicit measure as both groups performed significantly above chance (one-tailed for the 3-year-olds). Furthermore, there was no significant correlation between the two types of FBU, suggesting that implicit FB is an independent way to reason about an agent's behavior than explicit FB. Interestingly, Grosse Wiesmann et al. (2017) found competence in implicit false belief in 3-year-olds whereas Grosse Wiesmann et al. (2018) only found competence in the FB1 condition in 4-year-olds. The difference in children's performance on the anticipatory looking tasks of these two studies might be explained by aspects of the paradigms. One of these aspects might be salience of the object which was constantly located in front of the wall in the paradigm in Grosse Wiesmann et al. (2018), which might have made it difficult to focus on the agent. Focusing on the agent has been argued to be one of the crucial factors that enable children to pass a FB task, be it implicit or explicit (Rubio-Fernández & Geurts, 2013, 2016). However, in the task by Grosse Wiesmann et al. (2017), the agent left the scene for an extended period which might have also made it difficult to focus on his beliefs, but still 3-year-olds were able to pass the task. Thus, it is difficult to explain the discrepancy between the two tasks with agent-salience alone. Overall, only longitudinal studies appear to identify a relationship between implicit and explicit FBU whereas concurrently there is no evidence for both skills measuring a similar underlying concept. Thus, concurrent evidence might either be taken as a sign of two unrelated systems (dual systems account), a sign for unreliable implicit FB tasks or a low level account in which children only judge behaviors, while longitudinal evidence is more difficult to tease apart as even a dual systems account might show longitudinal relations between the two systems as one develops early and the other later and both are concerned with mental states.

In summary, children develop an explicit understanding of FBs around the age of 4, with certain tasks showing similar competence earlier by reducing task demands (Rubio-Fernández & Geurts, 2016; Setoh et al., 2016). Implicit FBU, however, appears to be present in certain tasks, but not in others and has been difficult to replicate. Some tasks show predictive longitudinal relations to later explicit understanding, but mixed results in cross-sectional work. Varying performance on these tasks at different ages emphasizes the importance of additional research on the third year of life as a transitional period and effort to establish a reliable implicit ToM paradigm. Independent of whether implicit FBU is a reliable concept, children's ability to assess knowledge states preverbally and how it relates to their explicit understanding of knowledge states remains to be investigated. In addition, further research is needed to determine if children assess knowledge states through concepts or behavioral rules.

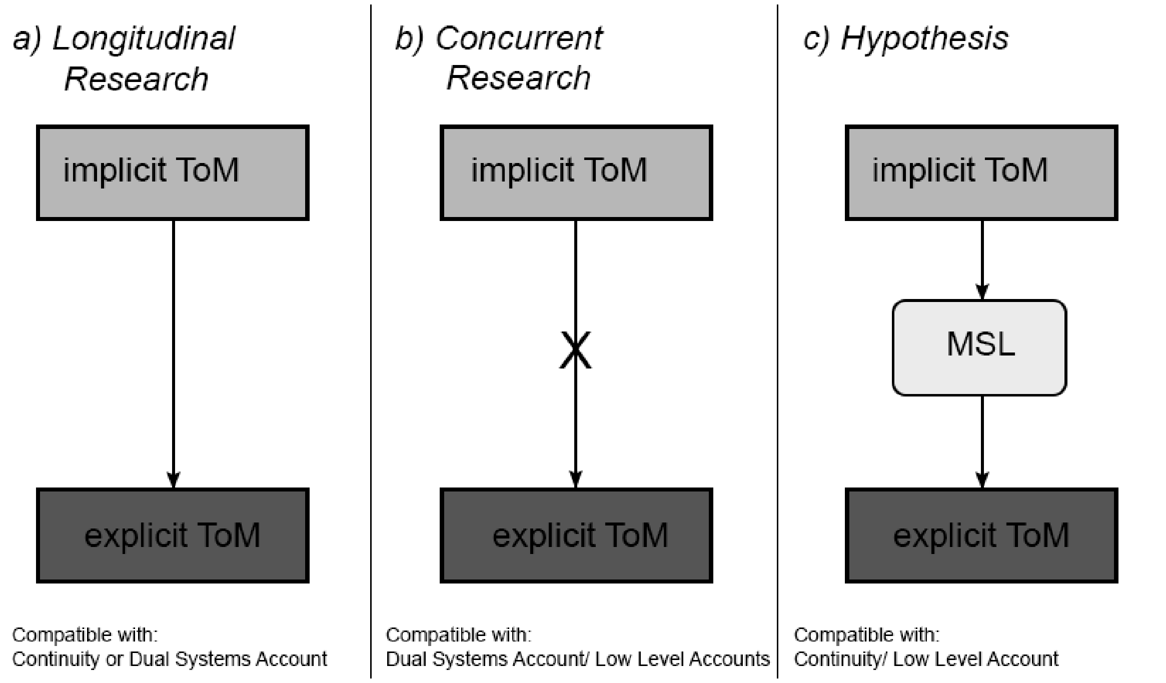

Fig. 2.2: Overview of longitudinal and concurrent research on the relationship between implicit and explicit ToM and a possible hypothesis. *a)* longitudinal studies find relationships between early implicit ToM and later explicit ToM which can be explained by continuity in the concepts or by two systems that develop subsequently. *b)* concurrent studies do not find a relationship between implicit and explicit ToM, which can either be explained by two independent systems at work or behavioral rules being at play for implicit ToM that do not relate to later explicit ToM *c)* one possibility is that mental state language (MSL) aids the transition from implicit ToM to explicit ToM by making mental representations that are present early on explicit. Concurrent relations between implicit ToM, explicit ToM and MSL might be explained by continuity within the concepts that is aided by language, or a low level account in which MSL verbalizes behavioral rules and thereby helps to develop explicit ToM.

The way children transition from their preverbal competence to later mastery of explicit ToM tasks can be investigated by examining children's mental state understanding during the third year of life. In the following section, the inconsistencies between early preverbal competence and later verbal competence will be elaborated in more detail before mental state language will be introduced as an important ability that may play a role for the transition. See Figure 2.2 for a brief overview of the discrepancy between implicit and explicit ToM and the hypothesis that mental state language may play an important role.

2.3 Ways to Bridge Inconsistencies Between Early Understanding and Later False-Belief Performance

Despite children showing early competence in understanding their caretaker's knowledge states and inferring their intentions (O'Neill, 1996), the relationship between preverbal competence and explicit MSU remains unclear. Even if we assume that children possess nonverbal concepts of mental states and develop explicit concepts around the age of 4, there are no clear conclusions on how exactly those concepts relate to each other and whether there might be abilities that aid the transition. When looking at children's development during the third year of life, implicit MSU is not the only way to address the transition between early nonverbal competence on mental states and later ToM performance. Two other abilities are relevant for ToM, namely children's perspective-taking abilities and their mental state language (MSL). Perspective-taking abilities and their relevance in the third year of life will be addressed in the following, whereas MSL and its relevance for ToM and the transitional period shall be addressed in the next section.

2.3.1 Perspective-Taking

Before passing FB tasks, children seem to have an ability that also requires them to differentiate between their own and somebody else's perspective, particularly in the domain of what they can see and what another person can see. Flavell (1974) has termed this skill 'level-1 visual perspective-taking'(VPT) as 2-year-old children are able to identify that people with different lines of sight see different things. For example, in a task with an occluder, children were able to understand that they might be seeing something that another person does not (Masangkay et al., 1974; Moll & Tomasello, 2006). However, 2-year-olds and many 3-year-olds struggle at 'level-2 visual perspective-taking', namely the idea that one and the same object can look different from different perspectives (Flavell, 1974; Moll, Meltzoff, Merzsch, & Tomasello, 2013). A pencil might look elongated to somebody looking at it from the side, but to a person that looks at it from the front it might resemble two circles with the pencil lead making up the smaller circle.

The main thing that differentiates perspective-taking from FBU, is that instead of addressing epistemic states (e.g. what a person knows or does not know) it is mostly concerned with visual perception. However, 'seeing' something can have both an epistemic and a non-epistemic meaning. Dretske (1990) argues that seeing in the non-epistemic sense may involve only observing an object without the knowledge that it is exactly that object. I can observe a piano without the awareness that it is indeed a piano that I am looking at, but instead, only perceive the single features. Seeing in the epistemic sense, however, involves acknowledging that the piano has certain properties that make it a piano and belong to it. The idea is that being able to perceive in the epistemic sense is an important step towards understanding the desires and beliefs of others, because through perceiving in the epistemic sense we categorize and think about objects. Therefore, FBU has often been termed 'epistemic perspective-taking'(EPT) (Moll & Meltzoff, 2011; Perner, Stummer, Sprung, & Doherty, 2002). From a theoretical perspective, examining perspective-taking is relevant as the relationship between preverbal and verbal perspective-taking may serve as a model for determining the relationship between children's preverbal and subsequent verbal MSU abilities and thus give us insight into whether children possess concepts of perspectives early on (ST) or whether production rules might be the better explanation (TT).

Flavell (1974) found level-1 perspective-taking competence in older 2-year-old children, but the emergence of level-1 perspective-taking was not examined. Moll and Tomasello (2006) conducted a novel study in which they investigated children's level-1 perspective-taking in a natural play situation. After playing with toys, an experimenter entered the room and looked for one of two toys that she and the child played with. Both toys were visible to the child whereas one was occluded behind a bucket for the experimenter. Overall, 24-month-olds, but not 18-month-olds handed the experimenter the occluded toy, suggesting that they understood what the experimenter could and could not see. The understanding could be due to a preverbal concept of 'seeing' or due to children recognizing behavioral patterns and learning that if an adult looks for an object, it is most likely the one close to a larger object.

Despite level-1 perspective-taking emerging early on, there is no conclusive evidence that it serves as a precursor for FBU. There is evidence for a relation between level-2 perspective-taking and FBU (Bigelow & Dugas, 2009), however, evidence on a relation between level-1 perspective-taking and FBU is inconclusive. A study by Carlson et al. (2004) found no relation between children's performance on a VPT task at 2 years of age and later performance on several ToM measures at 39 months. However, general performance on the ToM measures was low, which makes the interpretation questionable due to low variance. In a study by Thoermer et al. (2012) there was also no correlation between VPT in a verbal task at 30 months and FBU at 48 months. Similar results were found by Sodian and Kristen-Antonow (2015), who also used a verbal judgment task and did not find any relation between VPT at 30 months and FBU at 50 months. However, the study identified mirror self recognition at 18 months and declarative joint attention at 15 months as relevant predictors for later VPT. This corresponds to a view proposed by Moll and Meltzoff (2012) that postulates sharing perspectives and implicitly representing them during joint attention is a first step towards differentiating between perspectives. Sodian and Kristen-Antonow (2015) add that mirror self recognition might contribute by making representations that remained implicit during joint attention explicit. Thus, there appears to be a relationship between an important precursor of FBU (i.e. joint attention) and level-1 VPT, but no direct relation between perspective-taking and later explicit FB performance. The only study that found a direct relation between VPT at the age of 25 months and children's later FBU at the age of 50 months was conducted by Yeung, Müller, and Carpendale (2019) and used a traditional perception production task. In the task, children were given a picture card that was attached to a wooden stand and could be turned around. One side of the card depicted a rabbit, whereas there was nothing on the other side. Children's performance was scored based on whether children spontaneously showed the experimenter the bunny side when asked, or whether they only did so after a prompt to show the bunny. Yeung et al. (2019) computed a model in order to identify which factors explained a significant amount of variance for FBU around 50 months, including level-1 perspective-taking, mental state language and comprehension of pretense play. Level-1 VPT was the only significant predictor after controlling for number of siblings, verbal ability, working memory and conflict inhibition scores.

At this point, it is important to note that performance varies greatly depending on whether the VPT task is verbal or not. The longitudinal studies by Thoermer et al. (2012) and Sodian and Kristen-Antonow (2015) used a verbal task, whereas Yeung et al. (2019) did not. When looking at how well children perform on different perspective-taking tasks in general and how early they master them, the differences between tasks are easy to identify. Moll and Tomasello (2006) found competence in 24-month-olds in their action-based task, but Gonzales et al. (2018) only found competence around 3 years of age in a traditional verbal task where children had to explicitly state whether they and another person could see an item or not. In a nonverbal study by Sodian, Thoermer, and Metz (2007), even 14-month-olds were shown to look longer during a looking-time task, when an agent reached for a new object instead of the familiarized goal object even though the latter was visible. This was not the case when the goal object was invisible to the person, but visible to the infant. In the verbal task by Gonzales et al. (2018), 2-, 3- and 4-year-old children were presented with two stuffed animals, one of which was hidden behind an occluder. Subsequently, they were asked whether they could see the first animal (which was visible) and whether they could see the second animal (which was occluded). The same was repeated for a puppet on the experimenter's lap to identify whether children were aware of the perspective of a third person. The goal of the study was to identify whether children acquired their own VPT before they acquired that of a third person. Gonzales et al. (2018) showed that verbally reporting on one's own perspective was possible for 2-year-olds whereas only 3-year-olds were able to correctly report on the perspective of the puppet. Thus, only a nonverbal perspective-taking task has been able to find predictive relations with FBU, but not verbal perspective-taking tasks. One of the reasons for this discrepancy might be children's overall low performance on verbal perspective-taking tasks at a young age. Nonverbal tasks are easier and might therefore be better predictors early on. Whether that is because children already possess a nonverbal concept of perspectives or rely on production rules to pass them remains to be answered.

In summary, visual perspective-taking is an ability that develops around the age of 2 and although verbal VPT tasks do not have a definite relation to later FBU, there is evidence that VPT is related to other relevant precursors of ToM, among them joint attention (Sodian & Kristen-Antonow, 2015) and mental state language (Chiarella et al., 2013). Furthermore, as mentioned in earlier sections FB is not the only criterion for MSU and perceptive-taking on its own describes a valid mental state, namely 'seeing' that can serve as a helpful model for MSU. One may thus argue that examining both verbal and non-verbal perspective-taking enables us to investigate the relationship between preverbal MSU and subsequent verbal MSU, in particular for 'seeing' in the epistemic sense. In the following part of this literature review, I will focus on mental state language, which has been found to relate to both verbal perspective-taking and mastery of explicit ToM tasks.

2.4 Mental State Language as Early Theory of Mind

One of the biggest challenges of mental state understanding is that mental states are something we are not able to explicitly observe in the outside world. It comprises states and processes that occur in our own minds and in the minds of other people. When acquiring general language, children are able to associate an object with the respective term for it, e.g. 'car' with a toy car that they own or the family car. Mental states on the other hand, cannot be easily associated with an entity in reality, as they can neither be observed as an object, nor directly in behavior (Montgomery, 2002). For example, a person can feel differently than their face suggests and children learn to identify this discrepancy only much later, even after acquiring basic FBU (Wellman et al., 2001). What makes mental state language (MSL) so special in comparison to general language, is that MSL introduces terms that denote something that cannot be grasped, a concept that is not observable e.g. 'to know', 'to believe', 'to guess' or 'to want'[6]. Through the existence of these terms, children have the opportunity to generalize from multiple occurrences of the terms until they find something all behaviors that involve the term have in common, namely the mental state. There is a strong interactive relationship between children talking about and using mental states and their caregivers explaining others' and their own mental states (Slaughter, Peterson, & Carpenter, 2008; Taumoepeau & Ruffman, 2008). Explanations make learning mental state terms easier, because children get an insight into what is happening within another person without having to observe it directly. Ruffman and Taumoepeau (2016) showed in a longitudinal study that it is not the amount of times a mother uses 'want' that aids children's understanding, but rather that it is the amount of different contexts in which the verb is used. Research in autism supports the idea that conversation is what makes mental state terms applicable, as autistic adults have similar mental state term vocabularies, but use the terms much less frequently in personal narratives, suggesting a lack of qualitative conversational input (Bang, Burns, & Nadig, 2013).

MSL has been defined in different ways over the years and sometimes only includes desire and epistemic terms like 'want' and 'know' and other times includes everything from physiological terms like 'hungry' to emotional terms like 'sad'. In this thesis, MSL is always taken to mean a wide range of categories, among them physiological terms, emotional terms, obligation terms, volitional terms and cognition terms.

When considering MSL in general and its association with MSU, it becomes apparent that not all mental state terms are equally relevant for understanding beliefs and intentions. Children's use of cognition terms like 'think' and 'know' in particular is thought to be a hint of their internal MSU, a sign that children understand what these words denote and are subsequently able to use them in everyday conversation (Ruffman, Slade, Rowlandson, Rumsey, & Garnham, 2003). As all early signs of in-

[6] One may argue that 'want' is more straightforward behaviorally as children might associate grasping motions with desire.

tention understanding and the attribution of knowledge states in infancy have been based on inferences and were not verbal, MSL is considered to be the earliest sign of an explicit understanding of mental states (Bartsch & Wellman, 1995; Bretherton & Beeghly, 1982). Two studies have famously investigated the emergence of MSL in early childhood. The earliest was an analysis of MSL in 2-year-olds by Bretherton and Beeghly (1982) in which they asked parents to fill out which words their children used. Several categories were investigated, among them physiological terms (e.g. 'hungry' and 'tired'), perceptual terms (e.g. 'see' and 'look'), emotional terms (e.g. 'happy' and 'sad'), desire terms (e.g. 'want') and cognition terms (e.g. 'think' and 'know'). The authors were particularly interested in which categories children had already mastered at the age of 2. Overall, children frequently used physiological terms, perceptual terms, emotional terms and also desire terms, however, their usage of cognition terms was much less frequent. The difference between the usage of desire terms and cognition terms like 'think' and 'know' is very much in line with the usual explicit trajectory of children's ToM performance in which desire tasks are much easier to solve than tasks about belief (Wellman et al., 2001). Here, it is important to note that even though cognition terms were used much less frequently, there were still children who used them. One may wonder if the mere usage of the terms 'think' and 'know' suggests MSU in these few children who were able to use them. Bartsch and Wellman (1995) looked at children's MSL in more detail by analyzing utterances found in the CHILDES database, a collection of mental state terms produced by children between the ages of 18 months and 5 years. In particular, they looked at in which contexts children commonly use mental state terms. In their analysis, Bartsch and Wellman (1995) found that children already use desire terms in the right context before the age of 2, but genuine references to beliefs and knowledge only began around 3 years of age. Sabbagh and Callanan (1998) showed that when looking at utterances of 'think' and 'know' in detail, children only began to use the terms commonly and in the right contexts around the age of 4 and 5, which corresponds to the typical age at which children's belief comprehension emerges (Wimmer & Perner, 1983). This late development might also be due to a selection of terms their mothers use most commonly. Nelson and Shaw (2002) investigated situations in which children use 'think' and 'know' and most of these situations corresponded to the way adults use these terms e.g. talking about oneself in terms of 'think' and denying one's own knowledge in terms of 'know'. One might argue that using the mental state verb 'know' in the formulation 'you know what?' is not a genuine reference to a mental state, as it can be used as an attention grabber in daily conversation. Furthermore, Bretherton and Beeghly (1982) found that children used 'know' most commonly in the formulation 'I don't know', which might indicate genuine awareness of one's own knowledge state, but could also be used as a refusal to answer or a means to change topic.

In summary, MSL is a crucial concept for MSU as mental states cannot be easily observed and MSL provides words in order to talk about them. Cognitive terms are considered to be especially important and analyses of spontaneous speech have concluded that children acquire desire terms before belief terms and that genuine references to beliefs and knowledge occur earliest around the age of 3. Before the age of 3 most utterances involving cognition terms do not seem to be referencing a mental state.

However, more recent research on children's use of mental state terms at the age of 2, suggests that genuine references to mental states are much more common than has previously been assumed (Harris, Ronfard, & Bartz, 2017; Harris, Yang, & Cui, 2017), which will be presented in detail in a later section. Since children's mental state terms begin to develop earlier than expected and might reveal competence in the third year of life, MSL seems to be a good measure to fill the gap between early MSU in preverbal communication and later ToM performance on tasks.

In the following sections, I will outline the relationship between MSL and ToM, arguing why MSL can be considered a precursor and early sign of ToM and why it is a reliable measure. Next, I will elaborate on the early development of MSL and current research on when children show genuine references to mental states. A genuine reference is defined as the correct use of a mental state verb in order to denote a mental state including correct inferences depending on the context. As genuine references to mental states would suggest children possessing early MSU, experimental studies are particularly important to identify children's true competence. These studies will be presented in the subsequent section with a focus on factivity understanding. Finally, I will address the relationship between MSL and basic cognitive and social skills as these are believed to substantially drive its development.

2.4.1 Developmental Trajectories between Mental State Language and Theory of Mind

As mentioned above, MSL is a way to explicitly communicate what is going on inside one's own mind and the minds of others. This relationship between MSL and MSU becomes apparent very early on. Kristen et al. (2011) investigated how early precursors of ToM in infancy and later MSL are related in a longitudinal study. The results showed a predictive relation from joint attention at the age of 9 months to children's intention-based imitation at 15 months, but also later MSL at the age of 24 months. More specifically, 12-month-old's comprehension of an agent's imperative motive was predictive of their usage of desire terms at the age of 24 months. This finding shows that specific acquisition of MSL terms appears to be based on early joint attention abilities (i.e. imperative joint attention for desire terms). However, joint attention is not the only ability that MSL is strongly related to. Chiarella et al. (2013) examined the concurrent relationship between MSL production and VPT at the age of 30 months. So far this is the only study to do so. The authors tested children on their VPT abilities with two tasks adapted from Flavell, Everett, Croft, and Flavell (1981). In the first one, children were presented with a card that had an animal depicted on each side. Test questions were what animal the child could see and what animal the experimenter could see. In the second task a picture of a turtle was placed flat on the table and one half was occluded by a blank card that was held perpendicular to the picture. Subsequently, children were asked which part of the turtle they were able to see (i.e. feet or shell) and which part the experimenter was able to see. Children's

MSL production was measured with a parental questionnaire. Chiarella et al. (2013) found a significant correlation between VPT abilities and MSL at the age of 30 months, independent of general language abilities. This finding suggests that independent of children's language abilities, there is common conceptual ground between taking one's own perspective and that of somebody else and being able to produce mental state terms.

Furthermore, apart from an early relationship between precursors of ToM and MSL, there have also been several studies that showed direct predictive relationships from MSL to later ToM performance (Brooks & Meltzoff, 2015; Devine & Hughes, 2019; Olineck & Poulin-Dubois, 2007), indicating that MSL is not only developmentally related to ToM precursors, but at the same time it is important for later ToM mastery. As mentioned in earlier sections, the MSL input of mothers is crucial for children's later MSL and differs depending on the age of the child (Taumoepeau & Ruffman, 2008).

In summary, MSL is not only conceptually relevant for MSU, but also relates well to both, precursors of ToM and subsequent explicit ToM mastery. Interestingly, when it comes to relations to ToM, precursors are related to the child's later ToM performance, but predictive relations to later ToM also exist from mothers' mental state talk. Since not only the mental state talk of mothers (Devine & Hughes, 2019), but also the mental state talk of children themselves (Brooks & Meltzoff, 2015) appears to be predictive of later ToM performance, I am going to present research on mother's mental state talk and children's mental state talk in two separate sections.

Mother's Mental State Talk

One of the first longitudinal studies that showed the importance of mothers' mental state talk for children's later FBU was conducted by Ruffman et al. (2002). Mothers were asked to describe pictures to their children at three different occasions over the course of 1 year. Concurrently, children's MSU and their mental state vocabulary was assessed. Results showed reliable predictions from mothers' mental state talk at earlier time points towards children's MSU at later time-points. The relationship was unidirectional and did not depend on other factors like children's own mental state talk, their verbal abilities, their age, mothers' education or any other utterance apart from mental state talk spoken during the picture descriptions. A more recent longitudinal study by Kristen-Antonow, Licata-Dandel, Müller, and Sodian (2018) showed that mothers' production of cognition MSL was an important predictor of children's later first- and second-order FBU. Mother-child dyads were investigated in a free play situation at 7 months and a book-reading session at 24 months, followed by measures of first-order FB at 50 months and second-order FB tasks at 70 months. Mothers' mental state talk during the book-reading session at 24 months was significantly related to children's first- and second-order FB performance at 50 and 70 months, independent of mother's verbal IQ. If mothers' mental state talk is relevant for children's MSU, how about mothers' own MSU? Devine and Hughes (2019) investigated this hypothesis and found that only mothers' mental state talk and not their ToM skills was a significant predictor of children's ToM performance. Additional studies investigated which other aspects of the mother's behavior and abilities might play a role apart from mental state talk. Ruffman, Slade, Devitt, and Crowe (2006) addressed whether mothers' mental state talk or parenting style predict children's later MSU. Children were tested once at the age of 3 and another time at the age of 4. The assessment involved several ToM tasks and a task during which mothers helped their child draw items. Mothers' mental state talk significantly predicted children's MSU, whereas their parenting style had no effect on children's performance. Gola (2012) went even a step further and examined whether it had to be the mother that gives children input on mental state talk. In an experimental study, preschoolers were exposed to 128 mental verb utterances that were shown to them in a video over the course of two weeks. Mental verbs were presented under different conditions: a) in the form of a statement or a questions, b) said in first person or third person, and c) in an interactive manner or overheard by the child. The form of the mental state verbs did not matter to children's performance on FB tasks, however, when they overheard a statement about a third person, they performed significantly better on FBU. This suggests that mental state talk does not need to be directed at the child and can scaffold their MSU in passing as long as it emphasizes the mental state of a third person.

All studies presented until now have been concerned with children's explicit MSU. A recent study by Roby and Scott (2018) is the first study on mothers' mental state talk and MSU that found a significant relationship between the input given by mothers and 2,5-year-old children's performance on a verbal non-standard anticipatory looking task. The authors argue that the relationship supports the idea that MSL is important for assigning beliefs and intentions and tracking the behavior of an agent, even before the usual age of 4. However, also this non-standard FB task has not been replicated yet.

In summary, mothers' mental state talk has been shown to play an important role for children's subsequent MSU and it is not mothers' own MSU or parenting style, but the amount of MSL they use that predicts performance. Aside from mothers' mental state talk, also the mental state vocabulary of the child has been found to predict later ToM skills. In the following section, I will elaborate on children's mental state talk and how it relates to their mental state understanding.

Children's Mental State Talk

For a long time, research has mostly focused on mothers' mental state talk as children appeared to only begin to use mental state terms appropriately around the age they solve FB tasks (Bartsch & Wellman, 1995, pp.187-190). However, not only mothers' mental state talk has been shown to be related to FBU, but also children's own spontaneous mental state talk (Symons, Peterson, Slaughter, Roche, & Doyle, 2005). When examining younger children, most studies do not use spontaneous measures to elicit mental state talk as it only reflects a single situation, but rather ask parents to fill out questionnaires on which terms their child is already familiar with (Kristen et al., 2011). These measures avoid the bias of a single measurement.

Recent longitudinal work suggests that children's MSL is important for later ToM even before the age at which they master cognition terms. In a longitudinal study by

Brooks and Meltzoff (2015), the relationship between gaze-following, MSL and FBU was investigated. Children's gaze following at 10,5 months significantly predicted their mental state terms at 2,5 years, which in turn predicted their FBU at the age of 4,5 years. This study suggests continuity from a nonverbal understanding of gaze, to the production of mental state terms and finally to the ability to assign FBs to others. Another study by Olineck and Poulin-Dubois (2007) showed a predictive relationship from MSL at 32 months of age to ToM mastery at the age of 4, additionally supporting the claim that children's MSL is predictive of their later ToM performance. Not only longitudinal studies in neurotypical children emphasize the importance of MSL for ToM. A study by Siller, Swanson, Serlin, and Teachworth (2014) with autistic children at the age of 7 showed that their vocabulary of emotional terms was significantly related to their concurrent ToM abilities.

From these results it becomes apparent that there is a strong relationship between MSL and MSU. When looking at the relationship between early MSL and later ToM, production of mental state terms appears to play an important role. Interestingly, this relationship appears to switch in early childhood, as primary school children's understanding of mental state terms is a better predictor of their ToM abilities than their production of the terms (Grazzani & Ornaghi, 2012). This is probably due to the fact that production is limited to the situations children are familiar with, whereas comprehension can extend beyond familiar contexts, especially in primary school children.

An advantage of studying MSL is that it enables us to understand the mechanisms of MSU better through making mental states explicit by addressing them verbally. Thus, studying MSL enables us to study the transfer. Furthermore, MSL is a very reliable measure as it remains stable throughout the third year of life and does not show big differences across cultures (Kristen, Sodian, Licata, Thoermer, & Poulin-Dubois, 2012; Pascual, Aguado, Sotillo, & Masdeu, 2008; Tarchi, Bigozzi, & Pinto, 2019). Additionally, it is easier to measure than other abilities as it can be reliably assessed with a parent questionnaire (Harris & Jones, 1997).

In summary, both mothers' and children's MSL are important predictors of subsequent ToM mastery. In the following, I will elaborate on the early development of mental state terms in toddlerhood and in particular on the question whether spontaneous usage of mental state terms reflects MSU.

2.4.2 Early Mental State Language

As mentioned above, children were believed to use terms like 'think' and 'know' rarely before the age of 24 months, with a significant increase between 24 and 33 months (Taumoepeau & Ruffman, 2008). It has been argued that this early usage of 'know' does not clearly reference epistemic states, but may serve various other functions in regulating children's early conversations (Shatz, Wellman, & Silber, 1983). Only around the age of 36 months, clear evidence for reference to beliefs and knowledge states was found in an analysis of children's utterances (Bartsch & Wellman, 1995, pp.25-47).

Recent work by Paul Harris (Harris, Ronfard, & Bartz, 2017; Harris, Yang, & Cui, 2017) challenged this traditional view by reanalysing a naturalistic corpus of child speech for 3 children (pattern confirmed for 8 children through analyses by Bartz (2017)). The authors argued that prior analyses have underestimated children's understanding of knowledge and ignorance. Their reanalyses of spontaneous speech data showed that even young 2-year-olds use the word 'know' flexibly in various functions, namely to affirm, to query and also to deny knowledge, in particular, by differentiating between what a person knows and does not know and only rarely to question their own knowledge. The authors argued that the asymmetry children portrayed in their queries, namely queries about the knowledge of others, but not their own, constituted additional evidence for the ST. Children's own mental states are more accessible to them, therefore they only rarely have to question their own knowledge, whereas the knowledge state of another person is mostly unknown to them. Still, across all occurrences of 'know', 2-year-olds seemed to use the term mostly appropriately to describe their own and other persons' knowledge states, thus indicating an early, according to the ST, conceptual understanding of 'know'. As mentioned above, using a mental state term in the appropriate context is considered to be a sign for children having grasped the concept as they are able to make appropriate inferences. Additional evidence for an early implicit sensitivity for knowledge states comes from an analysis of children's questions: Harris, Yang, and Cui (2017) found that children pose both interaction-seeking questions and information-seeking questions, but that the latter outnumber the former at the age of 30 months. Asking another person for information indicates at least a rudimentary idea of what this person knows and does not know. Children's usage of the mental state verb 'know' was also demonstrated in a task that investigated metacognition of ignorance (MCI) in 16- to 37-month-olds (Bartz, 2017). In this study children were shown pictures of real objects (e.g. a bird, a car) and fake objects that had vaguely familiar elements, but were unknown as a whole. For each picture, the child was asked whether they knew what was depicted. Children were found to produce a lot of incorrect terms or invented words for the fake objects, but they also showed signs of uncertainty, be it shaking their heads, making longer pauses and saying 'um' or explicitly stating that they did not know the object. Harris, Ronfard, and Bartz (2017) propose that this suggests children have an early understanding of their own ignorance, even if they do not explicitly state it in every case. Furthermore, young children appear to use 'know' in an appropriate context, namely to deny their own knowledge.

Rubio-Fernández (2019) objects and argues that using the verb 'know' in appropriate situations does not necessarily undermine the traditional view of late acquisition. The author criticizes Harris et al's liberal interpretation and argues that 'I don't know' could be used as an alternative for 'I can't answer' that was established through conversation and frequent use by the caregivers. Rubio-Fernández (2019) also points out that children's preference for questioning others' knowledge states compared to their own knowledge state can additionally be explained by the caregiver's usage of mental state terms. The argument is that also in adult speech it is more common to question someone else's knowledge instead of one's own. Another study that supports the idea that 'don't know' is only a conversational chunk instead of a meaningful unit was done by Brandt, Verhagen, Lieven, and Tomasello (2011). In this study, 4- and 5-year-old

German-speaking children were asked to switch first person sentences to third person sentences. The motivation was that a corpus study revealed mental state verbs to be mostly used with the first person (i.e. 'I believe' instead of a scheme like 'X believes'), which might suggest that children memorized them as chunks. Four-year-olds struggled to separate person and verb for mental state verbs, independent of their frequency, whereas they were able to do it with frequent transitive verbs and verbs of communication. Only 30-35% of the children succeeded at least once. The 5-year-olds, however, were able to exchange the first person for the third person for all types of verbs. The authors argue that children only rarely hear mental state verbs being used in third person and are therefore not able to create a person-general scheme. This fits well with research mentioned above that suggests third-person formulations are particularly what drives FBU (Brandt, Buttelmann, Lieven, & Tomasello, 2016; Gola, 2012).

It is therefore conceivable that in a naturalistic situation children mainly respond to a behavioral request and show their inability to comply by replying with 'I don't know' or use a chunk they have learned to apply in similar situations. Thus, children's spontaneous usage of mental state verbs, particularly epistemic verbs, does not shed light on children's genuine understanding of the verbs.

In summary, recent research suggests that already 2-year-olds might have MSU as they use the verb 'know' flexibly and in appropriate contexts, in addition to demonstrating metacognition of ignorance. However, authors like Rubio-Fernández argue that children merely copy their usage of the verb 'know' from their caregivers without understanding the implications. Also experimental evidence seems to argue for mental state verbs to be memorized as chunks. Experimental evidence is needed to investigate whether children's early use of mental state terms remains appropriate in a controlled and constrained environment and thus, provides evidence for children being capable of mental state understanding.

2.4.3 Experimental Evidence on Mental State Language Comprehension

In order to investigate children's understanding of mental state terms, experimental studies are crucial. When looking at the types of verbs that are usually investigated, a large portion of research focuses on factivity, the degree to which a mental state verb like 'know' or 'believe' entails the known or believed proposition to be a fact. Several researchers argue that ToM cannot be reduced to passing the FB task only (Apperly, 2012, pp.15-33) and that the comprehension of factivity is a much better measure for children's understanding of belief. Nagel (2017) points out that instead of differentiating between an agent having a true belief or a false belief, what matters is whether the agent is in a factive or a non-factive state. The idea is that the conclusions drawn from a non-factive mental state verb are less certain than those drawn from a factive one. When an agent claims that he knows something, we can assume that the following statement is true, whereas an agent thinking something does not give us any information on whether the statement is true or not. Let's take the example 'Lisa knows a ball is in the box'. In the case of 'know', we can be certain that the ball is in the box, as the verb 'know' is factive. In contrast, the statement 'Lisa thinks a ball is in the box' does not give us any certainty about the content of the box as 'think' is a non-factive verb and the box could contain a ball, but also something else. A typical FB task does not assess factivity, as children are explicitly shown or told whether the agent's belief is true or false. This makes the assessment of certainty obsolete, whereas the judgment of a statement as the one above without any additional situational cues would still require an intact understanding of factivity (Kristen-Antonow et al., 2019). An effective method to study factivity are therefore tasks that assess speaker certainty and require children to choose between speakers using factive and non-factive mental state terms.

Children's ability to distinguish between the verbs 'know', 'think', and 'guess' has been studied experimentally since the 1980s (Dudley, Orita, Hacquard, & Lidz, 2015; C. N. Johnson & Maratsos, 1977; Léger, 2007; Moore, Bryant, & Furrow, 1989). Commonly, 5-year-old children achieve competence in one contrast of know with a non-factive verb (guess, think) but not both, whereas full competence is only reached at the age of 8 years (Kristen-Antonow et al., 2019). In the first study that investigated factivity in 3- to 8-year-olds, Moore et al. (1989) presented children two puppets that talked about the location of a hidden object and either used the factive verb 'know' to express their certainty, or the non-factive verbs 'guess' and 'think'. Only at the age of 4, children began to prefer the agent who 'knew' the location over the agent who only 'guessed' or 'thought' the object was there. An important point is that even at the age of 8, children were not able to differentiate between 'think' and 'guess'. Kristen-Antonow et al. (2019) investigated factivity longitudinally by using Moore et al's paradigm and testing German children at the age of 50, 60, 70 and 94 months. Even though children were able to perform significantly above chance on the factive/non-factive verb contrasts around 60 months, this ability was only consistent starting 70 months. Only at 94 months more than 50% of the sample reached competence in differentiating factive from non-factive verbs. Furthermore, Kristen-Antonow et al. (2019) investigated the relationship between beginning understanding of factivity at 70 months and children's FBU at 60 months, finding a predictive relation from early FBU to factivity understanding at 70 months. It is important to note that the relationship between FBU and factivity was one-directional, suggesting that an understanding of false beliefs develops before children are able to differentiate between factive and non-factive verbs. This is an important finding as it suggests that understanding what knowing something entails in terms of factivity develops only much later. However, there have also been studies that suggest an earlier understanding of factivity than the studies presented above. Brandt et al. (2016) have found competence to distinguish 'know' and 'think' in 4-year-olds when they were asked to choose between two informants. The authors compared children's performance in a scenario in which they had to judge first person utterances (i.e. 'I know' vs 'I think') and another in which they judged third person utterances ('the pig knows' vs. 'the cow thinks'). Interestingly, 4-year-olds performed better on the third person version of the task compared to the first-person version. This might also be in line with children learning better from hearing mental state terms used in the third person (Gola, 2012). A recent study by Dudley et al. (2015) suggests competence in the know-think distinction in 3-year-olds, however, when looking at consistent responses over trials, the majority of children used a non-factive interpretation of ‘know’ (see Rubio-Fernández, 2019).

Even if children do not yet possess an understanding of factivity and consistent competence only develops at the age of 70 months (Kristen-Antonow et al., 2019), the question remains whether there is an early understanding of reference to mental states that is not based on a full grasp of factivity. To claim that even 2-year-olds have some understanding of knowledge and ignorance as expressed by ‘know’ and ‘don’t know’ (Harris, Yang, & Cui, 2017) does not entail the claim that 2-year-olds understand ‘to know’ as a factive verb. There is no need for an understanding of factivity if one assumes early preverbal informing behaviors around 18 months (O’Neill, 1996) to be conceptually connected to an understanding of mental state verbs. Rather, the assumption of conceptual continuity would lead to the prediction that when 2- and 3-year-old children are confronted with a verbal contrast of a mental state term, they are able to respond appropriately. For example when presented with the contrast of ‘I know x/I don’t know y’, they understand the inference that the speaker needs information about y, not about x. Thus, a listener, who is supposed to cooperate with (‘help’) the speaker, will show the speaker the object the speaker claimed to be ignorant of (y), rather then the one she is knowledgeable about (x). If there is conceptual continuity between preverbal informing behaviors and the later understanding of mental state verbs, it should be possible for 2- and 3-year-olds to draw the correct inferences from ‘know’ and ‘don’t know’ earlier than they distinguish between ‘know’ and ‘think’. Until now there have been no studies that investigated children’s ability to draw the correct inferences from verbal knowledge statements without comparing different factive verbs. If 2- and 3-year-old children were able to draw the correct inferences from a verbal statement with a mental state term, this would suggest understanding of what the mental state term entails and thereby an understanding of mental states. Of course this understanding does not correspond to FB reasoning or factivity yet, however, it would be a first sign for a relation between comprehension in preverbal communication and subsequent verbal comprehension.

In summary, experimental evidence on children's understanding of mental state verbs has mostly been done in the area of factivity and consistent understanding of factivity only begins to develop around the age of 5. Several studies have also been able to find competence earlier, but have not been replicated yet. I argued that in order to examine children's understanding of mental state verbs it is not necessary to examine factivity, but also simpler contrasts could be sufficient to investigate whether children know how to draw the right inferences from a mental verb.

When looking at mental state language in general, it cannot be ignored that no understanding of mental state terms can emerge without a minimal understanding of language. However, not only language appears to matter for the acquisition of mental state terms, but also other basic cognitive functions. Similarly to ToM, a few basic cognitive and social abilities have been suggested to be crucial in the development of mental state language and subsequent mental state understanding. In the following section, research on these abilities will be presented before the literature review will be concluded with the main research issues that remain and the hypotheses of this thesis.

2.4.4 Relations between Mental State Language and Basic Cognitive and Social Abilities

The relations between MSL and basic cognitive and social abilities have not been investigated much over the past few decades. Of course many studies have found relations between general language and subsequent acquisition of mental state terms (Kristen et al., 2011; Ruffman et al., 2002), but relations to other domains are less obvious and have been investigated less. Carlson et al. (2004) examined preschoolers longitudinally, once at the age of 2 and another time at the age of 3. Children were assessed in a battery of executive function tasks and ToM tasks at both time points and parents were asked to fill out questionnaires about their child's temperament, MSL and vocabulary. The authors were particularly interested in the relation between executive functions and ToM, but included MSL as an early measure that might shed light on mental state understanding before children master ToM tasks. At the age of 24 months, there was little evidence suggesting a relationship between early executive functions and ToM abilities. Only children's comprehension of pretense correlated with one of 5 executive functions tasks after controlling for age, verbal ability, and sex. However, there was a significant relationship between children's MSL and a task for inhibition skills independent of age, sex and verbal ability. At 39 months, the pattern changed and there was a significant relationship between ToM tasks and executive function tasks, while the significant relation between executive functions and MSL also remained in place. This suggests a relation between executive functions and MSL early on that stays during the third year of life. In particular inhibitory control might be necessary in order to differentiate between using mental state terms to talk about oneself and others as one's own mental state may need to be inhibited before we are able to focus on the mental state of another person. A study that investigated the concurrent relationship between MSL and inhibitory control was done by Bellagamba et al. (2014). The authors measured 18- and 24-month-olds' performance on an inhibitory control task, a word production task and assessed their mental state vocabulary. Similar to previous studies, there was a relationship between word production and MSL. However, Bellagamba et al. (2014) also found a significant correlation between MSL and inhibitory control when controlling for vocabulary size. This suggests that shifting from one perspective to another in a cognitive control inhibition task is related to the ability to talk about mental states of oneself and others.

Apart from executive functions, Tarchi et al. (2019) has investigated the relationship between mental state talk in children of various ages, beginning at toddlerhood until primary school age, and their narrative competence. In particular, the authors were interested in whether children who have better narrative structure while telling a story, are also more likely to use an abundance of mental state terms. The results of the study confirmed the hypothesis that children with better narrative structure were more likely to use emotional, cognitive, moral and socio-relational terms. Of course the causal effect of the relation does not become apparent through a correlational study and it is possible that children with better narrative structure use more mental state terms, or that children who use more mental state terms have a better narrative structure. However, prior research on the importance of syntax for the processing of beliefs (J. De Villiers, 2007; J. G. De Villiers & De Villiers, 2000) suggest that both abilities require metarepresentations and therefore share conceptual demands.

In summary, research on MSL and other basic cognitive and social skills remains sparse. Inhibitory control has been found to play a role and concurrent relationships to narrative structure have been identified. Language has been used as a control variable in some studies in order to determine the specificity of the relation between MSL and mental state understanding, but the relation between language acquisition in general and the development of MSL has rarely been addressed. Therefore, it would be important to investigate predictive relations between executive function, language and MSL and be able to determine which basic abilities particularly foster MSL. This is especially important in light of the relationship between early MSL and explicit ToM, suggesting that improving MSL through intensive input and conversation might improve children's later ToM abilities. In the following I shall address issues that remain in current research and how this thesis aims to address them experimentally.

2.5 Main Research Issues

In summary, there appears to be a discrepancy between children's preverbal MSU and their subsequent mastery of explicit ToM tasks. It appears that children already possess a basic concept of agency and knowledge states preverbally, be it in the form of production rules or an actual concept. However, this comprehension of knowledge states is not expressed explicitly in ToM tasks until much later. One possible explanation is that additional representational resources are necessary in order to make these production rules or preverbal concepts explicit. One such representational resource is language, which has been found to predict later performance on explicit ToM tasks (Milligan et al., 2007). A particular subcategory of language, MSL, may be of particular importance as a representational resource that aids the transition between preverbal competencies and subsequent explicit ToM mastery for several reasons.

MSL production is considered to be the earliest sign of explicit MSU as it is the first time children begin to verbalize mental states, be it in appropriate contexts or not. Evidence suggests that MSL is a crucial precursor for ToM because it 1) relates well to known earlier precursors (Kristen et al., 2011) and 2) relates to later explicit ToM performance (Brooks & Meltzoff, 2015; Olineck & Poulin-Dubois, 2007). But, there is still little experimental research on how closely it reflects children's MSU. Harris, Ronfard, and Bartz (2017) argue that children use mental state terms early and that they already possess a (basic) concept of their own knowledge and ignorance, thus, suggesting a (basic) representation of mental states that is present early on and may already be associated with the respective mental state terms. In contrast, Rubio-Fernández (2019) argues that children might be learning verbal formulae instead of using terms for their mental representations (or production rules) and only later on through additional contextual input children are able to associate their mental representations (or production rules) with mental state verbs. Also, a differentiated understanding of the mental state terms 'know' and 'think' with which children can draw correct inferences about factivity only develops much later (Kristen-Antonow et al., 2019).

Thus, is it important to investigate whether and to what extent the appropriate use of mental state terms (i.e. 'know') in conversation, reflects an understanding of the semantics and pragmatics of these terms. Identifying the role that MSL plays for MSU at a young age enables us to address its relevance for bridging the discrepancies that are present in ToM development. One way to address the relation between MSL and understanding of semantics and pragmatics of mental state terms is by 1) assessing whether children use the terms in appropriate situations as reported by Bartz (2017) and Harris, Ronfard, and Bartz (2017) and 2) determining whether children are able to draw correct pragmatic inferences from verbal statements, similar to the non-verbal inferences they draw in infancy in experimental settings. An additional aspect is that individual differences in children's language have been shown to be relevant for later explicit ToM performance, but there is little research on the relation between production and comprehension in children's early mental lexicon. A study of concurrent relations between children's MSL production and their performance on comprehension tasks would give us insight into the relevance of production for comprehension and how MSL production might fit into the bigger picture of ToM development.

The third year of life is a transitional period that enables us to shed light on the discrepancy between preverbal MSU and the mastery of explicit ToM tasks by investigating the concurrent relationship between children's preverbal ToM-related abilities, explicit ToM-related abilities and MSL production. By examining abilities that develop in the third year of life and are known to be associated with preverbal MSU or the subsequently developing mastery of explicit ToM tasks, one may gain insight into the transition from one to the other and whether MSL production is relevant for both, thus showing conceptual continuity and arguing for children's preverbal competence to be the result of conceptual understanding (ST) or only showing relations to verbal MSU and suggesting that preverbal MSU might rather be an accumulation of production rules (TT). Thus, a second research issue that remains unaddressed is the role that MSL plays for the transition from preverbal MSU to subsequent explicit MSU.

Several preverbal ToM-related abilities and explicit ToM-related abilities have been shown to play an important role in the third year of life, in particular preverbal and verbal level-1 perspective-taking and implicit FBU. As MSL production has been suggested to play a mediating role in the transition from preverbal MSU to later explicit (and also verbal) MSU, it is of great value to examine its relation at a time point in which relations to both preverbal and verbal might be identified. Most commonly FBU is chosen as a model for ToM and used in order to examine relations. Many different precursors of ToM and also later explicit FBU has been associated with MSL, but the relation to nonverbal tasks such as implicit FBU have not been investigated yet. It is important to examine whether there is a significant relationship between children's implicit FBU and their concurrent MSL production as FBU is considered to be one of the gold standards for ToM. This would support an association between a conceptual understanding of mental states and the appropriate mental state terms to express them and thereby speak for both the simulation account as it assumes concepts to be present and the continuity account as preverbal abilities would be associated with the verbalization of these abilities. However, as the validity of implicit FB is currently under debate, it is important to have an alternative model for preverbal and subsequent explicit MSU. Level-1 perspective-taking can serve as such a model and therefore help to determine the role that MSL may play in the transition between preverbal (implicit) and verbal (explicit). Furthermore, investigating the concurrent relationship between verbal level-1 perspective-taking and MSL production makes it possible to replicate the conceptual relation between the tasks (as has been found by Chiarella et al. (2013)) before the age of 20 months, or whether relations are strongly dependent on general language abilities and only become language-independent later on. As cognition terms are expected to be the best predictors of MSU, the question remains whether it is cognition terms that drive relations to other ToM relevant abilities early on or whether also physiological, emotional and volitional terms matter for development.

As mentioned above, MSL may be considered a representational resource for the emergence of explicit ToM mastery, however, MSL may be dependent on several developmental factors itself that have not been investigated previously. MSL depends on vocabulary acquisition, but the relation of mental terms and vocabulary in general has rarely been studied. Similarly, mental state terms are associated with complex syntax (e.g. complement sentences), but children begin to use mental state terms before they acquire a full syntactic frame of superordinate and subordinate clauses. A more basic set of competences that can be measured before complement syntax are language comprehension and production and in particular sentence length. The role they play in the development of MSL production is yet to be explored. Thus, the third research question that should be addressed is how MSL production relates to early basic abilities, among them general language and executive functions. Several studies have emphasized the importance of general language and in some cases inhibition skills, but all relations have been cross-sectional until now and it was not possible to determine which early abilities predict later MSL production. Investigating predictors of MSL production might make it possible to foster early interventions and thereby improve children's subsequent MSL production and possibly also their ToM skills.

In order to answer the questions outlined above, this thesis reports a large-sample longitudinal study in which children were assessed at two time-points, once at the age of 24 months when various aspect of general language, cognitive development and inhibition were measured, and another time at the age of 27 months where their mental state language production, perspective-taking abilities, metacognition of ignorance, implicit FBU and pragmatic inferences from information on what a speakers knows vs. wants were measured (Study 1). The latter issue was followed up by an additional cross-sectional study with children at the age of 2-, 3- and 5-years of age (Study 2). In the following, the research questions and hypotheses shall be presented in detail.

2.5.1 Research Questions and Hypotheses

The core aims of this thesis are the following, 1) to determine whether and to what extent 2-year-olds' usage of MSL reflects their MSU, 2) to address the gap between early signs of ToM in preverbal communication and later mastery of explicit ToM tasks by examining production of MSL and its relationship to the concurrent development of preverbal and verbal level-1 perspective-taking and implicit FBU, and 3) identifying which basic linguistic and cognitive abilities contribute to the development of MSL. More specifically, the present thesis aims to illuminate the following research questions:

1. How is young children's production of MSL related to measures of their comprehension of MSL and mental states themselves?

    a) Is the production of mental state terms a sign of children's comprehension of the semantics and pragmatics of these terms and thereby MSU? This question includes three sub-questions. First, do children who produce mental state terms according to a parental questionnaire, also produce them in an appropriate situational context i.e. to indicate their own lack of knowledge by responding with 'I don't know'? Second, do children who produce the terms 'want' and 'know' understand the pragmatic inferences that are associated with a speaker wanting/not wanting something and knowing/not knowing something? And third, are children who produce the terms 'know' and 'don't know' capable of using these terms to answer questions about their own knowledge and that of others in an epistemic perspective-taking task?

b) At which age are children capable of drawing correct pragmatic inferences to inform an agent based on their usage of the terms 'want' and 'know' and thereby demonstrate a reliable association between mental state verbs and the respective concepts?

2. What is the role of MSL production for the transition between preverbal MSU and the mastery of explicit ToM tasks? Specifically, does MSL relate to both nonverbal and verbal aspects of MSU and thereby demonstrate continuity in concepts (continuity account) and at the same time suggest presence of mental state concepts preverbally (ST) or does it only show a relation to verbal MSU and thus, suggest preverbal understanding to be only due to production rules (TT)? Overall, the following sub questions are included:

   a) Does MSL production relate to verbal level-1 perspective taking independent of language and demonstrate additional conceptual continuity beyond the acquisition of language? A general language independent relation would support the view of MSL as the first step towards explicit concepts of mental states.

   a) Does MSL production relate to preverbal MSU in the form of implicit FBU or preverbal level-1 perspective taking (a simpler model)? A relation would support continuity and the presence of concepts in preverbal MSU.

3. Which language and cognitive abilities predict MSL competence early in the third year? Are the contributing factors for MSL production the sames ones as for MSU later in life (i.e. executive functions and language)?

Approach

In order to address the questions above, the following studies were conducted:

a) Study 1 is a longitudinal study that examined the relationship between basic cognitive and linguistic abilities at the age of 24 months and their relationship to later MSL production at the age of 27 months, in addition to the concurrent relationship betweenMSL production, perspective-taking and implicit FBU. Finally, it also investigated the relationship between MSL production and children's comprehension of mental state terms. (Questions 1a, 2 and 3)

b) Study 2 is a cross-sectional study that examined the age at which children are able to draw correct pragmatic inferences from an agent using the terms 'know' and 'want', by administering a newly developed task from study 1 to 2-, 3- and 5-year-old children. (Question 1b)

Hypotheses

1. Relation between young children's MSL production and measures of MSL comprehension:

   a) Because Harris, Ronfard, and Bartz (2017) have argued that children not only use mental state terms like 'know' early on, but also use them in appropriate contexts, it was expected that 27-month-olds would show signs of their own ignorance and use the formulation 'I don't know' to indicate it. In addition, in case the production of mental state terms shows children's comprehension of mental states, it was expected that children with an extensive productive MSL vocabulary would be more likely to use the formulation 'I don't know' to express their uncertainty.

Furthermore, based on studies on preverbal communication that suggest children are able to draw correct inferences from an agent's knowledge states (O'Neill, 1996) and newer studies suggesting that children use mental state terms early and in appropriate contexts (Bartz, 2017; Harris, Ronfard, & Bartz, 2017; Harris, Yang, & Cui, 2017), it was hypothesized that children who are reported to produce mental state terms in the mental state language parental questionnaire, would also perform better on an experimental task that measures children's pragmatic comprehension of 'want' and 'know'. The main idea was that children who are able to produce the terms adequately, should also be able to draw appropriate behavioral inferences from the terms, especially if they already draw the same inferences preverbally, prior to production of the terms.

Additionally, if children's usage of mental state terms reflects their comprehension of mental states, it was expected that children who have more extensive MSL vocabulary would perform better on verbal EPT tasks that use the mental state verb 'know', compared to children with a less extensive MSL vocabulary. If children already produce the word 'know' adequately and it is associated with a mental representation of knowledge states (a concept), it is feasible that they should be able to report their own knowledge. However, since avoce chance performance in the implemented task was only observed at the age of 4 (Gonzales et al., 2018), weak associations without general above chance performance were expected.

b) Concerning factivity, children are only able to differentiate between 'know' and 'think' around 4-5 years of age (Kristen-Antonow et al., 2019; Moore et al., 1989) and EPT tasks show competence in verbally reporting the knowledge state of another person around 4 years of age (Gonzales et al., 2018). However, as children are able to draw similar inferences non-verbally before the age of 2 (O'Neill, 1996) and drawing pragmatic inferences can be considered easier than assessing factivity, accurate performance in drawing pragmatic inferences from the mental state verb ‘know’ was expected to manifest itself before the typical age of 4-5. If children are indeed capable of drawing the right inferences from ‘know/don’t know’ around the age of 2, then MSL production can be argued to be a valid sign of children possessing functional concepts of mental states which aid them in inferring behavior. However, if children are only able to infer an agent’s need for information much later than they produce the associated mental state terms, it is difficult to argue for young children possessing functional representations of mental states as soon as they acquire the terms.

2. Concurrent relationship between MSL and preverbal and verbal MSU:

    Given that MSL production is argued to be a first indicator of explicit understanding of mental states, this question aimed at addressing the role that MSL plays for early implicit ToM abilities and later explicit ToM abilities. While the ST account argues that children possess early concepts of mental states when they solve preverbal tasks, the TT presupposes that children use production rules which appear similar to concepts but only work for a limited set of predictions. Only later when children acquire metarepresentations they have concepts and are able to solve explicit tasks.

    Considering preverbal MSU, an association with MSL production would be more likely to suggest the presence of a concept, as words would only need to be associated with a conceptual structure. In the case of production rules, separate associations for each production rule appear more likely as they are not part of a common concept and require separate verbal highlighting of perspectives. If no relation to preverbal measures, but only a relation to verbal measures is present, then one might argue that preverbal MSU does not require conceptual understanding and only MSL enables the emergence of a conceptually-based ToM because language is needed to highlight the presence of mental states.

a) MSL production was expected to be associated with verbal level-1 perspective-taking, namely, understanding whether someone sees or does not see something. Because a previous study by Chiarella et al. (2013) found a relationship between MSL production and level-1 perspective-taking in 30-month-olds, independent of general language, similar results with 27-month-olds were expected. Furthermore, as the relationship was not specific to MSL categories, a category-general relationship between the two variables with perception and physiological terms correlating to the same degree as cognition terms was expected. An early concurrent relationship between these two variables might suggest continuity in the mental domain, with verbal comprehension of perspectives relating to usage of verbs denoting mental states (e.g. seeing).

b) As a preverbal measure, action-based level-1 perspective-taking and implicit FBU were investigated in relation to MSL production. Implicit FBU is commonly considered a gold standard for ToM, but is at the same time rather unreliable. Therefore, both FBU and a simpler model, perspective-taking, which shows both preverbal and verbal competence in the third year of life were chosen. Children were expected to perform better on the preverbal perspective-taking task, than the verbal perspective-taking task, as Moll and Tomasello (2006) have found competence in 24-month-olds. Given that MSL production was found to relate to a verbal perspective-taking task, it was hypothesized that the same relationship may be present for a preverbal task if MSL production is indeed a representational resource for early MSU. For implicit FBU hypotheses were as follows: Similarly to preverbal level-1 perspective taking, a relationship between implicit FBU and MSL production was expected in case implicit FBU could be replicated and showed above chance performance.

The presence of a relationship between preverbal perspective-taking, implicit FBU and MSL production would imply continuity between preverbal competence and MSL production, whereas the lack of a relationship leaves open several possibilities, among them that preverbal ToM does not involve MSU, that preverbal ToM is pre-conceptual (TT) or that it is indeed conceptual (ST), but children at that age have not yet developed the appropriate parts of their mental lexicon.

3. Basic cognitive and linguistic predictors of MSL production:

   As language and executive functions have been found to play an important role in ToM development, it is to be expected that these basic abilities influence MSL production in a similar way. Based on earlier studies that have found a relationship between inhibitory control and MSL (Bellagamba et al., 2014), inhibitory control was expected to be a relevant predictor. Furthermore, as language is a relevant precursor of mental MSL (Kristen et al., 2011), language production was expected to be the main linguistic precursor for the emergence of MSL production. It is possible that a particular size of vocabulary is necessary for children to begin using mental state terms, or that only when they can speak in 2-word sentences they have the basis to speak of knowing or wanting something. There has been no prior research on the importance of general cognitive skills or fine motor skills, but it was expected that cognition might play a bigger role for MSL production than fine motor skills and serve as a control for the necessity of a general developmental stage before mental state terms can be acquired.

   See Figure 2.3 for a brief summary of hypotheses.

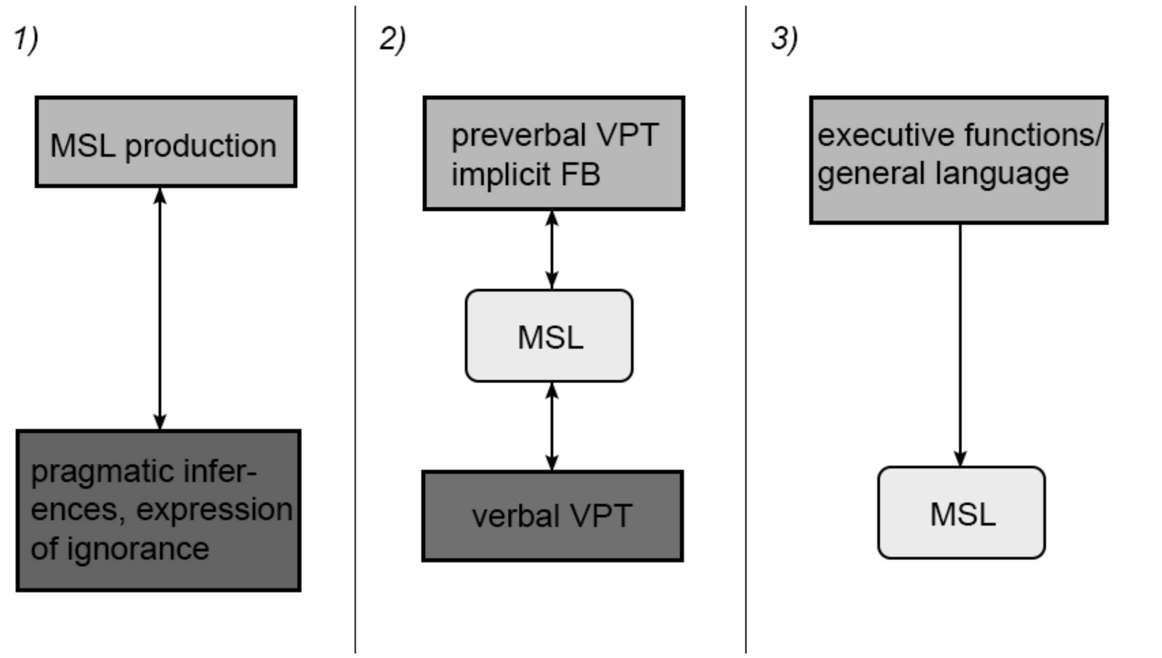

Fig. 2.3: Overview of the three main hypotheses. *1)* Mental state language production (MSL) is related to children drawing the right pragmatic inferences and expressing their ignorance *2)* preverbal visual perspective-taking (VPT) and implicit false-belief (FB) is related to MSL, which in turn is related to verbal VPT *3)* executive functions and general language at 24 months serve as predictors for MSL at 27 months.

3. STUDY 1: EARLY PREDICTORS OF MENTAL STATE LANGUAGE PRODUCTION AND THE RELATIONS BETWEEN MENTAL STATE LANGUAGE PRODUCTION AND MENTAL STATE LANGUAGE UNDERSTANDING AT THE AGE OF 27 MONTHS

3.1 Objectives and Hypotheses

Study 1 had three objectives, as outlined above. The first objective was to investigate whether children's production of mental state terms corresponds to understanding mental states, i.e. usage of the word 'know' goes along with conceptual understanding of 'knowing'. This question was addressed via three sub-questions that focused on different aspects of mental state term comprehension. First of all, it was assessed whether children who produce mental state terms are also able to use them appropriately in contexts where they denote a mental state. Second, it was investigated whether children are able to draw inferences about a speaker's knowledge state and desires from their usage of the terms 'want' and 'know' and whether this ability is related to their mental state verb production. Third, it was investigated whether children who produce mental state verbs are also better at EPT, more specifically, reporting their own and someone else's knowledge. It was hypothesized that children who produce mental state terms, in particular 'know', would be able to use mental state terms in appropriate contexts and draw the correct pragmatic inferences and additionally be more likely to show competence in verbal EPT.

A second objective was to determine the role that MSL plays for a transition from preverbal ToM to verbal ToM by investigating the concurrent relationship between MSL production, level-1 perspective-taking and implicit FBU. Theoretical accounts differ in their predictions on whether children already possess a conceptual understanding of mental states preverbally (ST) or whether a conceptual understanding only develops later on (TT). Examining the relationship between MSL and preverbal and verbal competences allows to shed light onto that question. As shown by a previous study that suggested a relationship between MSL production and verbal perspective-taking (Chiarella et al., 2013), a language- and category-independent correlation between the two abilities was expected. This would demonstrate continuity between the production of mental state terms and verbal MSU in the realm of seeing in the epistemic sense. More importantly the relationship between preverbal competences and MSL production was examined, in particular with implicit FBU and preverbal level-1 perspective-taking which served as a simpler model for preverbal MSU. It was hypothesized that there would be a significant relationship between children's MSL production and their performance on a preverbal perspective taking task, indicating continuity between preverbal concepts and the verbal terms associated with them. In the case of implicit FBU a similar relationship was expected in case the tasks showed reliable performance.

A last objective was to identify basic cognitive and linguistic skills that are relevant for the emergence of MSL production. For later ToM skills, executive functions and language have already been shown to be crucial, but the role that these basic skills play in MSL production has not been investigated systematically. Executive functions, in particular inhibitory control, have been argued to be relevant in addition to general language skills, however, which of these is particularly relevant and to what degree has not been addressed yet. It was hypothesized that especially language production and inhibitory control would be important predictors. Concerning cognitive and fine motor skills, it was assumed that cognitive skills would play a bigger role and function as a control measure as children might need to reach a certain developmental stage before they are able to acquire additional skills such as MSL.

3.2 Methods

3.2.1 Participants

For this study, 160 monolingual children (77 male) were recruited from local birth registers (24 months: M_{age} = 24;12 months, range = 23;19 – 25;26; 27 months: M_{age} = 27;07 months, range = 26;06 – 28;04). Thirty-eight additional 24-month-old children were recruited, however, 11 of those did not participate in the 27 months measurement time point due to personal reasons, 8 moved or changed their address and could not be reached and another 19 showed delay in language development at both time points (24 and 27 months) according to the criteria set by the SETK-2 (Grimm & Aktaş, 2000), and were therefore excluded.

All participating children came from the lower- to upper-middle class in an urban to suburban region in the south of Germany. Mothers' average age was 36.1 years (N = 156, SD = 4.2, range: 24–48 years) and 4.5% of mothers had a lower secondary level degree comprising 9 years, 5.8% a lower secondary degree comprising 10 years, 15.4% had passed the Abitur examination (equivalent to A levels), 73.7% a university degree and 0.6% had completed their doctoral degree. Fathers were on average 39.3 years old (N = 155, SD = 6.0; range: 29-63 years) and 3.9% had a lower secondary level degree comprising 9 years, 12.9% a lower secondary degree comprising 10 years, 11.6% had passed the Abitur examination (equivalent to A levels), 71.0% a university degree and 0.6% had completed their doctoral degree. Overall, 46.8% of the children had no siblings at the time of the study, 44.9% one sibling and 8.3% had more than one sibling (% out of$N$= 156). In the case of 4 children, socio-economic data was not assessed due to parents not completing the questionnaire.

The study was conducted according to the ethical standards for conducting experiments involving humans as stated in the 1964 Helsinki declaration and its later amendments and comparable ethical standards and was approved by the local ethics committee. Participants were given monetary compensation for travel costs and children received gifts for participation at each measurement time point. For the testing at 24 months, children were gifted a picture book and at 27 months a finger puppet.

The tasks were embedded in a longitudinal study that extended beyond 27 months until the age of 36 months. Following best practice recommendations (Simmons, Nelson, & Simonsohn, 2012), I report how the sample size, all data exclusions, all manipulations and all measures in the study were determined. An apriori power analysis using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) showed that a sample size of n = 46 would be sufficient to detect a simple correlation ($r = 0.4$; $\alpha = .05$; $(1 - \beta) = .8$). As a drop-out rate of about 20-30% is common for longitudinal designs, the total sample size was set for N = 70. However, as the study consisted of two parts and an additional training group with N = 80 was planned, the total intended sample size for the first two testing points was set for N = 150. See Table 3.1 for the number of included children for each task. Details about measures, manipulations and reasons for data exclusions are noted in the descriptions for the respective tasks.

3.2.2 Procedure and Design

Children participated in two test sessions, one at the age of 24 months, in which control variables were assessed, namely general language abilities (SETK-2, Grimm, 2000), cognition and fine motor skills (Bayley Scales III, Reuner & Rosenkranz, 2014), and executive functions in the form of learning and inhibitory control (Kovács & Mehler, 2009), and one at the age of 27 months in which key variables were assessed, namely visual and epistemic level-1 perspective-taking abilities (Gonzales et al., 2018), action-based level-1 perspective-taking (Moll & Tomasello, 2006), metacognition of own ignorance (Bartz, 2017), MSL vocabulary and production (MSLQ, Olineck & Poulin-Dubois, 2005), understanding of the pragmatics of 'know' and 'want', and implicit FBU (Grosse Wiesmann et al., 2017). Upon arrival at the research laboratory for the first session, children were given a short warm-up in which the experimenters showed every child a picture book. For the second session at 27 months, children and experimenter examined stickers as a warm up. Performance was coded online and double checked via video recordings of the sessions by two independent coders. Inter-rater

Tab. 3.1: Number of included children for each measurement point and task.

Measurement Point	Task	*N*
24 months	General language	Comprehension: 149 Production: 146 Both: 145
	Cognition and fine motor	157
	Learning and inhibition	Pre-switch: 118 Post-switch: 100
27 months	Know-Want task	Want: 104 Know: 102
	Visual and epistemic perspective-taking	VPT Self: 76 VPT Other: 72 EPT Self: 74 EPT Other: 118
	Action-based perspective-taking	Pointing: 72 Giving: 60
	Metacognition of ignorance	136
	Mental state language production	147
	Implicit false-belief understanding	FB1: 148 FB2: 145

Note. VPT = visual perspective-taking; EPT = epistemic perspective-taking, 'level-1' was omitted for brevity.

reliability is reported for each task separately.

3.2.3 Statistical Analysis

All major statistical analyses have been performed with SPSS 25 (IBM Corp. Released 2019. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.). In the case of the implicit FB task, part of the analysis was also conducted in R (R Core Team, 2017) as in the original study. In order to correct for multiple comparisons where appropriate the false discovery rate (FDR) (Benjamini & Hochberg, 1995) was used as it is less conservative than the Family Wise Error Rate (FWER) approach, provides more power the higher the number of non-null hypotheses and is broadly applicable (Benjamini & Yekutieli, 2001). The FDR is defined as the expected proportion of falsely rejected hypotheses. Adjusted p-values according to the FDR are denoted as $*q$.

3.2.4 Tasks at 24 months

General language at 24 months

The SETK-2 is a German standardized test for language abilities in children between 24 and 35 months of age (Grimm & Aktaş, 2000). It consists of two subscales, comprehension and production, which can be split into 'words' and 'sentences'. The reliability lies between .88 and .95 for the productivity scale, and around .70 for the comprehension scale, indicating lower reliability. The test has a high construct validity as can be concluded from the high intercorrelations of the different scales (.49 to .86). For the administration of the measure, the experimenter was seated across from the child and introduced the task by saying the following: "I have brought a few pictures for you. I am going to say a word and you point at the correct picture. Only one picture is correct. Let's try it out!" The child was then shown cards with four objects or animals depicted on each of them. The experimenter uttered a word denoting an object or an animal and the child was then asked to point at the corresponding picture, e.g. 'cow' among a bunny, a dog, a duck and a cow (comprehension: words). The same was repeated for sentences with four complex scenarios in which the child had to choose the corresponding one e.g. 'The pen is in the cup' with pictures which depicted the pen inside, outside, on top and next to the cup. In the production subscale, the child was presented with a small bag that contained 6 objects: a puppet, a pen, a book, a key, a ball and a knife. The child was asked to name each of the objects. The subtest was introduced with the words "I have brought my surprise bag for you, let's see what's inside!" The second part consisted of drawn pictures of items like a car, a pear or more complex ones such as a fridge. Again, the child was asked to name them. Finally, the child was shown pictures of scenes and asked to describe them to assess sentence production, for example a horse standing on a table. According to the SETK-2 manual there are two ways of coding sentence comprehension, one involving further questions for the verb and the subject of sentences and another which only allows the question "What can you see on this picture?" The simpler version without any further questions was used. All words that were grammatical and corresponded to the items on the presented pictures were counted with one point. In order to determine if performance was critically below the expected norm, DAWA scores were calculated by adding up all valid words for the sentence production measure and dividing them by the total number of sentences (16). Children who received a score below .39 were considered to be below the norm. For the comprehension scale it was possible to obtain a total of 17 points (9 for words and 8 for sentences) and for the production scale there was no limit as children were free to describe the pictures (30 for words and unlimited for sentences). Overall, 11 children refused to participate in the task all together and an additional 2 children refused to participate in the production part of the task. Two independent raters coded the performance and agreed on 98% of the cases. Cases in which the raters disagreed were discussed to reach agreement. For subsequent calculations, the raw scores were used to compute a score for language comprehension, production and a total score as the sum of both the comprehension and the production scores.

General cognitive and fine motor skills at 24 months

The Bayley Scales III are a standardized measure to examine development on the individual level (Reuner & Rosenkranz, 2014) and include three scales: language, cognition and motor skills. The language scale is divided into receptive and expressive language and the motoric skills scale is divided into gross motor skills and fine motor skills. The test is applicable to children between the age of 0;16 and 42;15 months (months;days). In the current study, the screening version of the Bayley Scales, which is shorter than the full scale, was used and only the cognition and the fine motor scale were administered. Since the SETK-2 was used as a separate standardized language task, there was no need for the language measures of the Bayley and fine motor skills were considered sufficient as a control measure of motoric abilities. Reliability for the screening measure lies between r = .68 and r = .83. The advantage of the screening procedure is, among others, time as it only takes 30 minutes for its full application. Additionally, one is able to identify possible delay and obvious delay in the participating child for each subscale. On the cognitive scale, 34 items are administered in the form of playful tasks. Among others, information processing, color identification and play are measured. On the fine motor scale, 28 items are administered, among them holding a pen, building a tower out of small blocks and grabbing small objects with index finger and thumb. For each solved item the child received one point, which was then added up into a total score per scale. For the cognition scale a value below 29 is considered 'at risk' and a value below 25 as critical. For the fine motor scale a value below 22 is considered 'at risk' and a value below 19 as critical. However, none of the children were excluded based on this categorization, since the screening test is known to identify children as 'at risk' that do not have developmental delay and since the Bayley scales were only used as a control measure. Two independent raters coded the performance from video recordings and agreed on 90% of the cases. Cases in which both raters disagreed were discussed to reach agreement. For both the cognition and the fine motor scale the raw scores were used for subsequent computations.

Learning and inhibition at 24 months

Children's learning and inhibition skills were measured with the help of the so-called *'SWITCH' task*, developed by Kovács and Mehler (2009). The SWITCH task is an eye-tracking task that was developed to measure executive functions in the form of learning patterns as well as the subsequent inhibition of these patterns. In this task, children were taught the association between a cue and an event. After learning the first association, they observed a switch after which the previously learned cue triggered a different event. The main measure of this task was how quickly children learned the first association and how quickly they were able to inhibit it for learning the second association.

Task

The task was split into two blocks divided in 9 trials before the switch (pre-switch phase) and 9 trials after the switch (post-switch phase). The duration of each trial video was 5200 ms. Each trial was built up as follows: Two white squares were displayed on the left and right of the screen while a colorful fixation cross served as a visual attractor which was displayed for 500 ms until an auditory cue was presented. The offset of the cue was followed by 1000ms of silence during which only two white squares were shown. Subsequently, a reward appeared for 2000 ms until it disappeared from the screen and only the boxes remained. The reward was a puppet that increased in size and was emphasized by the sound of a bell. The side of the reward was kept constant across trials for each phase, switching for the post-switch phase. Apart from the side on which the reward appeared, the pre- and post-switch phases differed in the structure of the auditory cue. On which side the reward appeared for the two phases was counterbalanced across children. See Figure 3.1 for a display of the task.

Cues

The auditory cues used were trisyllabic words in which one syllable was repeated twice. Before the switch the syllables repeated in an AAB pattern, whereas after the switch the pattern changed and turned into an ABB pattern. All cues were created from individual syllables without combined meaning and were presented for 1700 ms.

Procedure

Children were seated on a car seat in front of a Tobii T60 Eyetracker with approximately 60cm distance to a 1280x1024 pixel screen. Parents were sat next to them with occluding glasses on or remained further away on a couch if possible. Throughout the task, parents were instructed not to interact with their children with the exception of returning their attention to the screen. The videos were presented and recorded using Matlab (Version 2015b, The MathWorks Inc), the Psychtoolbox and the Talk2Tobii toolbox. Prior to the experiment, children's eye-gaze was calibrated with a five-point calibration involving a small cartoon figure that wiggled accompanied by a ringing sound and moved from one point to the next. If necessary, calibration was repeated until each calibration point had sufficient gaze information. After calibration a colorful attention grabber was shown on the screen until the experimenter confirmed with the space bar that the child was attending to the screen. This was done for each trial to avoid data loss due to the child not paying attention to the screen.

Processing and Analysis

The eyetracking sampling rate was set to 60Hz and the child's left eye-gaze coordinates were analyzed. Missing eye-gaze data within each trial was linearly interpolated if possible. Data was analyzed based on two rectangular Areas of Interest (AOIs). The first AOI covered the left third of the screen whereas the second AOI covered the right third, corresponding to the two white boxes but also including the area below and above. If children's gaze was interrupted for less than 200ms and disappeared from an AOI, it was nevertheless coded as within the AOI, as such interruptions are more likely due to signal loss and eye-tracking errors than refocusing of attention. Children's performance was scored by identifying the time window of interest, namely the 1000 ms interval beginning 150ms after the end of the auditory cue and ending 150ms after the appearance of the reward. The interval began before the onset of the reward to analyze saccades that occurred in anticipation of the reward. Children received a score of 1 if they looked to the side where the reward was going to appear and a score of 0 if they looked at the side of the screen without a reward. The pre-switch and post-switch phase were split into two blocks each to differentiate between immediate learning within the first few trials and subsequent learning after the second half of the trials. Block 1 consisted of trials 2-5 and block 2 consisted of trials 6-9. The first trials of each phase were not considered for analysis as it was not yet possible for children to learn before observing the relation between cue and reward. Blocks were also split for the post-switch phase, in order to differentiate between children who inhibited their response immediately in the first half of the post-switch trials vs the latter half. Children received a score with a proportion of anticipatory looks for the trials before the switch and the trials after the switch separately. Overall, out of 160 children, 31 did not participate in the task due to technical difficulties or refusal. An additional 11 did not provide valid gaze data for any of the pre-switch trials, while 10 did not provide valid gaze data in any of the post-switch trials. Further, only children who did not score 0 on the pre-switch trials were included for the evaluation of the post-switch trials as these children did not learn the association between cue and reward in the first place and did not have a learned association to inhibit ($N = 18$). This resulted in a total of 118 children for the pre-switch phase and 100 children for the post-switch phase. The number of children with valid trials in each of the sub blocks differed. In order to keep the variance between children who learned quickly and inhibited quickly and children who learned slower and needed time to relearn the new strategy, both blocks are used separately in subsequent analyses. Coding and analysis were conducted in Matlab.

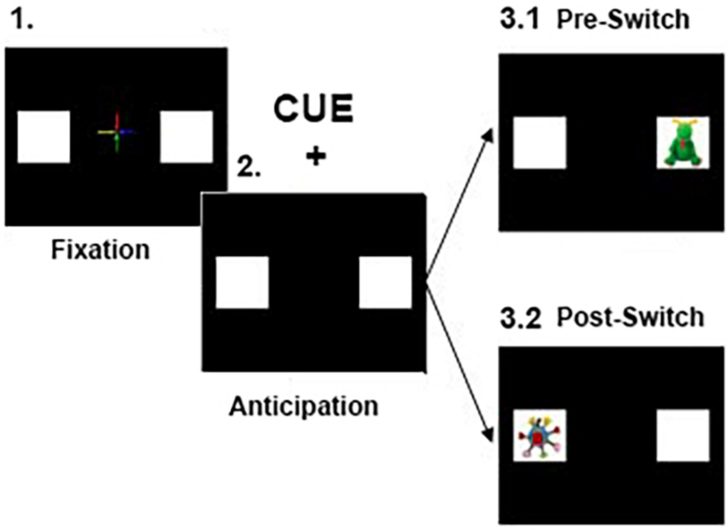

Fig. 3.1: Sequence of the SWITCH task: 1. Fixation period with a colorful and moving fixation cross. 2. Addition of auditory cues and anticipation of reward. 3.1. Appearance of reward on one side during the learning phase (pre-switch) and 3.2. Appearance of the reward on the other side during the inhibition phase (post-switch).

3.2.5 Tasks at 27 months

Know-Want task at 27 months

The *Know-Want task*[1] was used to measure children's pragmatic comprehension of the terms 'want' and 'know' in the form of a hiding game. Two experimenters took part in the task with the first experimenter (E1) sitting at a table across from the child and a second experimenter (E2) kneeling at the edge of the table.

In a short familiarization phase, E1 asked E2 to identify a book and a plate by saying "Point at the book. Can you help me?" After E2 answered with "yes" and pointed at both objects, the same procedure was repeated with the child. All children identified the book and the plate correctly.

Before the beginning of the test phase, E1 arranged four colorful boxes (8cm x 15cm) at equal distance in front of the child. The order from left to right was red box, green box, white box and blue box. In the test phase, E1 showed two objects to

[1] See 4.2 for a more elaborate description of the task. Since study 1 comprises several tasks, details that are necessary for a replication were included in study 2.

the child and asked her[2] to name them. The first two objects were a ball and chair, followed by a bird and a cup, a dog and a car, and a toy figure and a spoon. See Table 3.2 for the object box allocation. E2 was present during the introduction of the objects. If the child used a different name for one of the objects e.g. 'boy' for the toy figure, E2 adapted the name during the rest of the trial. After both objects were named, E2 left the room. E1 hid one object each in one of four boxes in front of the child. Upon return, E2 either requested the object of which she did not know the location, "I know where the ball is, I do not know where the chair is" or the object that she desired, "I want the car, I do not want the dog". There were two *want* trials and two *know* trials. For each verb, there was a trial in which the negation was mentioned first (i.e. "I do not want the ball") and one trial in which the assertion was mentioned first (i.e. "I want the chair"). The position of negation and the order of *want* and *know* trials was counterbalanced in four orders over all children. See Figure B.2 in the supplementary material for an overview.

Children's performance was rated with 1 in the *want* trials if the child handed the experimenter the box with the object desired by E2 and with 0 when the other box with an object was chosen. For the *know* trials, performance was rated with 1 if the child handed E1 the box that contained the item E2 did not know the location of and 0 if the box with the other object was chosen. Cases in which the child chose an empty box were coded as -1, but not included in the final analysis of task performance. Choosing a distractor box indicated either a deficit to remember the location or failure to understand the task. Children had to pass both trials for each condition to be rated as competent in the respective condition. Children who were missing one trial for a condition did not receive a final score as a single trial was not considered sufficient to assess performance and were therefore coded as missing. Overall, 4 children were excluded because they refused to participate in the task and 10 children were excluded for refusing to reply in one *want* trial and 7 children for refusing to reply in one *know* trial. Additionally, 28 children were excluded for choosing an empty box in the *want*

[2] The female pronoun will be used inclusively for males and females.

Tab. 3.2: Know-Want task: Object-box allocation and order of appearance.

Trial				
Trial 1		**Ball**	Chair	
Trial 2	Bird		**Cup**	
Trial 3		Dog		**Car**
Trial 4	**Toy Figure**			Spoon

Note. The order of object appearance and hiding location remained constant. Object marked in bold was introduced first.

trials and 33 children for choosing an empty box in the *know* trials. A total of 14 children were excluded as they chose empty boxes for both conditions. Overall, 104 children remained for the *want* condition and 102 children for the *know* condition. Performance was rated by two independent raters that agreed on 100% of the cases. Scores used for further analyses were the total *know* performance score and the total *want* performance score.

Visual and epistemic perspective-taking at 27 months

Visual and epistemic perspective-taking were measured with the help of the *Knowledge-Perception task* developed by Gonzales et al. (2018). It consists of two different measures, namely perception (visual perspective-taking) and knowledge (epistemic perspective-taking). Each of those measures was split into two conditions, the *self* and the *other* condition. In the *self* conditions, the child was asked to answer questions concerning her own perspective and in the *other* conditions about the perspective of a puppet that was seated on the experimenter's lap. Experimenter and child were seated at a table across from each other. See Figure 3.2 for an overview of all conditions.

Visual perspective-taking. In the perception measure, the child had to differentiate between an item that was in plain sight and another item that was hidden behind an occluder, both either for themselves (*self* condition) or for a puppet on the experimenter's lap (*other* condition). In the *self* condition, the experimenter showed the child a toy cat and asked the child to name it, followed by a toy dog. This served as a control to make sure the child was able to identify both animals correctly and would recognize the labels when they were used later in the task. In case the child named the animals differently, the experimenter either adapted the child's names if they were consistent, or made sure the child learned the correct terms by repeating them. After this familiarization, the experimenter placed an occluder in front of the cat, such that it was still visible to her but not to the child with the words: "I will put this here and now I will ask you what you can see." She then said: "Say yes if you can see it and no if you can't see it." Thereafter the test questions were asked, firstly "Can you see the cat? Say yes if you can see it and no if you can't see it." and then "Can you see the dog? Say yes if you can see it and no if you can't see it." In the *other* condition, a new set of toy animals was introduced, a pig and a horse, followed by a puppet named Martin. After the child named both animals as above or learned the terms, the experimenter placed an occluder in front of the pig, such that it was still visible to the child, but not the experimenter and the puppet. The occluder was shown with the words "I will put this here and now I will ask you what Martin can see. Say yes if he can see it and no if he can't see it." Then the test questions were asked, regarding the perception of the puppet. "Can Martin see the pig? Say yes if he can see it and no if he can't see it." followed by "Can Martin see the horse? Say yes if he can see it and no if he can't see it." In case the child did not respond to the test questions, a prompt was given that changed the order of affirmation and denial, "Say no if you/he can't see it and yes if you/he can see it."

Epistemic perspective-taking. In the knowledge measure the child had to differentiate between having knowledge about the content of one box and having no knowledge about the content of another box, once for their own knowledge (*self* condition) and once for a puppet's knowledge (*other* condition). In the *self* condition the experimenter showed the child two boxes, a yellow one and a brown one with the words: "I will show you what is inside of the yellow box, but I will not show you what is inside the brown box." Following the introduction, the experimenter opened the yellow box and showed the child a coin that was located inside, asking "What is inside?" Once the child answered, the yellow box was closed and returned to the table. Again the child was asked about the content of the yellow box to ensure that they remembered. Then the test phase began. The experimenter said, "Now I will ask you if you know what is in each of those boxes, say yes if you know and no if you don't know." For each box the experimenter asked whether the child knew the contents, using the prompt: "Do you know what is inside the yellow box? Say yes if you know and no if you don't know." The same was done with the brown box. For the *other* condition a puppet named Kathi was introduced, followed by a pink and a grey box, "This is Kathi. I am going to show Kathi what is inside the pink box, but I will not show her what is inside the grey box." After opening the lid of the pink box, such that the child was not able to look inside, the experimenter pretended to show the puppet the contents and confirmed by speaking for the puppet: "I see what is inside the box." After showing the puppet the contents, the experimenter closed the box and returned it to the table. The test questions were introduced by the following: "I will now ask you if Kathi knows what is inside each of those boxes. Say yes if she knows and say no if she doesn't know." For each of the boxes the child was reminded of whether Kathi had looked inside the box or not and asked the test question, "Kathi has looked inside the pink box. Does Kathi know what is inside the pink box? Say yes if she knows and say no if she doesn't know." And "Kathi has not looked inside the grey box, does Kathi know what is inside the grey box? Say yes if she knows and no if she doesn't know." In case the child did not respond to the test questions, a prompt was given that changed the order of affirmation and denial, "Say no if you/she don't/doesn't know it and yes if you/she know/s it."

The order of knowledge and perception trials was varied across children in addition to the order of the *self* and *other* conditions. Furthermore, it was varied whether the experimenter started with 'say yes' or 'say no' to control for order effects. Overall, there were four possible randomizations of the task.

Scoring

A child passed a condition if she answered both test questions correctly, resulting in a score of 1 for the respective condition. Out of the originally tested 160 children, 12 refused to participate in the task altogether, 2 refused to participate in the perception conditions and 9 refused to participate in the knowledge conditions. In addition, children who did not answer both control questions correctly were excluded from analyses. Overall, 6 children failed to answer both control questions in the perception *self* condition, 23 in the perception *other* condition and 33 in the epistemic *self* condition. As there was no control question for the epistemic *other* condition, no children were excluded. Further, children who did not answer one of the test questions were not included in the analysis. Overall, 39 children did not provide an answer to one of the test question in the perception *self* condition, 22 in the perception *other* condition, 15 in the knowledge *self* condition and 21 in the knowledge *other* condition and were therefore excluded from further analyses. An additional 25 children were excluded in the perception *self* condition due to glancing over the wall during the test question and an additional 29 were excluded in the perception *other* condition due to children removing the barrier before replying to the test question, being unfamiliar with the introduced animals or experimenter errors in prompts. Lastly, 17 children were excluded for the knowledge *self* condition due to experimenter errors in naming the content of the box or not ensuring the child was able to name the content of the box (i.e. money). Overall, the remaining sample size was 76 children for the perception *self* condition, 72 children for the perception *other* condition, 74 children for the knowledge *self* condition and 118 children for the knowledge *other* condition. Performance was rated by two independent raters that agreed in 93% of the cases. Discrepancies were discussed among the raters and resolved. Four scores were used in total, a visual perspective-taking self score (VPT Self), a visual perspective-taking other score (VPT Other), an epistemic perspective-taking self score (EPT Self) and an epistemic perspective-taking other score (EPT Other).

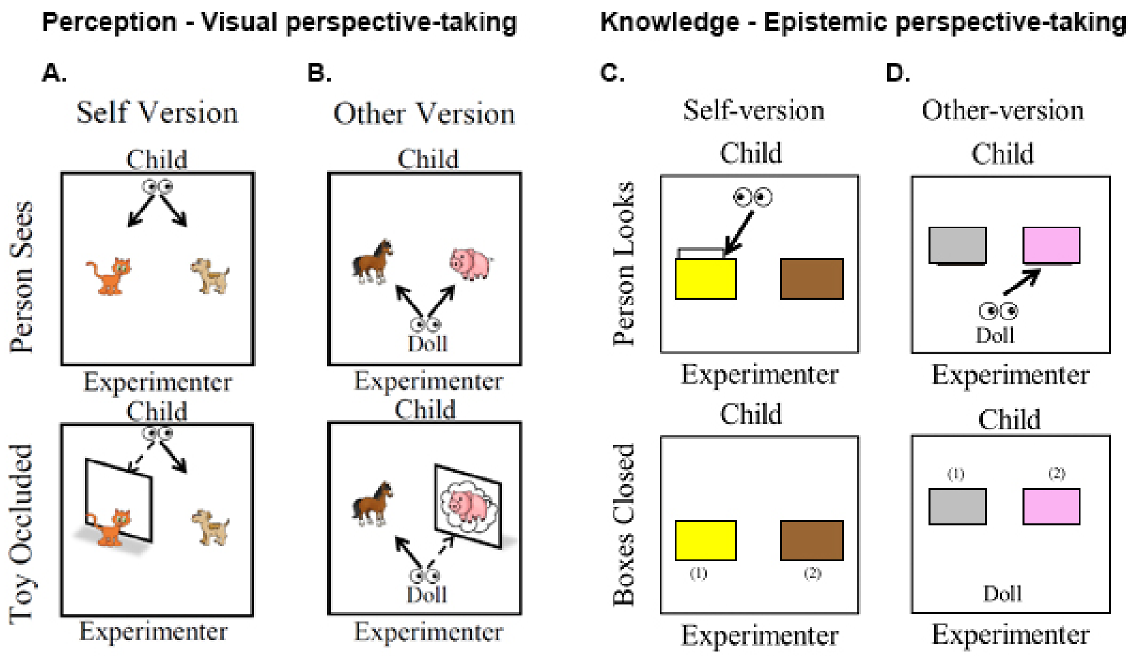

Fig. 3.2: Overview of conditions for the visual and epistemic perspective-taking tasks (adapted from Gonzales et al. (2018)). On the left is the VPT part and on the right the EPT part. A. and C. are the self-versions and B. and D. the other versions with a doll on the lap of the experimenter.

Action-based perspective taking at 27 months

The action-based perspective-taking task was developed by Moll and Tomasello (2006) and assesses children's ability to infer an agent's perceptual information. In the task, the child was asked to help an experimenter find 'the other toy' while one toy was located in plain sight and another was hidden behind a bucket and only visible to the child. The parent was seated on a chair in the middle of the room and two locations were specified at equal distance from the chair, one on the right and one on the left. In on of these locations, a toy was hidden behind a bucket (20x15.5) while the other remained in open sight at the other location. One experimenter (E1) remained in the room, ready with a puppet on the side while another (E2) was outside before the beginning of the trial and walked in through the door which was right across from the location of the chair. There was a brief familiarization phase and a test phase that included 6 trials (see Figure 3.3). For each trial a bucket and two toys were used. The toys matched in all features apart from color.

In the familiarization phase, E2 waited outside while the child was seated on the parent's lap on the chair. E1 took two balls and placed them in the two designated locations, one of them hidden behind the bucket. From the perspective of the parent and the child both balls were visible at all times, whereas the ball behind the bucket was not visible from the door for E2. Once both balls were in place, E2 accompanied the child to the door and pointed at both locations. The parent was then asked to reveal the occluded ball, showing the child that objects behind the bucket were not visible from the door. The revelation of the ball was not accompanied by words to avoid mental state terms such as 'to see'. Afterwards, the child returned to the parent, E2 entered the room and the test phase began.

In the test phase, E2 played with the child to familiarize her with the toys. Each playing session lasted for 50 seconds. In order to avoid a preference due to exposure, E2 made sure to play with both toys equally. After 50 seconds, E2 excused herself with the words "I have to leave shortly, but I'll be back soon." and left the room. During E2's absence, E1 used a puppet to take both toys and switch the location of the bucket, pushing it to the opposite side. Afterwards, the toys were arranged at the two locations, one behind the bucket and the other in open sight from the perspective of the door and both in plain sight with one in front of the bucket from the perspective of the child. The parent was asked to hold onto the child at the location of the chair until the return of E2. E1 told the child "E2 will be back soon. You have to help her!" and then called E2 back into the room by knocking on the door lightly. Upon her return, E2 looked at both locations for 2 seconds each before saying "Where is the other [toy]? I cannot find it." After a short pause, "Can you give it to me?" followed. As soon as the child attempted to give E2 the toy or at least pointed at one of the locations, E2 thanked the child and the trial ended. The next trial began with E2 taking the next set of toys and playing with the child for 50 seconds. The pairs of toys used were cars, bathing frogs, small plastic mice, building blocks, spinning tops and bathing fishes. The starting position of the bucket was varied over children and then switched from trial to trial.

In contrast to the study by Moll and Tomasello (2006), both pointing and giving were coded. For pointing it was coded whether children pointed at the correct toy (behind the bucket), the incorrect toy (in plain sight) or both toys simultaneously (invalid trial). For giving, it was coded whether children handed the experimenter the correct toy (the one located behind the bucket), the incorrect toy (located in plain sight) or both toys (invalid trial). Handing the toy was defined by placing it in the experimenter's hand, rolling it towards her or placing it in front of her feet. Two separate performance scores were calculated for pointing and giving respectively. For pointing it was calculated in how many trials children pointed at the correct location divided by the total times they pointed at one of the two locations. Performance for giving was computed by dividing the trials in which the child handed the experimenter the correct toy by the total number of valid trials in which a toy was given. Invalid trials were defined as trials in which the child did not respond or pointed at or handed both objects simultaneously (27% of the pointing trials, 34% of the giving trials) and also trials in which the child did not remain with the mother and played with one of the toys before the beginning of the test question (2% of the pointing trials, 6% of the giving trials). Overall, 10 children refused to participate in the task and 32 children were excluded because of mistakes in the familiarization. Mistakes ranged from involvement of parents to the usage of incorrect objects or comments about the 'visibility' of the toys (e.g. 'now you can see the ball'). Furthermore, children who did not have at least 3 valid trials after the removal of invalid trials were excluded for the respective scores, resulting in a total of 72 children for the point score and 60 children for the give score. Performance was coded by two independent raters that agreed in 86% of the cases. All discrepancies were discussed and resolved between the two raters. Scores used in subsequent analyses are the action-based perspective-taking pointing score (APT Pointing) and the action-based perspective-taking giving score (APT Giving).

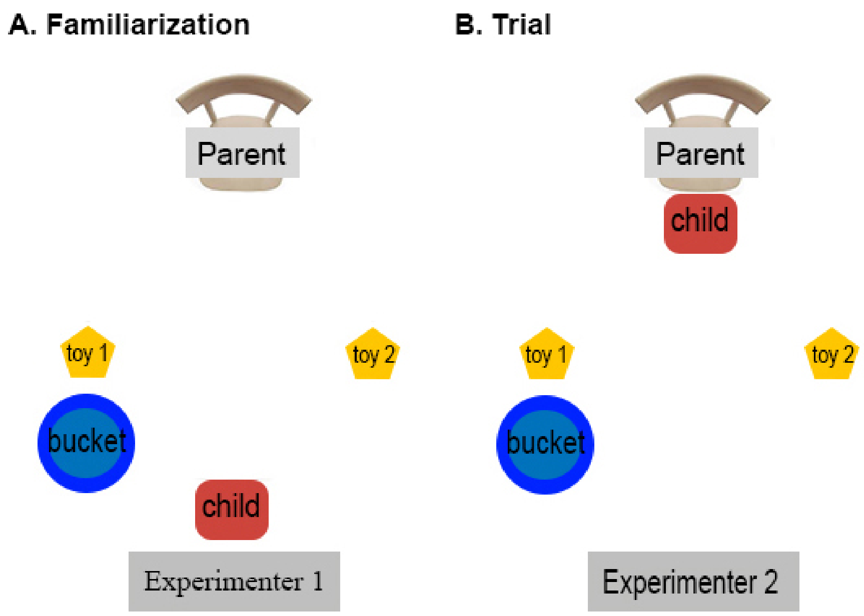

Fig. 3.3: Arrangement during the action-based perspective-taking task. **A.** Familiarization trials in which the child is shown that an object is not visible from the door when occluded by the bucket. **B.** Critical trials during which experimenter 2 returns and requests to be handed the 'other' toy.

Metacognition of ignorance at 27 months

Metacognition of ignorance (MCI) was assessed via the Object Naming task developed by Bartz (2017). In the task children were shown pictures of familiar and unfamiliar objects while signs of uncertainty were coded. Familiar pictures were taken from items listed in the MacArthur Short Form Vocabulary Checklist: Level 1 (Fenson et al., 2000) and depicted a book, a bird, a car, socks, a dog, a spoon, a chair and a shoe. The first two items were familiarization items. The unfamiliar pictures were developed by C. J. Johnson (1992) and were created by combining several features of different objects into one object. Signs of uncertainty were defined as looking at the mother, looking at the experimenter, saying "hmm", saying "no", shaking the head and saying "don't know / I don't know". In the original experiment a few others signs were included, but as they occurred rarely they were not coded in the current implementation of the task. The pictures were shown on an A4 page and drawn in black on a white background. The task began with the experimenter saying "We are going to look at a few pictures together. So, look at this." Upon opening a folder with the various pictures, the experimenter showed the child a picture and asked, "What is that?" The first two pictures served as familiarization and depicted a book and a bird. Children were included as long as they named one of the familiarization objects correctly. In the original study, children received feedback after naming the objects, however, as that appeared to result in more misapplications and word inventions due to children feeling competitive and pressured to give an answer, this was not done in the current study. After naming the book and the bird, the child was shown familiar objects that depicted animals and household items or unfamiliar objects as described above (see Figure 3.4). For each picture the child was asked "What is that?" and signs of uncertainty were coded by a second experimenter that was located sideways behind the child. If the child used an invented word or misapplied an existing word it was noted, but not coded as a sign of uncertainty. In case the child did not respond, the question was repeated once again after three seconds. If the child did not respond a second time, the next picture was shown.

Two types of scores were calculated, a trial-based one that indicated on how many trials a sign of uncertainty was shown for familiar and unfamiliar objects and a qualitative one that created a score based on points for each sign of uncertainty for unfamiliar objects. In the original study by Bartz (2017), each type of sign of uncertainty was compared with sign tests over the two conditions and between two age groups. This was done in order to determine which signs were the most common. In the current study, the focus lay on the amount of trials in which children showed signs of uncertainty, thus a trial-based score was computed. In the trial-based score the number of trials in which a child showed a sign of uncertainty was divided by the total number of trials, establishing a comparable score for both types of trials. Thus, if a child showed any type of sign of uncertainty on a trial she received a point for that trial and a sign on all six trials resulted in a 100% score. The qualitative measure was an additional score that was computed for the current study. This was done for the following reason: A qualitative score enables us to filter out the importance of explicit signs of uncertainty (saying

"I don't know"), semi-explicit signs that could correspond to denial (saying "no" or shaking the head) and implicit signs (saying "hm"). Thus, computing a qualitative score enables us to identify children who are consistently explicit in their expressions of ignorance, versus children who are mostly implicit. For the qualitative score the child received 1 point for saying "hmm", 2 points for shaking the head or saying "no" and 3 points for saying "I don't know". In case a child showed more than one sign of uncertainty on one trial, the highest valued sign of uncertainty was counted. The total score for the qualitative evaluation was therefore 18 points with a child saying "I don't know" for each unfamiliar picture. The reasoning was that replying with 'hmm' is the most subtle sign of uncertainty (apart from looking at the mother or the experimenter) as children do not refuse to reply or admit to being unable to reply. It might only be a sign of hesitation to give an answer. Replying with 'no' and shaking the head on the other hand, are one step further since children refuse to reply, possibly because they are unable to. Lastly, replying with 'I don't know' shows the most metacognitive awareness of uncertainty and therefore received the most points. A total of 12 children refused to participate in task. Further, 12 children who were unable to name both test items in the beginning of the task were excluded from analyses, resulting in a total of 136 children for the task. Responses were coded by two independent coders that agreed on 100% of the cases. Scores used in future computations are the trial-based score in which the percentage of trials is computed in which children showed at least one sign of uncertainty (MCI : Trial-based), and the qualitative score in which children were given more points for more difficult and also more verbal signs of uncertainty as outlined above (MCI: Qualitative). In order to determine the source of a correlation, also the individual impact of signs of uncertainty was investigated.

Mental state language production at 27 months

MSL production was assessed with the help of the MSLQ parental questionnaire based on Bretherton and Beeghly (1982), Olineck and Poulin-Dubois (2005) and Klann-Delius (1998). Parents were asked to note whether their child actively used words from the fol-

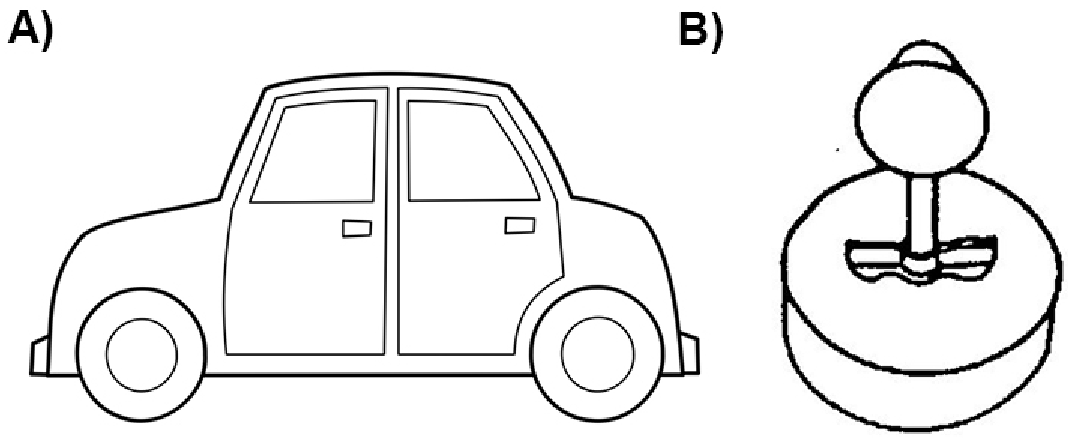

Fig. 3.4: Examples of objects presented during the object naming task. **A.** A real world object, in this case a car **B.** A generated object that has no real-world counterpart and can therefore not be named.

lowing categories: emotion, volition, physiology, ability/obligation, cognition, cognitive particles, modifying particles and knowledge. For each category, with the exception of knowledge, several words were presented and parents were asked to check whether the word had been actively used by their child or not. The number of included words was 51 for emotion, 8 for volition, 26 for physiology, 14 for ability, 16 for cognition, 12 for cognitive particles and 12 for modifying particles. In the knowledge category, parents were asked if their child used the formulation “I know” and “I don’t know” in the following contexts: a) when they know or don’t know the location of a toy, b) when they are shy or do not want to say something, and c) when they were asked about their knowledge literally. See 7.1 in the supplementary material for the original questionnaire used. The questionnaire was scored by calculating a percentage score for each category by dividing the number of actively used words by the number of total words. For the knowledge category the score was divided by the number of scenarios. A total of 13 questionnaires were not filled out in time or were not sent back to the laboratory, resulting in 147 questionnaires that were included for further analyses. Responses were coded by three independent raters who agreed in 100% of the cases. Two types of scores were computed for the MSLQ. The first type of score were percentage scores quantifying the percentage of words used per category and the total percentage for the whole questionnaire. The second type of score assessed variability of vocabulary by giving children a 1 for each category in which they were able to produce at least 1 word and 0 otherwise.

Implicit false-belief eye tracking task at 27 months

The anticipatory looking measure for implicit FBU was adapted from Grosse Wiesmann et al. (2017) and showed computer-animated scenarios. In this task, the child observed a hiding game between two agents (chasee was always a mouse, chaser varied). The hiding game took place in a y-shaped tunnel with two boxes located at the end of each exit. The mouse disappeared in the tunnel and the child's anticipatory looks towards the tunnel exits were used as a measure of expectation.

Task

The task was split into several trials, 10 familiarization (FAM) trials and 12 false-belief (FB) trials. In the original study, 6 true-belief (TB) trials were also presented, however, as there was no difference between FAM trials and TB trials in the original study, TB trials were removed in the current study to reduce the length of the task.

Each trial began with an introduction of the protagonists, first the mouse and then the chaser (either a cat, a dog, a crocodile, an elephant, a fox or a lion). The chaser observed the mouse disappearing within the tunnel and reappearing at one of the two exits, followed by the mouse hiding inside the box at the exit. Depending on the trial, the following events differed.

In the FAM trials, the chaser followed the mouse through the tunnel and reappeared at the other end after 2500 ms, opening the box and discovering the mouse. Before the chaser reappeared at the exit of the tunnel, and 500 milliseconds after entry, a light illuminated both exit areas and the sound of the chaser was played to gain attention (e.g. a bark for the dog).

In the FB trials (see Figure 3.5), two courses of action were possible. In the FB1 trials, the chaser observed the mouse cross over to another box before leaving the scene. During the chaser's absence, the mouse left the box and disappeared. Then the chaser returned, entered the tunnel and as in the FAM trials, the exits were illuminated and a sound followed. However, opposed to the FAM trials, the chaser did not exit the tunnel and the trials ended after 2940 ms. The FB2 trials were similar with the exception that the chaser did not observe the first transfer of the mouse and left beforehand. During the chaser's absence the mouse crossed over to the other box and entered it. Afterwards the mouse left the box and disappeared from the screen. In the FB1 trials, the chaser believed the mouse to be in the last location before the mouse disappeared and in the FB2 trials, the chaser believed the mouse to be in the first box it entered. In both cases the chaser had a FB about the mouse's location. The mouse left the scene in both conditions in order to prevent children from simply looking at the location of the mouse.

The trial order of presentation was randomized; however, before showing a FB trial with a new chaser, a FAM trial was presented first to familiarize the child with the fact that the chaser was following the mouse. Also the correct exit was randomized over trials and balanced for every chaser and every trial type. Children were shown the trials in two blocks with two different randomization orders.

Procedure

As in the SWITCH task, children observed the hiding game on a Tobii T60 Eyetracker while either seated in a car seat or on their parent's lap. The parent was requested to wear opaque sunglasses (to avoid them giving cues about the video on the screen) and to not interact with their child during the viewing, apart from motivating them to look at the screen. The videos were presented on a 1280x1024 pixel screen and the eyes tracked using TobiiStudio. As in the SWITCH task, a five-point calibration was used, however, instead of a cartoon character a simple red dot moved from one calibration point to the next. In case there was no data for one or more calibration points, the procedure was repeated for the missing points. After calibration, the task began. The child was assigned to experimental randomization A or B, which were each separated in two blocks. After the first block, a short break followed in addition to new calibration before the second block of trials was presented.

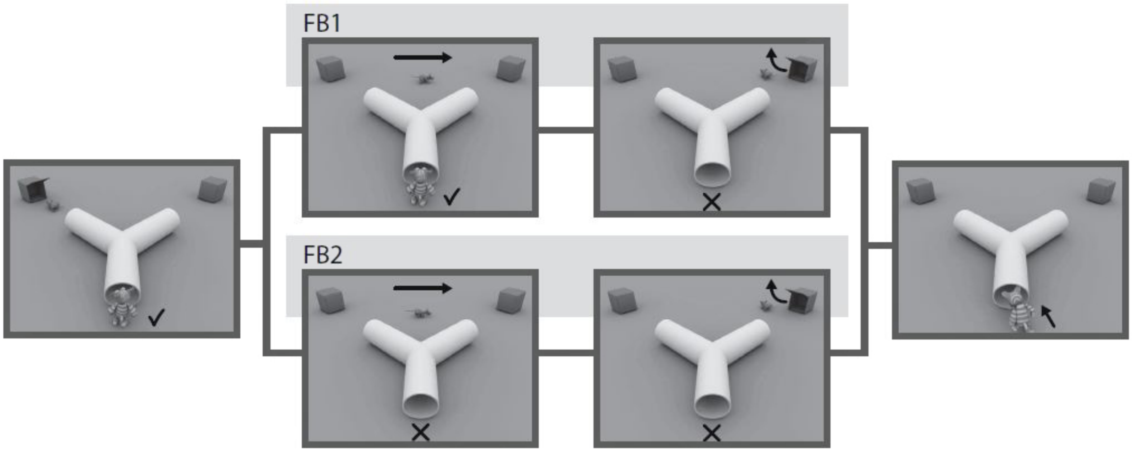

Fig. 3.5: Overview of FB1 and FB2 trials, adapted from Grosse Wiesmann et al. (2017).

Processing and Analysis

Sampling rate was set to 60Hz and the mean of left and right eye-gaze coordinates was analyzed. Preprocessing was conducted with the freeware statistical analysis software R (R Core Team, 2014), followed by statistical analysis with SPSS 25 (IBM Corp. Released 2019. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.) as in all other tasks. For data analysis two Areas of Interest (AOIs) of identical size were created, capturing one tunnel exit and box each. The AOIs were defined by a tangent at the light reflection of the circle and vertically through a perpendicular line which was located in the middle between the two tunnel arms (see Figure B.1 in the supplementary material). Each AOI was 415x374 pixels. All eye tracking videos were visually inspected for a) cases in which important transfer moments of the video were not observed by the child (6.8% of trials), b) the child did not pay attention at the critical moment (2.7% of trials), c) there was no gaze data for the critical interval (1.2% of trials) and lastly, c) trials that were excluded due to bad calibration (1.1% of trials). After these trials were removed, the critical segment of each trial was extracted, which spanned 2500 ms for FAM trials and 2940 ms for FB trials. They began as soon as the animal disappeared in the tunnel and lasted until the end of the trial. After visual inspection and definition of the AOIs, the data was exported and analyzed with the help of an R script. As in the original study, children's first fixations and longest looks were computed with DLS scores as an additional measure. See the results section for the detailed procedure. Overall, 11 children refused to participate in the task. One child did not provide enough trials based on the above mentioned exclusion criteria for all conditions and an additional 3 children did not provide enough trials for a FB2 score. The number of FAM trials ranged from 2 to 12 and the number of FB trials ranged from 3 to 10. The visual assessment of trials was done by two independent raters who agreed on 97% of the cases. Discrepancies were discussed and resolved between the two raters. Scores used in future computations were a mean score between percentage of correct first fixations and percentage of correct longest look as in the original paper with a separate score for the familiarization phase (False Belief: FAM), FB1 trials (False Belief: FB1) and FB2 trials (False Belief: FB2) respectively.

3.3 Results

This section is structured according to the research questions and objectives introduced earlier. First, the descriptive results for the control measures at 24 months and the key measures at 27 months will be presented. Then the question whether MSL production corresponds to MSL comprehension will be addressed. Following, MSL and concurrent relationships to preverbal and verbal tasks will be presented. Finally, a regression model on the predictive relationship between the control measures at 24 months and MSL production will be presented. All additional questions and computations will be addressed towards the end of this section.

3.3.1 Missing values

Not all children were able to participate in all tasks due to fuzziness, refusal or task-specific exclusion criteria. In order to control whether the missing values occurred at random, a missing-completely-at-random (MCAR) test was computed including all relevant test variables for each task and control variables such as sex. In case of non-significance, the test indicates that the missing values are unlikely to depend on other variables and are therefore taken from the same sample distributions. Furthermore, it is unlikely that the missing values depend on unobserved variables in case of a non-significant result. For the MCAR test, all performance variables at both time points were included in addition to control variables like sex, parental education and number of siblings. The test was not significant ($\chi^2(2014) = 2018.24$, $p = .469$), suggesting that values were missing at random and the current subpopulation can be regarded as representative for the larger sample.

3.3.2 Descriptive analyses – 24 months

In the following section the descriptive results of the various tasks will be presented, beginning with the control measures at 24 months and finishing with the key measures at 27 months. See Table 3.3 for an overview of descriptive results for the control measures and Table 3.5 for the key measures.

Tab. 3.3: Descriptive statistics of control measures (number of participants (N), mean (M), standard deviation (SD), range).

Task	N	M	SD	Range
SETK-2: Comprehension	149	9.61	2.99	2-17
SETK-2: Production	146	31.68	15.02	0-84
Bayley Scales: Cognition	157	28.36	2.59	19-33
Bayley Scales: Fine Motor	157	22.51	2.28	16-27
SWITCH task: Learning	118	73.09	34.16	0-100
SWITCH task: Inhibition	100	54.47	34.57	0-100

Tab. 3.4: Descriptive statistics of the employed SET-K scales including the number of participants (N), mean values (M), standard deviation (SD) and average T-values (T) and their range.

Scale	N	M	SD	Range	T	T Range
Word-comprehension	152	6.44	1.71	2-9	52.32	28-69
Sentence-comprehension	146	3.20	1.90	0-8	48.17	26-78
Word-production	146	17.73	6.45	0-31	48.74	26-70
Sentence-production	141	14.54	10.30	0-59	43.08	30-68

SETK-2

The subscales of the SETK-2 were summed up into comprehension and production scores with the comprehension score including word and sentence comprehension and the production score including naming of objects, pictures and description of sentences. The T values and means for the single scales can be found in Table 3.4. There was a significant difference between boys and girls with regard to their language comprehension ($t(147)$ = -2.67, p = .008, Cohen's d = .44, *q = .016). On average, girls responded correctly in 10.22 items out of 17 (SD = 2.89), while boys responded correctly to 8.93 items out of 17 (SD = 2.99). Post-hoc analyses revealed that the difference was mostly driven by sentence comprehension ($t(144)$ = -2.62, p = .011, Cohen's d = .44, *q = .015), but not word comprehension ($t(150)$ = -1.55, p = .124, Cohen's d = .25[3]). Concerning language production, there was no significant difference between boys and girls ($t(144)$ = .30, p = .765, Cohen's d = .05). Children's language production and comprehension were significantly correlated (N = 142, r = .42, $p < .001$, *q = .003) and this correlation remained significant when controlling for sex (N = 142, r = .44, $p < .000$, *q = .004).

[3] The effect size suggests that there may be a difference in word comprehension, but the current study does not have enough data to estimate it with sufficient precision.

Bayley scales

The descriptive results for the cognition and fine motor scale of the Bayley scales can be found in Table 3.3. There were no sex differences for either the cognition scale ($t(155)$ = -.436, p = .664, Cohen's d = .07) or fine motor scale ($t(155)$ = -1.28, p = .201, Cohen's d = .20). In addition, a significant correlation was found between the cognition scale and the fine motor scale (N = 155, r = .46, p < .001, *q = .009).

SWITCH task

In the pre-switch phase, children anticipated the side of the reward following the cue on average in 70% of the trials in the first block (SD = 38.82, range = 0-100) and 77% of the trials in the second block (SD = 38.37, range = 0-100). A Wilcoxon signed-rank test revealed no significant differences in children's learning in the first block of the pre-switch phase compared to the second block, Z = -1.656, p = .102, Cohen's d = .18. See Table 3.3 for the combined performance on both blocks.

In the post-switch phase, children adapted to the change in rule on average in 47% of the trials in block 1 (SD = 38.28, range = 0-100) and 65% of the trials in block 2 (SD = 42.11, range = 0-100). A Wilcoxon signed-rank test revealed a significant difference in children's inhibition of the previously learned association between the first and the second block, Z = -2.881, p = .004, Cohen's d = .45, *q = .02, indicating that children significantly improved from the first four trials in the post-switch phase to the last four trials. See Figure 3.6 for the development over the two phases and chance performance. As prior work has identified sex differences in inhibitory control, this was also tested in the current study (Gagne & Saudino, 2016). A Mann-Whitney U test revealed no significant difference between boys and girls on either block in the pre-switch phase (Block 1: U = 1099, p = .972, Cohen's d = .06; Block 2: U = 1453, p = .311, Cohen's d = .15) and not in the first block of the post-switch phase (U = 840, p = .37, Cohen's d = .18). Considering the adjusted p-value *q, there was also no difference between boys and girls in inhibiting their previously learned rules in the second block of the post-switch phase (Girls: M = 72.55, SD = 40.98, range = 0-100; Boys: M = 55.85, SD = 42.10, range = 0-100; U = 666, p = .038, Cohen's d = .40, $*q$ = .114).

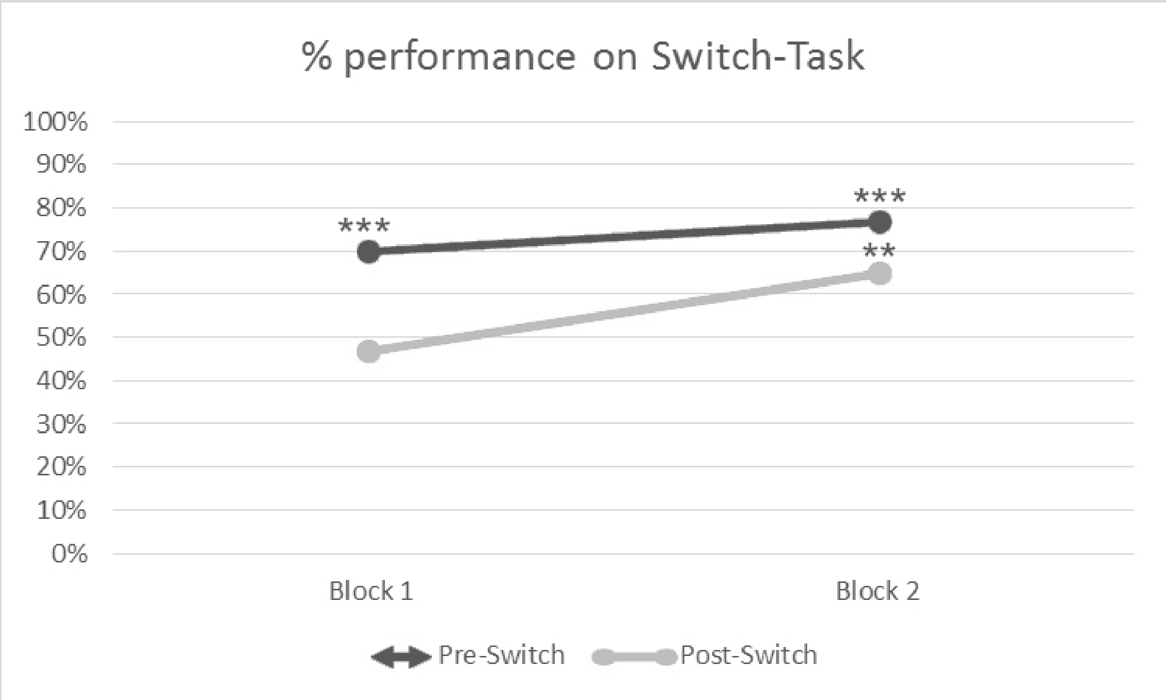

Fig. 3.6: Mean performance on the two blocks of the pre- and post-switch phase. Mean scores that significantly differ from chance performance were identified with a one sample Wilcoxon test. ** $p < .01$ *** $p < .001$

3.3.3 Descriptive analyses – 27 months

Know-Want task

The descriptives for children's performance on the *know* and the *want* condition can be found in Table 3.5. As children were required to pass both trials in each condition in order to be considered competent, chance performance was defined as 0.25. Binominal tests revealed that children performed above chance in the *want* condition ($p < .001$, Cohen's d = .56) but not in the *know* condition (p = .329, Cohen's d = -.06). A Mann-Whitney U test revealed that boys and girls significantly differed in their performance on the *want* condition (Girls: M = 62.71, SD = 48.77, range = 0-100; Boys: M = 40.00, SD = 49.54, range = 0-100; U = 1026, p = .022, Cohen's d = .46, $*q$ = .04), but not on the *know* condition (U = 1194, p = .334, Cohen's d = .19).

Tab. 3.5: Descriptive statistics of key measures (number of participants (N), mean (M), standard deviation (SD), range).

Task	Subtask	N	M	SD	Range
Know-Want task	Want	104	52.88	50.16	0-100
	Know	102	22.55	42.00	0-100
Visual (VPT) & epistemic perspective-taking (EPT)	VPT Self	76	30.26	46.24	0-100
	VPT Other	72	9.72	29.83	0-100
	EPT Self	74	13.51	34.42	0-100
	EPT Other	118	8.47	27.97	0-100
Action-based perspective-taking (APT)	Pointing	72	56.09	25.36	0-100
	Giving	60	57.50	21.85	20-100
Metacognition of ignorance	Trial-based	136	52.06	32.71	0-100
	Qualitative	136	2.93	5.01	0-18
Mental state language production		147	44.05	18.47	3-95
Implicit false-belief understanding	FAM	148	70.39	18.58	25-100
	FB	145	51.85	18.68	0-100

As there was a difference between boys and girls, sex was included as a predictor of performance for the calculation of a generalized estimating equations model (GEE). A GEE with an unstructured working correlation matrix, a binominal distribution and a logit link function was used. Condition (want, know), sex (female, male) and trial order (see methods) were entered as predictors of task performance. See Table 3.6 for detailed results of the model. Only condition and sex were significant predictors of task performance.

Tab. 3.6: Generalized estimating equations results assessing the influence of condition, sex and trial order on performance on the know-want task.

Parameter	B	SE	Wald	df	p value	Exp(B)	95% CI for Exp(B)	
							Lower	Upper
Condition	-1.73	.84	4.25	1	.04	.18	.03	.92
Sex	-.70	.32	4.62	1	.03	.50	.26	.94
Trial Order	-.10	.44	.05	1	.83	.91	.39	2.13
Trial Order by Condition	.16	.30	.27	1	.60	1.17	.65	2.12

To further investigate the developmental pattern for understanding 'know' and 'want', cross tables were generated to examine whether passing the *want* condition can be considered significantly easier than passing the *know* condition. Table 3.7 shows the results of a McNemar test on the difference in passing the *want* condition vs. passing the *know* condition. The test showed that passing the *want* condition can be considered significantly easier compared to the *know* condition.

Visual and epistemic perspective-taking

The proportion of children who passed each separate task is depicted in Table 3.5. Overall, the 27-month-olds in this study did not perform significantly above chance (.25) in any of the four conditions. A generalized estimated equations model (GEE;

Tab. 3.7: McNemar test results on the developmental trajectories of children who passed 'want' before 'know' versus children who passed 'know' before 'want'.

	Know = 0	Know = 1	N	p-value
Want = 0	27	6		
Want = 1	28	9	70	$p < .001$

A McNemar's Chi-squared test with continuity correction was performed.

with an unstructured working correlation matrix, a binominal distribution and a logit link function) was used to assess the effects of task type (visual, epistemic), condition (self, other) and sex (male, female), in addition to the interaction between task type and condition on children's performance. See Table 3.8 for detailed results of the model. Condition and task type significantly predicted children's performance, whereas sex did not. Also the interaction between condition and task type did not contribute significantly to the model. Overall, *self* conditions were easier to solve than *other* conditions and the VPT condition was easier to solve than the EPT condition. Additionally, a McNemar test was computed to determine whether children were more likely to respond with yes for the two trials of the VPT *self* condition. The test was significant, suggesting a bias in responding with yes independent of whether the object was visible.

Tab. 3.8: Generalized estimating equations results assessing the influence of task type (visual, epistemic), condition (self, other), and sex for visual and epistemic perspective-taking.

Parameter	B	SE	Wald	df	p value	Exp(B)	95% CI for Exp(B)	
							Lower	Upper
Sex	.14	.35	.17	1	.68	1.15	.59	2.28
Task	-1.63	.94	3.00	1	.08	.20	.03	2.28
Condition	-1.95	.97	4.03	1	.04	.14	.02	.95
Task by Condition	.62	.65	.90	1	.34	1.85	.52	6.58

Action-based perspective-taking

The average performance on the action-based perspective-taking task is depicted in Table 3.5. As both the pointing score and the giving score were not normally distributed, a one-sample Wilcoxon signed rank test was used to test children's performance against chance-level. Children performed significantly above chance in the giving score (Z = 871.50, p = .022, Cohen's d = .24), but only with a trend towards significance in the pointing score (Z = 1289.00, p = .052, Cohen's d = .24). Furthermore, both the giving and pointing score were highly correlated (r = .755, p < .000, *q = .005) and did not differ significantly from each other (Z = -.587, p = .557, Cohen's d = .06). There were no significant differences between girls and boys on either the pointing score (Z = -.802, p = .422, Cohen's d = .17) or the giving score (Z = -1.209, p = .227, Cohen's d = .30)[4].

Metacognition of ignorance

The average performance on the task can be seen in Table 3.5. In order to determine whether children showed signs of uncertainty significantly more often for the unfamiliar objects compared to the familiar objects, a Wilcoxon signed rank test was computed. Children were significantly more likely to show signs of uncertainty for the unfamiliar objects, 52% of the trials (SD = 32.71%, range = 0-100%), than for the familiar objects, 7% of the trials (SD = 14.87%, range = 0-100%), Z = -9.33, p < .001, Cohen's d = 1.48. The distribution of signs of uncertainty can be seen in Figure 3.7 for both the familiar and the unfamiliar objects. Interestingly, 20.6% of children used the formulation 'I don't know' on at least one unfamiliar object trial, whereas 3.4% of children used the formulation on at least one familiar object trial.

Concerning the qualitative assessment of uncertainty, on average children received 3 points out of 18 (SD = 5.01) with a range from 0 to 18, suggesting that children differed broadly in their expression of their own ignorance. The trial-based uncertainty

[4] The effect size suggests that there might be a significant difference and the current evidence is not sufficient to conclude the absence of an effect.

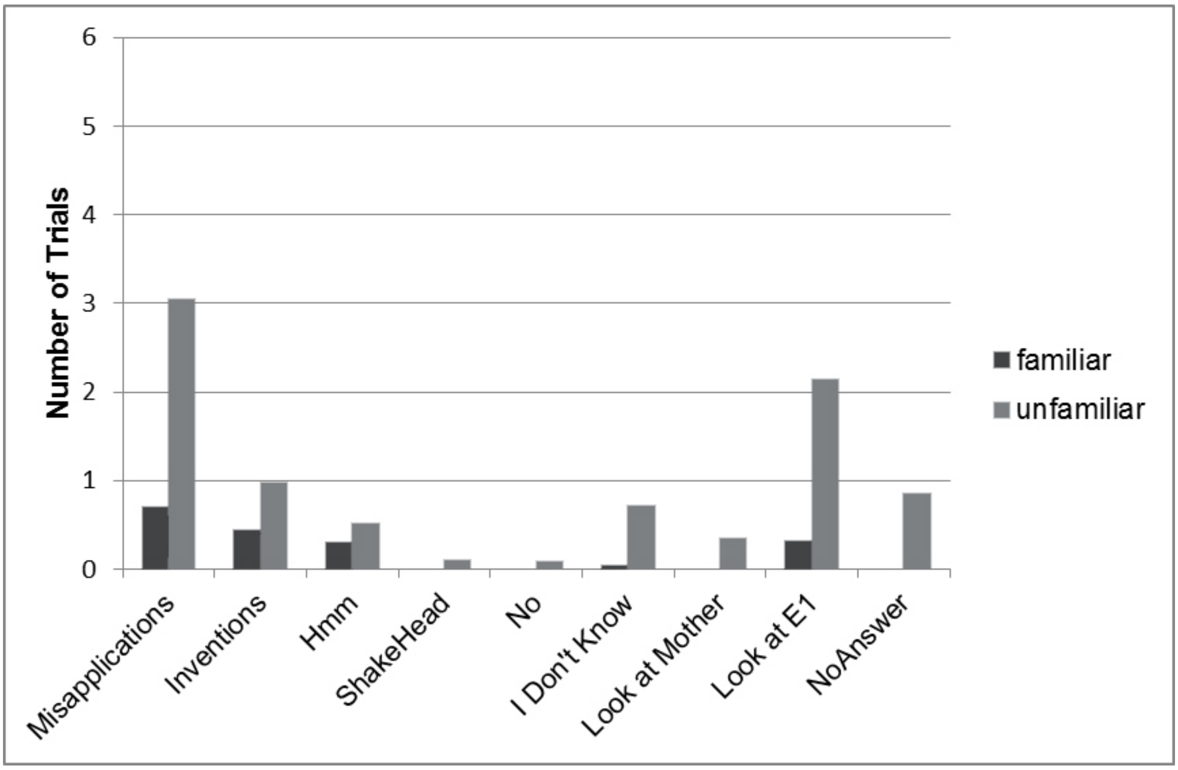

Fig. 3.7: Number of trials in which children showed a sign of uncertainty for familiar objects (blue) compared to unfamiliar objects (red).

measure and the qualitative uncertainty measure were highly correlated with each other ($r = .55$), suggesting that children who showed signs of uncertainty on more trials were also more likely to use explicit signs. Furthermore, there was a correlation between the number of trials in which children show uncertainty for the familiar objects with those for unfamiliar objects. See Table 3.9 for an overview.

Tab. 3.9: Overview of correlations between different scores for the MCI task assessing metacognition of ignorance.

	1.	2.	3.
1. MCI: Familiar Objects	-		
2. MCI: Unfamiliar Objects	.38***	-	
3. MCI: Qualitative	.10	.55***	-

*** $p < .001$, *$q < .004$. $N = 135$.

Mental state language production

Table 3.10 shows children's detailed performance on each of the subscales of the MSLQ. On average, children produced approximately 44% of the vocabulary asked for in the questionnaire. All subscales were strongly intercorrelated (see Table 3.11). As volition terms have been found to be acquired before cognition terms, a McNemar test was computed for children who had only acquired one of the two categories in order to determine whether there was a difference in the distribution of performance for volition and cognition terms. A total of 7 children had at least one word in the volition subscale, but none in the cognition subscale, whereas only one child had at least one word in the cognition subscale, but none in the volition subscale. The McNemar test was not significant with $p = .070$, however, this result has to be considered with caution as only 8 children were eligible for the analysis. An additional way to determine whether children are more versed in the volition subscale than the cognition subscale is to compare percentage of words they have mastered out of each scale. For this analysis 137 children were eligible as they did not perform equally on the subscales. A total of 82.5% of children performed better on the volition subscale than the cognition subscale, whereas only 17.5% of children performed better on the cognition subscale than the volition subscale. A binominal test revealed that the proportion of children performing better on volition than cognition was not due to chance ($p < .000$). Children produced 36% of the volition terms and 24% of the cognition terms.

In order to determine whether performance differences depended on sex, a t-test was calculated for the total score as the data was normally distributed, whereas a Mann-Whitney U test was used for the subscales. There was no significant difference between boys and girls on the total score ($t(145) = 1.526$, $p = .131$, Cohen's $d = .25$). A sex difference on the knowledge subscale was not significant after correcting for multiple comparisons ($U = 1918$, $p = .044$, Cohen's $d = .33$, $^*q = .088$). Overall, girls scored 58% on the knowledge terms ($SD = 33.13\%$, range = 0-100%) whereas boys scored 47% on the knowledge terms ($SD = 32.42\%$, range = 0-100%). Additionally,

Tab. 3.10: Descriptive statistics of the MSLQ scales including children's total performance on the questionnaire.

Scale	*N*	*M* (%)	*SD* (%)	Range (%)
Total	147	44.03	18.47	3-95
Emotions	146	48.54	17.87	4-96
Volition	146	35.96	19.18	0-88
Physiological	145	64.74	16.00	12-96
Obligation	146	45.72	24.96	0-100
Cognition	147	23.76	19.43	0-100
Cognitive particles	144	30.66	21.62	0-100
Modifying particles	144	52.15	27.57	0-100
Knowledge	138	53.14	33.12	0-100

the variety of children's mental state vocabulary was measured by determining the categories that children knew at least one word in. Overall, children knew at least one word in 7.3 categories out of 8 ($SD = 1.3$, range = 2-8), with 69.4% of the sample using at least one word in every category. There was no difference in vocabulary diversity between boys and girls ($U = 2439$, $p = .239$, Cohen's $d = .11$).

Implicit false-belief

As mentioned in the methods section, children's looking time was analyzed from the moment the chasing animal disappeared in the tunnel until it either reappeared (2500 ms for the FAM trials) or until the video ended (2940 ms for the FB trials). The measures computed were children's first fixation (FF; the first AOI they focused on as soon as the animal disappeared inside the tunnel), their longest look (LL; whether they looked longer at the target vs. the distractor) and finally the DLS (differential looking score) which is computed in the following way:

$$\frac{\text{(looking time to correct AOI) - (looking time to incorrect AOI)}}{\text{(looking time to correct AOI) + (looking time to incorrect AOI)}}$$

Tab. 3.11: Overview of correlations between the different subscales of the MSLQ.

	1.	2.	3.	4.	5.	6.	7.	8.
1. Emotions	-							
2. Volition	.64***	-						
3. Physiological	.72***	.57***	-					
4. Obligation	.75***	.60***	.69***	-				
5. Cognition	.71***	.70***	.64***	.75***	-			
6. Cog. particles	.69***	.59***	.66***	.76***	.73***	-		
7. Mod. particles	.74***	.56***	.58***	.74***	.67***	.70***	-	
8. Knowledge	.52***	.46***	.50***	.48***	.41***	.41***	.45***	-

*** $p < .001$. $N = 136\text{-}146$.

The DLS is a measure of preference for looking at the correct AOI in comparison to the incorrect AOI. Its values range from -1 to 1 with -1 indicating a preference for the distractor and 1 indicating a preference for the target. Since none of the measures were normally distributed, non-parametric tests were used in all analyses. In a prior analysis it was checked whether children differed in their performance on the task depending on which trial block they saw (A, B). There was no significant difference between the blocks (FAM: $U = 2724$, $p = .983$, Cohen's $d = .09$; FB1: $U = 2588$, $p = .585$, Cohen's $d = .19$; FB2: $U = 2617$, $p = .997$, Cohen's $d = .06$), therefore, analyses were collapsed. As there was no significant difference between the FF and LL measure for neither the familiarization (FF: $M = 70.39\%$, $SD = 18.58\%$, range = 25-100; LL: $M = 70.78\%$, $SD = 21.39\%$, range = 0-100; Wilcoxon signed rank test: $Z = -.920$, $p = .357$, Cohen's $d = .02$) nor the false belief trials, (FF: $M = 51.85\%$, $SD = 18.68\%$, range = 0-100; LL: $M = 52.74\%$, $SD = 20.38\%$, range = 0-100, Wilcoxon signed rank test: $Z = -.653$, $p = .514$, Cohen's $d = .05$), the results for both measures were collapsed as in the original study. Since DLS scores yielded similar results they will not be reported here in order to keep the current analysis as similar to the original as possible. See Figure 3.8 for a depiction of children's mean performance on the various trials types.

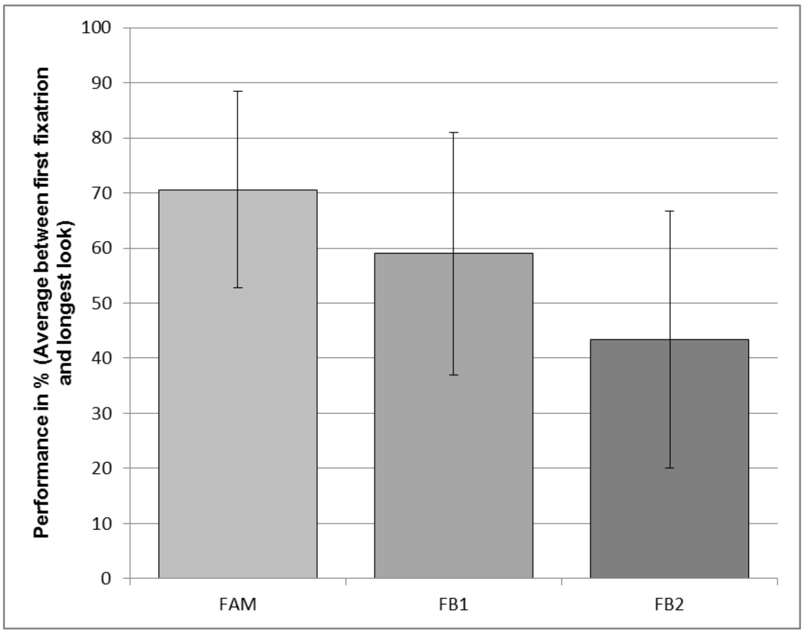

Fig. 3.8: Mean rate of correct anticipatory looks for all three conditions (FAM, FB1, FB2). The mean score is computed by taking the average of the first fixation and the longest look measure.

In the FAM condition, children looked at the correct AOI significantly more often than expected by chance (M = 70.58 %, *SD* = 17.85%, range = 13-100; One sample Wilcoxon signed rank test: $Z = 8.705$, $p < .001$, Cohen's $d = 1.15$), indicating that they correctly learned to anticipate the chasing animal's location. Children performed significantly above chance in the FB1 condition ($M = 59.00\%$, $SD = 21.98\%$, range = 0-100; $Z = 4.984$, $p < 0.001$, Cohen's $d = .41$) and significantly below chance in the FB2 condition ($M = 43.39\%$, $SD = 23.30\%$, range = 0-100; Z = -3,578, $p < .001$, Cohen's $d = -.28$). In contrast to the original study, there was a significant difference in children's performance on the FB1 and FB2 trials; therefore, they were not collapsed into a general false belief category ($Z = -5.925$, $p < 0.001$, Cohen's $d = .53$, $*q = .004$). Furthermore, there was a significant difference in performance depending on whether children saw the FAM, FB1 or the FB2 trials, $\chi^2(2)$ = 101,060, $p < .001$. Post hoc analyses with Wilcoxon signed rank tests were conducted with FDR corrected p-values. Overall, all three comparisons were significant. Children performed significantly better on the FAM condition than on both the FB1 (Z = -4.897, $p < 0.001$, Cohen's $d = .16$, $*q = .011$) and FB2 (Z = -7.989, p < .001, Cohen's $d = .35$, $*q = .005$) conditions.

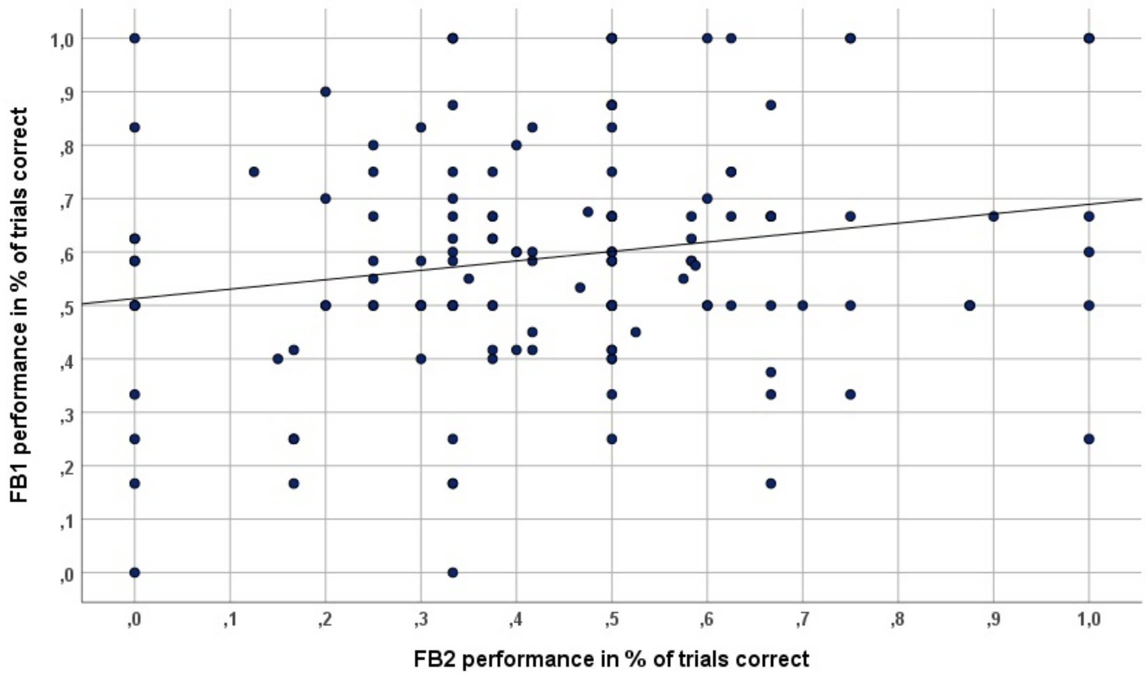

Fig. 3.9: Scatterplot of FB1 and FB2 performance measured in % of correct trials.

Lastly, there were no correlations between the FAM trials and FB1 (N = 147, r = .06, p = .454) or FB2 (N = 145, r = -.11, p = .203). Also the two types of false-belief trials did not correlate when controlling for multiple comparisons (N = 145, r = .164, p = .048, *q = .132), indicating that children who performed better on the FB1 trials did not necessarily perform better on the FB2 trials. See Figure 3.9 for a scatterplot of the FB1 and FB2 scores.

In the following section correlational results will be presented according to the research questions introduced in the literature review.

3.3.4 Mental state language production and mental state language comprehension

The first research question was concerned with the relationship between children's MSL production and measures of MSL comprehension. For this purpose, the relationship between MSL production and three tasks was analyzed. First of all, the relationship between MSL production (via MSLQ) and children's metacognition of ignorance (MCI), which would support children using mental state terms in mental state appropriate contexts. Second, the relationship between children's MSL production (via MSLQ) and their ability to draw the right pragmatic inferences from the verbs 'want' and 'know' (Know-Want task), which would argue for comprehension of what 'knowing' and 'wanting' means in context. And third, the relationship between children's MSL production (via MSLQ) and their ability to verbally report their own knowledge and the knowledge of others (epistemic perspective-taking), which might suggest that children who produce mental state terms are able to use them to denote knowledge states faster than children who do not produce them yet.

Mental state language production and metacognition of ignorance

The relation between MSL production and children's metacognitive awareness of their own ignorance (MCI) was tested through Spearman's rank correlations. There were no significant correlations between the trial-based measure of the MCI task and the subscales of the MSLQ (N = 118-124, r = .01 - .13, p = .194 - .686), neither for familiar nor for unfamiliar objects. However, there was a significant relation between the qualitative measure of the MCI task and several subscales of the MSLQ questionnaire. See Table 3.12 for a detailed listing of the correlational results.

In order to determine which signs of uncertainty drive the correlation the strongest, correlations between the MSLQ total score and the different signs of uncertainty were computed. Only children's usage of the formulation 'I don't know' correlated significantly with the MSLQ total score, indicating that higher scores on the MSLQ were associated with more frequent usage of 'I don't know' to express uncertainty (see Table 3.13 for the detailed results). A partial correlation was computed between the qualitative measure of metacognitive awareness and the total score of the MSLQ while controlling for general language and cognition. The partial correlation remained significant when controlling for general cognitive skills (N = 119, r = .20, p = .022), but not when controlling for both general language and general cognition (N = 113, r = .13, p = .16).

When examining the relationship between all MSLQ subscales and saying 'I don't know', several scales appear to play a role (see Table 3.14). Especially the knowledge scale is of particular interest as it assesses the scenarios in which children use the formulation 'I don't know' according to their parents. When only looking at the scenario in which children use the formulation to talk about knowing or not knowing the location of a toy, the correlation remained significant (N = 128, r = .25, p = .004, *q = .024). Again partial correlations were computed to control for general factors. Since replying with 'I don't know' was not correlated with general cognition (N = 133, r = .10, p = .253), only general language was controlled for in a partial correlation. The relation between saying 'I don't know' and MSL production remained as a trend when controlling for general language (N = 118, r = .17, p = .070). For the detailed correlations between saying 'I don't know' and the subscales of the MSLQ, only the knowledge scale significantly correlated after controlling for general language (N = 111, r = .19, p = .043). The partial correlation remained significant for the specific scenario of talking about the location of a toy, N = 119, r = .20, p = .029.

In order to assess whether it was easier for children to produce the word 'know' without using it in the context of their own knowledge, a McNemar test was calculated. For this purpose the pattern of children producing the word 'know' according to the MSLQ (72.6%) and the number of children using it in the MCI task to denote their own ignorance (20.6%) was arranged in a crosstable. The McNemar test revealed a significant effect, suggesting it was much more likely for children to be able to produce the verb while not using it in the task, compared to using it in the task and not being able to produce it (see Table 3.15). Furthermore, a chi-square test of independence was performed to examine the relation between denoting their own ignorance by using the mental state verb 'know' and actually producing it according to the MSLQ. The relation between these variables was significant, χ^2 (1, N = 124) = 8.29, p = .004, *q = .008. Children who did not produce the verb 'know' were less likely to use it to denote the lack of their own knowledge.

As many more children who were able to produce the word 'know' used it in the MCI task, than children who were not able to produce it, an additional McNemar test

Tab. 3.12: Overview of correlations for the qualitative measure of the MCI task and the subscales of the MSLQ. All p-values were corrected for multiple comparisons according to the FDR.

	MCI: Qualitative	N
1. MSLQ Total	.27**	126
2. Emotions	.23*	125
3. Volition	.26**	125
4. Physiological	.16	125
5. Obligation	.22*	125
6. Cognition	.22*	126
7. Cognitive particles	.21*	123
8. Modifying particles	.21*	123
9. Knowledge	.22*	119

* $p < .02$, $^{*}q < .03$ ** $p < .01$, $^{*}q < .02$

was calculated for the remaining children who did not use ‘I don’t know’ in the MCI task. For these children a crosstable was created to compare their production of the verb ‘know’ on the MSLQ and whether they showed any type of sign of uncertainty on the MCI task. The McNemar test was not significant, indicating that there was no difference between children who showed a sign without being able to produce the verb ‘know’ and the other way around (see Table 3.16). There was also no significant relationship between producing the verb ‘know’ and being respectively better at showing their own ignorance (chi-square test of independence, χ^2 (1, $N = 98$) = .90, $p = .342$).

Mental state language production and pragmatic inferences for ‘know’ and ‘want’

To determine whether there was a relation between children’s productive mental state vocabulary and their performance on the know-want task, Spearman’s rank correlations were computed for all the subscales of the MSLQ and the *want* condition. As performance for the *know* condition was at floor level, it was not included in the cal-

Tab. 3.13: Overview of correlations for total score of the MSLQ questionnaire and the different signs of uncertainty children showed in the MCI task. All p-values were corrected for multiple comparisons according to the FDR.

	MSLQ Total	*N*
1. 'Hmm'	.17	126
2. Shaking Head	-.17	126
3. 'No'	-.04	126
4. 'I Don't Know'	.30**	126

** $p < .01$, *q < .02

Tab. 3.14: Overview of correlations for children who used 'I don't know' during the MCI task and all the subscales of the MSLQ. All *p*-values were corrected for multiple comparisons according to the FDR.

	'I don't know'	*N*
1. MSLQ Total	.39**	126
2. Emotions	.24**	125
3. Volition	.24**	125
4. Physiological	.27**	125
5. Obligation	.20*	125
6. Cognition	.18*	126
7. Cognitive particles	.25**	123
8. Modifying particles	.20*	123
9. Knowledge	.30**	119

* $p < .04$, *$q < .04$ ** $p < .01.$, *$q < .01$

Tab. 3.15: McNemar test results on children who use 'don't know' according to the MSLQ and children who use it to denote their own ignorance in the MCI task.

	MCI: 'don't know' = 0	MCI: 'don't know' = 1	N	p-value
MSLQ: 'know' = 0	31	1		
MSLQ: 'know' = 1	67	25	124	$p < .000$

A McNemar's Chi-squared test with continuity correction was performed. The corrected p-value *q was .004.

Tab. 3.16: McNemar test results on children who use 'don't know' according to the MSLQ and children who show at least one sign of uncertainty in the MCI task. Only children who did not use 'I don't know' were included.

	MCI: Sig$N = 0$	MCI: Sig$N = 1$	N	p-value
MSLQ: 'know' = 0	4	27		
MSLQ: 'know' = 1	14	52	97	$p = .060$

A McNemar's Chi-squared test with continuity correction was performed.

culations. See Table 3.17 for the detailed results. Overall, several subscales of the MSLQ were significantly related to the *want* condition, among them emotional and physiological terms and the knowledge scale. Cognition terms were only at a trend with $p = .051$ and $^*q = .076$. Despite 86% of children already producing the verb 'want' according to the MSLQ and 73% producing the verb 'know/don't know', the respective scales did not correlate with performance.

In order to control for general cognitive skills, partial correlations were computed. The relation between MSL production and the pragmatic inferences for 'want' did not remain significant ($N = 95$, $r = .02$, $p = .823$). Out of the specific relations, only the correlation between the want score and the knowledge scale of the MSLQ remained significant when controlling for general cognition ($N = 86$, $r = .29$, $p = .006$).

To look at this relationship in more detail, a crosstable with children who used the formulation 'I don't know' in the context of not knowing the location of a toy (MSLQ: knowledge scale), and of children who passed or did not pass the *want* condition was created. A McNemar test was significant and indicated that it was more likely for children to use 'I don't know' in the appropriate situation and not pass the *want* condition, than the other way around (see Table 3.18).

Mental state language production and epistemic perspective-taking

As the fact that children are able to answer questions about their own and other's knowledge suggests comprehension of the mental state verb 'know', the relationship between the MSLQ questionnaire and the EPT task was investigated. However, as the seeing-knowing relation is known to only develop at the age of 3-4 (Pratt & Bryant, 1990) and children performed at floor for both conditions of the EPT task, it was not considered for correlations. Only 13% of children showed competence in the *self* condition and only 8% in the *other* condition.

However, a closer look at the pattern of children who showed competence was possible. In order to look at the usage of the verb 'know' in particular in relation to the EPT task, a crosstable was created with children who produced or did not produce

Tab. 3.17: Overview of correlations between the *want* condition of the know-want task and the subscales of the MSLQ. All *p*-values were corrected for multiple comparisons according to the FDR.

	Want	*N*
1. MSLQ Total	.24*	98
2. Emotions	.21*	97
3. Volition	.16	97
4. Physiological	.26*	96
5. Obligation	.11	98
6. Cognition	.20	98
7. Cognitive particles	.05	95
8. Modifying particles	.24#	97
9. Knowledge	.31**	92

* $p < .03$, $*q < .05$; ** $p < .01$, $*q < .02$; # $p = .015$, $*q = .067$.

Tab. 3.18: McNemar test results on children who use 'I don't know' to denote not knowing the location of a toy according to the knowledge scale of the MSLQ and children's performance on the *want* condition.

	Want = 0	Want = 1	*N*	p-value
MSLQ: LocationToy = 0	18	10		
MSLQ: LocationToy = 1	25	43	96	$p = .017$

A McNemar's Chi-squared test with continuity correction was performed.

Tab. 3.19: McNemar test results on children who produce the verb ‘know’ according to the MSLQ and children who pass or fail the *self* condition of the epistemic perspective-taking task.

	EPT Self = 0	EPT Self = 1	*N*	p-value
MSLQ: ‘know’ = 0	13	1		
MSLQ: ‘know’ = 1	46	8	68	$p < .000$

A McNemar’s Chi-squared test with continuity correction was performed.

Tab. 3.20: McNemar test results on children who use ‘don’t know’ when talking about the location of a toy according to the MSLQ and children who pass or fail the *self* condition of the epistemic perspective-taking task.

	EPT Self = 0	EPT Self = 1	*N*	p-value
MSLQ: LocationToy = 0	13	0		
MSLQ: LocationToy = 1	48	8	69	$p < .000$

A McNemar’s Chi-squared test with continuity correction was performed.

the verb ‘know’ according to the MSLQ and children who passed or failed the *self* condition of the EPT task. A McNemar test revealed that it was much more likely for children to be able to produce the verb, but not pass the EPT task than the other way around. Additionally, the descriptive numbers show that out of 9 children who passed the task, 8 were already able to produce the verb ‘know’ according to the MSLQ (see Table 3.19). Results were similar for a crosstable that compared children using the formulation ‘don’t know’ to talk about the location of a toy (MSLQ: knowledge scale) and children passing or failing the *self* condition of the EPT task (see Table 3.20).

3.3.5 Mental state language production and concurrent relationships to preverbal and verbal abilities

Mental state language production and perspective-taking

As relations between MSL production and level-1 perspective-taking were expected, correlations between the verbal visual (VPT) perspective-taking task and the subscales of the MSLQ were computed. Only the VPT *self* condition was significantly correlated with the MSLQ and all the subscales. No other condition of the task was related to the MSLQ. See Table 3.21 for the detailed results of the *self* condition of the level-1 perspective-taking task and all MSLQ subscales.

As both the MSLQ total score and children's score on the VPT *self* condition were highly correlated with general language skills, a partial correlation was computed to control for that (see also Table 3.21). The correlation did not remain significant when partialing out children's general language skills at the age of 24 months (N = 65, r = .19, p = .132). However, when examining the specific correlations between the VPT *self* condition and the subscales of the MSLQ, the relationship between VPT self and the knowledge scale remained significant when controlling for general language (N = 61, r = .25, p = .049).

Additionally, correlations between the MSLQ questionnaire and the action-based perspective-taking task were computed. However, neither the pointing score (N = 69, r = .01, p = .434), nor the giving score (N = 57, r = .20, p = .140) significantly correlated with the total score of the MSLQ questionnaire. The same was true for the subscales.

Mental state language production and implicit false-belief understanding

As children showed different performance patterns for the two FB trials in the anticipatory FB task, correlations were computed for FB1 and FB2 separately. There were no significant correlations between the total score of the MSLQ for neither the familiarization trials (N = 138, r = -.04, p = .673), nor for the FB1 trials (N = 137, r = -.024, p = .782) or FB2 trials (N = 135, r = -.01, p = .932). The same applied for the subscales of the MSLQ. In order to determine whether it was significantly more likely to pass the FB1 trial than passing the FB2 trial, a McNemar test was calculated. The test was significant and indicated that significantly more children passed the FB1 trials and not the FB2 trials than the other way around (see Table 3.22).

Tab. 3.21: Overview of correlations for the *self* condition of the visual and epistemic perspective-taking task and the subscales of the MSLQ, including partial correlations controlling for general language for VPT Self. All p-values were corrected for multiple comparisons according to the FDR.

	VPT Self (N)	VPT Self Partial (N)
1. MSLQ Total	.38** (70)	.19 (65)
2. Emotions	.32** (69)	.09 (64)
3. Volition	.35** (69)	.21 (64)
4. Physiological	.29* (69)	.10 (64)
5. Obligation	.33** (70)	.16 (65)
6. Cognition	.24* (70)	.10 (65)
7. Cognitive particles	.33** (68)	.15 (63)
8. Modifying particles	.34** (68)	.16 (63)
9. Knowledge	.37** (65)	.25* (61)

** p < .01, *q < .02; * p < .05, *q < .05

Tab. 3.22: McNemar test results on the developmental trajectories of children who passed FB1 before FB2 versus children who passed FB2 before FB1.

	FB2 = 0	FB2 = 1	N	p-value
FB1 = 0	48	12		
FB1 = 1	58	24	145	p < .001

A McNemar's Chi-squared test with continuity correction was performed.

3.3.6 Basic cognitive and linguistic predictors of mental state language

In order to predict children's productive mental state vocabulary at the age of 27 months, a linear regression was calculated which included the predictors general language comprehension and production (word production and sentence production), general cognition, general fine motor skills and inhibitory skills at the age of 24 months. Given that there were sex differences in children's inhibitory skills and language comprehension abilities, sex was also included as a predictor. Inhibitory skills were defined as performance in the second block of the post-switch phase, because only the second block was shown to be correlated with mental state language. All predictors were entered simultaneously. The model was significant ($p < .000$) and explained 45% of the variance. Children's sentence production, inhibitory skills and motor skills at the age of 24 months were identified as the only significant predictors. See Table 3.23 for a detailed overview of the predictors and regression statistics. In order to determine the amount of variance that only significant predictors contribute, a reduced model with sentence production, inhibitory skills and fine motor skills was computed. The model was significant ($p < .000$) and explained 42% of the variance with the adjusted R^2 remaining at .39.

Finally, in order to make sure that the relationships between MSL and sentence production and MSL and inhibition were not due to children differing in their fine motor skills, partial correlations were computed. Both the relationship to sentence production ($N = 130$, $r = .38$, $p < .000$) and the relationship to inhibition ($N = 89$, $r = .21$, $p = .044$) remained significant when controlling for children's fine motor skills.

3.3.7 Additional Questions

Relations between the preverbal and verbal level-1 perspective taking

In order to identify whether the two perspective taking tasks were measuring a similar underlying construct, correlations were computed between the pointing and giving scores and the VPT *self* score as the *other* score had floor performance. Only the

Tab. 3.23: Linear model of predictors of mental state language.

Parameter	*B*	*SE B*	β	*p* value	95% CI for B	
					Lower	Upper
Sex	.013	.040	.034	.749	-.068	.094
General Cognition	.005	.010	.064	.596	-.015	.026
General Fine Motor	.023	.010	.247	.032	.002	.043
Inhibition	.093	.046	.204	.049	.001	.184
Language Comprehension	.007	.007	.106	.373	-.008	.022
Word Production	.003	.004	.117	.392	-.004	.011
Sentence Production	.005	.002	.295	.022	.001	.009

R^2 = .45, adjusted R^2 = .39.

pointing score was significantly correlated with the *self* condition of the VPT task (N = 37, r = .35, p = .041), whereas the giving score was not (N = 31, r = -.05, p = .813). There was no control variable that both measures were highly correlated with, however, in order to control for the relationship between the two variables being based on children's verbal ability, general language was partialed out. The relationship between the two perspective-taking measures remained significant when controlling for general language (N = 32, r = .34, p = .046).

Relations between general language, cognition, fine motor skills and inhibition skills

In order to investigate whether the control measures in this study are interrelated, several Spearman rank correlations were computed among the different control tasks. There was a significant correlation between language comprehension and both, cognitive skills (N = 146, r = .30, p < .001, *q = .002) and motor skills (N = 147, r = .27, p = .001, *q = .006). No significant correlation was found between language comprehension and learning in the pre-switch blocks (Block 1:N= 87, r = -.04, p = .743; Block 2:N= 105, r = .04, p = .652), however, there was a significant correlation with inhibition skills in the second half of the post-switch phase (N = 76, r = .252, p = .028, *q = .045), but not in the first (N = 80, r = .09, p = .453). Similarly to language comprehension, language production was also significantly correlated with both cognitive (N = 143, r = .392, p < .001, *q = .013) and motor skills (N = 144, r = .314, p < .001, *q = .002), but not with the switch measures (N = 74-84, r = .03 - .13, p > .382). Both languages measures (N = 145, r = .441, p < .001, *q = .003) and cognitive and fine motor measures (N = 155, r = .46, p < .001, *q = .004) were highly intercorrelated among each other. For the fine motor scale there was a significant negative correlation between the first half of the post-switch trials, indicating that children who had better fine motor skills were the ones with worse inhibition skills (N = 86, r = -.279, p = .009, *q = .017).

Relations between control and key measures

See Table 3.24 for an overview of correlations between general language, general cognition, general fine motor skills, learning / inhibition skills and the key measures of this study. It is important to note that in spite of the lack of a significant correlation between giving in the action-based perspective-taking measure and total general language abilities, there was a significant correlation between giving and general language comprehension (N = 58, r = .28, p = .030). Correlations with children's performance on the *know* condition of the know-want task and the EPT task are not reported due to low variance.

General task relations

In order to investigate whether the key measures were interrelated, further correlational analyses were computed with the exception of tasks that had too little variance to be interpretable. See Table 3.25 for the overview of correlations. Most notably there was a significant correlation between the trial-based measure of the MCI task and the *self* condition of the VPT task. When controlling for general language, the relation remained (N = 62, r = .31, p = .012). This suggests that children who were more

Tab. 3.24: Overview of correlations for key measures and control measures. All p-values were corrected for multiple comparisons according to the FDR.

	Language (N)	Cognition (N)	Fine Motor (N)	Learning (N)	Inhibition (N)
1. Want	.05	.23	.11	-.09[1]	.03[2]
	(97)	(104)	(103)	(66)	(62)
2. VPT Self	.21	.13	.04	.04[1]	-.03[2]
	(74)	(73)	(74)	(41)	(38)
3. APT Pointing	.12	.07	-.01	-.35[1]	.12[2]
	(68)	(71)	(71)	(44)	(36)
4. APT Giving	.12	.19	.03	-.48**[1]	.05[2]
	(57)	(58)	(59)	(37)	{30}
5. MCI:	.14	.13	.13	.14[1]	.01[2]
Trial-based	(128)	(132)	(132)	(81)	(72)
6. MCI:	.23	.22	.14	-.11[2]	-.13[2]
Qualitative	(128)	(132)	(132)	(96)	(72)
7. MSLQ Total	.54***	.27**	.29**	-.05[2]	.25*[2]
	(136)	(144)	(145)	(103)	(78)
8. False Belief: FB1	.01	-.16	-.14	-20[1]	-.16[1]
	(134)	(144)	(144)	(91)	(84)
9. False Belief: FB2	.02	.08	-.03	.08[2]	-.16[2]
	(132)	(142)	(142)	(107)	(80)

For learning and inhibition scores, only correlations with the highest Spearman's r are reported. Superscripted numbers next to the learning and inhibition columns indicate which block the value was taken from. ** $p < .01$, *$q < .03$ *** $p < .001$, *$q < .04$.

likely to show any type of sign of uncertainty on MCI trials, were also more likely to be able to answer questions about their own perspective, independent of language skills. Furthermore, when looking at the relationship between the *want* condition and children's performance on the action-based level-1 perspective-taking task, only the giving score correlated significantly, but not the pointing score. This relation remained significant when controlling for general language skills ($N = 37$, $r = .35$, $p = .029$). This pattern is the opposite to the relationship between the two perspective-taking tasks as there it was the pointing score that correlated significantly.

Tab. 3.25: Overview of correlations for all key measures. *N* are noted in brackets. All p-values were corrected for multiple comparisons according to the FDR.

	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. Want	-	-	-	-	-	-	-	-	-
2. VPT Self	-.01 (46)	-	-	-	-	-	-	-	-
3. APT Pointing	.065 (49)	.35 (37)	-	-	-	-	-	-	-
4. APT Giving	.40* (41)	-.05 (31)	.76** (54)	-	-	-	-	-	-
5. MCI: Trial-based	.01 (89)	.41** (66)	-.08 (66)	.07 (53)	-	-	-	-	-
6. MCI: Qualitative	.09 (89)	.27 (66)	-.15 (66)	.01 (53)	.55*** (135)	-	-	-	-
7. MSLQ Total	.24* (98)	.37** (70)	.10 (69)	.20 (57)	.10 (125)	.29** (125)	-	-	-
8. False Belief: FB1	-.03 (98)	-.05 (68)	-.06 (69)	.15 (56)	.09 (126)	.10 (126)	-.02 (137)	-	-
9. False Belief: FB2	.04 (97)	.06 (68)	-.06 (68)	.06 (55)	.13 (124)	.03 (124)	-.01 (135)	.16 (145)	-

* $p < .05$, $*q < .05$ ** $p < .01$, $*q < .04$

3.4 Discussion

Study 1 of this thesis had the aim to investigate the role that MSL production may play in the transition from children's early mental state understanding in preverbal communication to their later mastery of explicit ToM tasks. In order to do that, a total of three aims were pursued. Since the early usage of mental state terms may not indicate a reference to mental states, 1) the relationship between children's production of mental state terms and their comprehension of said terms was investigated. In a second step, 2) the concurrent relations of MSL production and preverbal and verbal abilities like perspective-taking and implicit FBU were investigated, and 3) the basic cognitive and linguistic predictors of MSL production were identified. In discussing the results of this study, first the findings related to the study hypotheses are presented, followed by a general conclusion[5]. Descriptive findings and task-specific limitations will be discussed in the respective hypothesis sections and more generic limitations will be addressed in the general discussion of this thesis.

3.4.1 Mental state language production and mental state language comprehension

Many studies have shown the importance of MSL for later ToM performance (Brooks & Meltzoff, 2015; Olineck & Poulin-Dubois, 2007) and that children begin to differentiate between 'know' and 'think' in experimental settings around the age of 5 (Kristen-Antonow et al., 2019; Moore et al., 1989), but the degree to which children's early MSL production is used to denote early mental states remains under-studied. Harris, Yang, and Cui (2017) used spontaneous speech data in order to determine whether there are indicators for children using MSL specifically in order to denote mental states and have argued that 2-year-old children use the verb 'know' spontaneously and in order to talk about their own and somebody else's knowledge. Furthermore, Bartz (2017) claimed that children show signs of uncertainty when confronted with unfamiliar objects, which

[5] Results that were presented for the purpose of completeness will not be discussed in detail i.e. intertask correlations that do not relate to the main questions of this thesis and were only relevant for particular analysis steps.

suggests that they have some internalized concept of knowledge and ignorance. This early awareness of their own knowledge states is also supported by 2- to 3-year-old children asking information-seeking questions to their caretakers (Harris, Yang, & Cui, 2017). Assuming a simulation theorist position, an early representation of knowledge states and early production of the verb ‘know’, may be an indicator that children already associate mental state verbs with their early and nonverbal representations of mental states. However, until now there have not been any studies that systematically investigated the relationship between early indicators for a representation of mental states (preverbal competence) and children’s production of mental state terms correlationally. Therefore, even if one assumes children to possess early preverbal representations of mental states, it is difficult to argue that children’s usage of mental state terms is already based on those preverbal representations.

The first aim of this study was thus, to examine whether children who are already capable of producing mental state terms like ‘know’ and ‘want’, show better comprehension of the semantics and pragmatics of these verbs in experimental contexts. For this purpose, MSL production was correlated with performance on three different tasks. First of all, a task that assessed whether children showed signs of uncertainty and used explicit verbal labels for their own ignorance when confronted with pictures of unfamiliar objects. Second, a task that measured children’s ability to draw correct pragmatic inferences from utterances involving ‘know’ and ‘want’, and third, an epistemic perspective-taking task in which children were asked to report their own and someone else’s knowledge about a box whose content was known and another box whose content was unknown.

Before discussing the relationship between MSL production and each of these tasks, it is important to compare children’s performance on the mental state language questionnaire (MSLQ) with that of previous studies. In the current study, 27-month-olds knew approximately 44% of the terms presented in the questionnaire. This is in line with a similar study on German children conducted by Kristen et al. (2012) that found a total score of 37% for 24-month-olds and 61% for 30-month-olds. A study on Englishspeaking 30-month-olds identified a similar score of 66% for a short MSL questionnaire (Brooks & Meltzoff, 2015). The expected percentage according to the data by Kristen et al. (2012) would be approximately 49% at the age of 27 months, which is close to the identified 44%. However, it is important to note that the questionnaire used by Kristen et al. (2012) had a total of 69 items, whereas the one used in the current study had a total of 139 items plus the additional knowledge scale, which involved a situational assessment of children's usage of 'know'. Therefore, it is very well possible that additional words in the questionnaire reduced the overall performance by several percent.

Another point that speaks for the sample being comparable to prior studies is the diversity of children's vocabulary. The current sample had a similarly diverse vocabulary when considering the categories of the MSLQ, as has been found in Kristen et al. (2012). Overall, 69% of the 27-month-olds produced at least one term from each category, which is appropriate considering the 81% found by Kristen et al. (2012) for 30-month-olds. Also the hypothesis that children use mental state terms much earlier than previously assumed (Harris, Yang, & Cui, 2017) could be confirmed with 73% of 27-month-olds already producing the verb 'know' according to the MSLQ questionnaire. Additionally, 70% of children were already able to use the verb 'know' to talk about the location of a toy in the first scenario of the knowledge scale. The questionnaire only assessed mental state terms that children were actively using themselves.

A final finding in relation to MSL production that has already been identified by Bretherton and Beeghly (1982) and Bartsch and Wellman (1995) is the order of word acquisition. Volition terms have been commonly found to be acquired earlier than cognition terms, which has been confirmed by several subsequent studies (Kristen et al., 2012, 2011). This is in line with the order predicted by the TT and could also be confirmed in the current sample, with most children knowing more volition terms and performing generally better on the volition scale compared to the cognition scale. As has been mentioned in the introduction, one of the main reasons for the order in acquisition is believed to be the parallels to belief-desire reasoning (Wellman et al., 2001; Wellman & Liu, 2004). Children's mental state talk appears to mirror the order of acquisition from desire understanding to belief understanding. This is in line with what previous studies have found in mothers, especially considering that mother's mental state talk has been shown to predict children's mental state talk (Ruffman et al., 2002). Previous work by Ruffman et al. (2006) has shown that mothers adjust the mental state terms they use depending on the child's age and abilities. With 15-month-olds, mothers most commonly refer to desires, whereas those references switch to the cognitive states of others at the age of 24 months. Also Kristen-Antonow et al. (2018) found that at 7 months mothers commonly use more desire terms than cognition terms, which switches at the age of 24 months.

Overall, the current sample shows MSL production that is in line with previous studies and can therefore provide insight into children's early understanding of the semantics and pragmatics of mental state terms.

Mental state language production and metacognition of ignorance

In order to determine whether children not only use mental state terms, but also use them to denote mental states, MSL production was correlated with children's awareness of their own ignorance. For this purpose, they were administered a task with pictures of familiar and unfamiliar objects that they had to name while signs of uncertainty were coded. The idea was that awareness of one's own ignorance requires the ability to identify one's state of ignorance. If children are able to use the term 'to know' for a situation in which they are ignorant, this suggests that they have understanding of one scenario in which the verb 'to know' is appropriate and describes a mental state, namely the state of lacking information. This in turn provides evidence towards a relation between children's ignorance and the respective mental state verb.

Bartz (2017) found that young children are much more likely to respond with signs of uncertainty when confronted with unfamiliar objects compared to familiar objects and that this tendency increases with age. The author therefore concluded that children who show signs of uncertainty for unfamiliar objects have some internalized concept of knowledge and ignorance as they treat cases in which they are ignorant differently from cases in which they are able to answer. This finding was confirmed in the current study. The 27-month-olds showed significantly more signs of uncertainty when confronted with unfamiliar items compared to familiar items. Overall, 86% of children showed a sign of uncertainty on at least one unfamiliar trial. A total of 13% of children showed signs of uncertainty on every single unfamiliar trial. But despite children showing more signs of uncertainty for the unfamiliar objects, the most common response to unfamiliar objects were misapplications of words or inventions which were not considered signs that denote awareness of one's own ignorance. However, only 14% of children used misapplications or inventions without showing a single sign of uncertainty, whereas 74% used both, signs of uncertainty and misapplications/inventions. Furthermore, a total of 12% showed signs of uncertainty without misapplications or inventions. Cases in which children used misapplications and inventions may be explained by children's tendency to overestimate their own knowledge (Flavell, 1999). Furthermore, children may have been eager to participate and give an answer and therefore preferred to invent a word or use another in order to reply. Overall, in accordance with the conclusion that Harris, Ronfard, and Bartz (2017) and Bartz (2017) have drawn for 25- to 36-month-old children, 27-month-olds produce a number of signs of uncertainty for trials in which they lack knowledge and appear to have metacognitive awareness of their own ignorance.

If children possess awareness of their own ignorance, one may wonder if there are other mental states that children are able to demonstrate at such a young age and that might provide additional evidence towards either a concept of those mental states or a functionally equivalent accumulation of production rules. Children's general tendency to show signs of uncertainty on trials with unfamiliar objects correlated with their ability to answer questions about their own visual perspective in the verbal VPT task, independently of general language skills. It is important to note that the signs of uncertainty considered in this correlation are both verbal and nonverbal ones. The relation implies that children who show preverbal or verbal signs of uncertainty during situations in which they lack knowledge are also more likely to be children who show verbal understanding of their own perspective. As it is not language driving the relationship, one might argue for a common construct, which might be a concept of mental states as the ST suggests.

If children appear to have awareness of their own ignorance, the next question is in which way they express that ignorance, implicitly or verbally. In addition to the trial-based score, a qualitative score was computed to assess whether children used mostly implicit (nonverbal) signs of uncertainty or verbal signs of uncertainty that included the mental state term 'know'. The average score was 3 points out of 18, which indicates that children mostly used nonverbal signs of uncertainty when asked to name an unfamiliar object. When looking at the production of the mental state verb 'know' to indicate ignorance, only 20% used the formulation 'I don't know' at least once, suggesting that a majority of children did not use mental state verbs to express their MCI. This is in line with the findings by Bartz (2017) who also found that young children mostly used nonverbal signs of uncertainty and that the verbal signs increased with age. Additionally, Bartz (2017) analyzed children's flipping gestures between 14 und 42 months and found that around 22 months about a fifth of the sample produced such flips to indicate ignorance, which was closely followed by verbal statements about ignorance that increased in frequency between 26 and 30 months (see Harris, Bartz, & Rowe, 2017, for a review). Thus, nonverbal signs appear before children use explicit verbal labels to denote their ignorance.

If young children possess awareness of their own ignorance and produce mental state terms, the question remains how the production of mental state terms that indicate knowledge (i.e. 'to know') relates to this awareness. In the current study, there was a significant correlation between the metacognition of ignorance task (MCI) and children's production of mental state terms according to the MSLQ. However, this relationship was driven by children who used the formulation 'I don't know' and only remained significant for a single scale of the MSLQ after controlling for general language, namely the knowledge scale. First of all, the lack of a correlation with the trial-based score indicates that there is no language-independent association between having awareness of one's own ignorance (be it verbal or nonverbal) and being able to produce a mental state term. The correlation with the qualitative score on the other hand suggests that it is indeed about verbal signs of uncertainty as children received a higher score if they used 'I don't know'. Still, not even for a verbal sign of uncertainty the mere production of mental state terms appears to be enough, as the correlation was only independent of general language for the knowledge scale. The knowledge scale assesses whether children's production of 'know/don't know' occurs in various naturalistic scenarios and does not only ask for the production of the verb. Especially the first scenario is indicative as it is concerned with the child using 'know/don't know' spontaneously or when they are wondering about the location of a toy. A significant correlation suggests the following: first, that parents appear to be accurate at judging their children's flexibility with the usage of the verb 'know/don't know' and second, that children who use it in naturalistic situations to denote their own ignorance according to their parents, also use it to denote their own ignorance in the MCI task. In summary, it seems that 27-month-old children possess awareness of their own ignorance, but that awareness is not related to the mere production of mental state verbs, but rather to the production of mental state verbs in naturalistic contexts.

Given the relation that children who use verbal signs of uncertainty also use mental state terms in naturalistic contexts, a possible conclusion might be that children who show verbal signs of uncertainty including a mental state verb, have the appropriate associations between their metacognitive representation of their own ignorance and the mental state verb that can be used to express it. One objection to this conclusion may be that children's usage of 'I don't know' does not necessarily have to indicate a mental state for as long as it brings about a specific effect. Brandt et al. (2011) argued that children struggle to use 'to know' flexibly with first and third person and might therefore learn it as 'chunk' instead of as a verb with several applications. In addition, Rubio-Fernández (2019) brought forward that children who use 'I don't know', may use it as an excuse to avoid answering. Both studies suggest that 'I don't know' is only a verbal formula without any relation to mental states. However, in their analyses of spontaneous speech, Harris, Yang, and Cui (2017) have identified that children rarely use the mental state verbs for a third person. Thus, one of the reasons for the discrepancy in the study by Brandt et al. (2011) may be that caregivers are less likely to use third person formulas with mental state terms early on and mostly refer to themselves and the child during early development. The lack of experience with third person mental state verb usage might explain the difficulty of switching from a first person application of a mental state verb to a third person application. However, this difficulty does not need to indicate that children lack awareness of their own knowledge. In particular, Harris, Yang, and Cui (2017) found that children use mental state verbs to talk about their own knowledge frequently and do so spontaneously. The second objection, namely that children might be simply refusing to answer by reporting that they don't know the answer (Rubio-Fernández, 2019) would suggest that children only use 'I don't know' in uncomfortable situations during which they would prefer not to answer. There are several aspects speaking against this objection. First of all, the present correlation between the knowledge scale of the MSLQ and children's verbal signs of uncertainty emphasizes that children also use the formulation in order to talk about the location of their favorite toys and during a game. So there are naturalistic contexts in which children have no reason to avoid an answer as they use the mental state terms spontaneously. This is also supported by speech data by Harris, Yang, and Cui (2017) in which most uses of mental state verbs were spontaneous. Furthermore, one may argue that there are simpler ways for a child to avoid answering than to reply with 'I don't know', namely by remaining silent or even denying an answer by shaking their head. Replying by using a mental state verb might therefore be a sign for a representation of children's ignorance instead of a mere formula to avoid answering. Thus, one may conclude that children who use verbal signs of uncertainty including a mental state verb (i.e. 'I don't know') in the MCI task, do so in similar naturalistic situations and associate it with expressing their own ignorance.

Despite some children using the mental state verb appropriately to denote their own ignorance, it is important to keep in mind that the mere ability to produce the verb 'know' according to the MSLQ, does not automatically mean that children are able to apply the verb in the MCI task. A majority of children did not use verbal signs of uncertainty (79%) even though many of these children were already able to produce the mental state verb 'know/don't know' according to the MSLQ (73%). This was confirmed by a McNemar test showing that it was more likely for children to produce the verb 'know/don't know' according to the MSLQ and not use it in the task, than use it in the task and not be able to produce it according to the MSLQ. This is most likely due to the appropriate application of a verb being much more difficult than merely producing it, no matter if in imitation or spontaneously. Since a majority of children were already able to produce the verbs, one may argue that the acquisition of verbs occurs first and subsequently, through sufficient experience, children are able to apply mental state verbs to talk about their mental states. Still, it remains unclear whether children create associations between mental state verbs and their metarepresentations before they are able to use them in a task.

Thus, a question that remains is whether children who do not express their ignorance verbally in the MCI task, but already produce the verb 'know/don't know' according to the MSLQ, show more nonverbal signs of uncertainty than children who do not. Is the nonverbal representation of their ignorance related to their usage of the mental state verb? In the current study, children who showed signs of uncertainty, were not significantly more likely to produce the verb 'know/don't know' according to the MSLQ. This suggests that although children seem to have a nonverbal representation of their own ignorance in the MCI task and already have terms to denote it, there is no relationship between the two as of yet for these children. This leaves several questions.

First, whether children's nonverbal comprehension already represents conceptual understanding, or whether only through learning mental state verbs in various contexts children acquire conceptual understanding. Possessing awareness of one's own ignorance is no aid in the production of mental state terms for children, which might suggest that their awareness is not a fully developed concept which can easily be associated with the terms as might be expected from a ST approach. If children's awareness of their own ignorance is rather an accumulation of production rules as the TT suggests at this age, it seems feasible that each production rule needs to be associated with the mental state verb separately. Thus, their MCI does not provide a benefit in the acquisition of mental state terms as each situation would need to be experienced separately. To explain the missing relationship, ST might argue that preverbal concepts are only created in respect to the few situations that children are able to observe. Preverbally only a limited number of contexts can be experienced, therefore, the concept created is only preliminary and its association with a mental state verb requires additional contexts for it to become properly applicable. However, even then the situations that were already experienced preverbally should be easily associated with the mental state verb if they are included into a single concept. TT has less difficulties to explain the missing association as production rules are created for the few observed situations and each production rule has to be associated separately with a mental state verb for each context.

A second question is: Independent of whether children have concepts or production rules, which abilities are necessary for children to link their mental state terms with their nonverbal awareness of their own ignorance? Unfortunately, that question cannot be answered with the current data. Correlations between the qualitative score and general language and cognition at 24 months may suggest that children need basic language and cognition skills to show verbal signs of uncertainty. However, it is more likely that children require exposure to adults using mental state verbs to denote their knowledge and talk about mental states. Possibly, there is a need for more situational experience with the verbs before children are able to appropriately use them for situations in which they lack the knowledge to answer. Ruffman and Taumoepeau (2016) have found that not the frequency with which the mental state verb 'want' was used mattered for MSU, but rather the amount of different contexts in which it was used. In addition, Harris et al. (2005) have argued that conversations about different minds and perspectives are what aids children in understanding mental states. It is very well possible that through conversation about mental states and the usage of mental state terms by the caretakers, children acquire the association between their preverbal awareness of their own ignorance and the mental state terms that can be used to denote it. Through experiencing several situations it is possible to infer the true meaning of a verb, which is not dependent on the specific context it was used in. In a future study, it would be of major interest to investigate the relationship between contextual exposure of children to mental state verbs (e.g. by their mothers or in an intervention) and the subsequent application of these terms to experimental situations that require the expression of mental states.

Overall, there appears to be a relationship between MSL production and MCI for children who use verbal signs of ignorance (i.e. 'I don't know'). These children appear to have associated the metarepresentations of their own ignorance with the mental state verb. Children who only show nonverbal signs of ignorance (a majority), however, do not show any relations to their mental state vocabulary. This may be due to them being unable to associate their awareness of ignorance with a mental state verb, or because they do not do so in the current task. In case of the first, one might argue that MCI is not present in the form of concepts, but production rules. Therefore children need extensive experience to associate a cluster of production rules with the respective mental state verb. In conclusion, mere production of the verb 'know/don't know' does not go hand in hand with the ability to denote one's own ignorance through it. One reason for this might be a lack of exposure towards mental state verbs and in particular how they can be used to denote one's own or another person's mental state. The theoretical question whether children's awareness of ignorance is present as a concept (and only needs brief exposure) or whether it is present as production rules (and needs exposure for every single rule) remains to be answered more conclusively in future work.

Mental state language production and pragmatic inferences

In the pragmatic know-want task, children performed at floor for the *know* condition, whereas they were significantly above chance in the *want* condition. Children's performance was only dependent on whether they had to react to the experimenter using 'know' or 'want' and on whether the participating child was a boy or a girl, as girls performed significantly better than boys. This performance difference was not due to general language. Furthermore, passing the *want* condition was significantly easier than passing the *know* condition.

For the know-want task it was hypothesized that since children are able to draw correct inferences in nonverbal communication (O'Neill, 1996) and use mental state terms early and to talk about their own and others' knowledge states (Harris, Yang, & Cui, 2017), they should also be able to draw correct inferences from mental state terms. Especially, if children's preverbal competence is a sign of conceptual understanding and associations between the representation of another person's knowledge or desire and the respective mental state verb to denote it is already present, then children should be able to respond adequately to a speaker's desire or need for information. This was not the case in the current sample, because children only performed above chance on the *want* condition. One possibility is that children were not able to produce the respective terms yet and therefore inferences could not have been possible in the first place. However, children produced 36% of volition terms and 24% of cognition terms. For the separate knowledge scale that had three scenarios, children performed at 53%. Even when looking at the particular terms, 86% of children already produced the verb 'want' according to the questionnaire and 73% produced the verb 'know/don't know'. This suggests that the children of the current sample were able to use both verbs recruited in the task and about half of them used the word 'know' in more than one context. This is in line with Harris, Ronfard, and Bartz (2017) who suggested that children at the age of two are already able to use mental state terms. Therefore, lack of productive vocabulary could not have been the reason for chance performance on the *know* condition of the know-want task, especially considering that even for the *want* condition, children did not perform at ceiling.

Thus, despite production of MSL, children did not perform accordingly on the pragmatic task for 'want' and 'know'. Two possibilities might explain this discrepancy which will be elaborated in detail in the following. Either, 1) the task was operationally more difficult independent of the mental state verbs used, 2) children's production of cognition terms and in particular the verb 'know' does not reflect their comprehension of the inferences associated with the verb.

In order to solve the task, children were expected to infer the right action from the mental state verb used. In the case of the statement 'I want x, I don't want y' they were expected to hand the experimenter object x as that is the object that was requested. In the case of the statement 'I know where x is, I don't know where y is', the inference is a slightly different one. As the experimenter provides information about their own ignorance, the child has to infer the need for information and provide it for the item whose location is unknown. In order to solve both conditions, children need to be aware of the implied meaning of both verbs. The current results suggests that 27-month-old children were not able to make this inference for the verb 'know', but showed competence in the case of 'want'. In preverbal communication tasks, set ups usually involved a single unfamiliar hidden object that children were not able to reach on their own. The main effect was found for children's pointing actions and children pointed more when the parent was not aware of the location (i.e. looked away) (Behne et al., 2012; Liszkowski et al., 2007; O'Neill, 1996). Apart from additional boxes and verbal cues instead of visual cues, the tasks can be considered comparable. It is unlikely that additional boxes or the need to memorize locations was the main reason for children's inability to solve the task, especially since they were able to handle these demands for the *want* condition. However, one may argue that there were linguistic demands that kept children from passing the *know* condition and that were not related to their actual understanding of the verb.

One such possibility is the negation of a sentence (Nordmeyer & Frank, 2014), as in the *want* condition the affirmative sentence indicated the correct object, whereas in the *know* condition it was the negated sentence. But if the affirmative sentence had indeed been used as a cue for the correct answer, then performance on the *want* condition should have been higher, whereas performance on the *know* condition should have been significantly below chance, which was not the case. Furthermore, a study by Austin, Theakston, Lieven, and Tomasello (2014) found that 27-month-olds were able to comprehend negation in a search task when given both affirmative and negative cues. The same was the case for a study by Reuter, Feiman, and Snedeker (2018), who showed that 2-year-olds required an affirmative to establish context for a negation. Given that this was the case in the current study with both affirmative and negation present for each trials, negation is unlikely to be the reason for children's difficulty with the *know* condition.

Another possible objection concerning the structure of the task might be the requirement to infer knowledge from the statements of a third person. Harris, Yang, and Cui (2017) found in their speech analyses that children only rarely refer to the knowledge states of a third person while using the term 'know' frequently to converse about their own or an interlocutor's knowledge. However, as the first experimenter was not the one asking the test question and the second experimenter addressed the child by stating their knowledge in first person format (i.e. '*I* want the ball'), it is unlikely that children struggled with the format of the questions. Furthermore, children already ask information-seeking questions early in the third year of life, which also suggests that they have a basic idea of another person's knowledge as a source for their own, even if they do not explicitly contrast it with their own in their speech (Harris, Yang, & Cui, 2017).

Thus, it seems like the complexity of the task is not the reason for children's low performance, but rather something else[6]. Next, I shall address the interpretation that children's production of the verb 'know' does not reflect their comprehension of the inferences associated with the verb.

[6] See study 2 for more detailed answers to possible limitations of the task design

Harris, Yang, and Cui (2017) have argued that children use mental state terms to talk about their own knowledge states, but also those of others. Additionally, children's information-seeking questions imply that they are able to act on their own ignorance and find a way to gain information (Harris, Bartz, & Rowe, 2017; Harris, Ronfard, & Bartz, 2017). From these two findings one may expect that children should have early associations between their knowledge states and mental state verbs that can be used to talk about them. However, results in the previous section on the MCI task suggest that only a few children in the current sample are able to denote their own ignorance with a mental state verb, but the majority of children do not yet have associations between their metarepresentations of their own ignorance and mental state verbs. Thus, it is very well possible that only few children comprehend the semantics and pragmatics of the mental state verb 'know'. If an association is what children require in order to be able to draw correct pragmatic inferences from what a speaker knows or does not know, one would expect a correlation between verbal signs of uncertainty in the MCI task and performance in the know-want task. In the current sample, there was a floor effect for the *know* condition which makes it difficult to interpret relations between the two tasks. However, when looking at the distribution of children that show verbal signs of uncertainty and children who pass the *know* condition in a crosstable, expressing one's own ignorance verbally does not make it more likely to draw the correct pragmatic inferences from the verb 'know'. Thus, one may argue that an association alone does not go hand in hand with the ability to make pragmatically adequate inferences. The question remains what exactly it is that children require in order to be able to solve the task.

A peculiarity that has to be noted in relation to the verb 'to know' is that it may have two meanings. In the first, 'to know' can be meant in the sense of familiarity and does not require a source of knowledge, whereas the second postulates a present knowledge source. In the case of the children in this study, it may be that instead of using 'to know' with a source of knowledge in mind, children are mostly producing it in order to signify their familiarity with the topic in question. One might therefore argue that children understand their own lack of knowledge in relation to the production of 'know' before they are aware of sources of knowledge. However, if children were able to report their own lack of knowledge independent of knowledge sources, they should still be able to talk about what they know and do not know in an EPT task, which was not the case in the current study as shall be discussed in the next section.

If production of MSL and an association between MCI and a mental state verb are not sufficient in order to be able to draw the right inferences from the verb 'know', then what else is necessary? One finding that speaks for a relationship between MSL production and the ability to draw correct pragmatic inferences for 'want' is the correlation between the *want* condition and the MSLQ questionnaire. However, this relationship did not remain significant when controlling for general cognition with the exception of the knowledge scale, which assessed contextual usage and not only production. This is of particular interest as one would expect the volition scale to be strongly related to children's performance on the verb 'want' and not necessarily their ability to use 'know/not know' in various contexts. The correlation with the knowledge scale might be a further indicator of mere production not revealing children's comprehension and ability to draw inferences, but rather that there are additional aspects that matter for children's ability to differentiate between a desired and an undesired object and to help a speaker by identifying it. What differentiated the knowledge scale from the other subscales was that it not only assessed production of the mental state verb 'know' in general, but particular naturalistic contexts in which the verb was produced. A correlation between the knowledge scale and the *want* condition therefore indicates the need for a deeper understanding of mental state terms, one that goes beyond mere usage of the verb and involves production in relevant situational contexts. One might argue that cognitive aspects of the task and of requirements made by the knowledge scale might be what drove the correlation, however, it remained significant when controlling both, general language and general cognition. Ruffman and Taumoepeau (2016) have shown that the contexts in which 'want' is produced matter for children's MSU and not the frequency of production, so it is conceivable that the same applies for 'know'.

In particular, children who experienced more contexts for ‘know’ may have already encountered ‘want’ in variable contexts. Thus, these children are also the ones that master the *want* condition whereas children with less experience are less likely to do so. However, if application in the correct situational context is of great importance, the question remains why children who already use ‘know/don’t know’ in the appropriate context were unable to draw the right inferences in the *know* condition. In the case of the *want* condition, situational experience might be enough, but the *know* condition appears to require more than that. It is important to keep in mind that only because children show awareness of their own ignorance and use a mental state verb to indicate it, it does not mean that they are able to do the same for the state of ignorance of another person. Thus, it is possible that children are not yet able to identify the speaker’s ignorance based on the statement ‘I don’t know where the ball is’ and to draw inferences about pointing towards the right location as a pragmatically correct response. Children may need to learn to associate ignorance with the need for information, which is not the same as identifying their own state of knowledge. Moreover, it is not clear whether children are able to recognize ignorance in a third person from the mental state verb alone. Possibly, the task may have been easier for children if they had observed the speaker’s experience with one object but not another and thereby received nonverbal cues about the speaker’s knowledge state. This would have to be investigated in future research with the speaker observing the hiding of one object, but not another before requesting help. However, in this case it may be difficult to distinguish the effect of 1) the nonverbal observation and 2) conclusions that the child is able to draw from the mental state verb alone. If nonverbal observations are necessary to draw conclusions, this might once again argue for children using production rules inferred from behavior instead of concepts for their preverbal MSU competence.

The know-want task reveals that mere production of the mental state term ‘know’, even in addition to associations with children’s own ignorance, is not sufficient to draw correct pragmatic inferences from a contrasting statement about knowledge at the age of 27 months. But experience with situational contexts appears to play a major role for the mastery of pragmatic inferences for the mental state term ‘want’.

The order of acquisition with competence in ‘want’ appearing before competence in ‘know’, is in line with previous research that also finds desire terms before cognition terms. This also fits well with the above mentioned transition from desire to belief-understanding (Wellman et al., 2001). In general, it appears that MSL production reflects the pattern of desires before beliefs with pragmatic inferences about desires developing before pragmatic inferences about knowledge states.

Overall, MSL production alone is not sufficient for drawing correct pragmatic inferences for mental state verbs. Children at the age of 27 months seem to be able to make correct pragmatic inferences for the verb ‘want’, but struggle with the verb ‘know’, similarly to the commonly found order of volition before cognition terms. Despite drawing inferences for knowledge states much earlier preverbally, 27-month-olds appear to lack sufficient comprehension of the semantics and pragmatics of the verb ‘to know’ to do the same verbally. One possibility might be that children need additional situational experience with a verb before they can associate inferences that they draw preverbally with the corresponding mental state verb, as seems to be the case with ‘want’. But this explanation is not sufficient for ‘know’ as most children already showed experience with ‘know’ in naturalistic contexts. Given that drawing the right pragmatic inferences requires not only identifying the ignorance of another person, but also the awareness that they need information, it may be that the 27-month-olds of the current sample were not yet able to infer another person’s need for information from their ignorance[7]. In a future study, it would be important to investigate children’s ability to label their own ignorance versus labeling the ignorance of another person, followed by a test of their pragmatic inferences based on these labels of ignorance. Additional research might make it possible to identify the exact boundaries of the pragmatic inferences children are able to draw from mental state verbs.

[7] Adopting the view that children act according to production rules, it is feasible that separate rules code the other’s need for information and therefore need more experience to associate what children do preverbally with the mental state verb.

Mental state language production and epistemic perspective-taking

The epistemic perspective-taking (EPT) task was included into the study in order to determine how well children are able to report their own and somebody else's knowledge states[8]. It was expected that despite low performance on the task as has been found in the original study by Gonzales et al. (2018), 27-month-olds with higher scores in MSL production would perform better on the task. As in the original study, the current sample struggled with the task and performed low on both, the *self* and the *other* condition, with the *self* condition (13%) being marginally easier than the *other* condition (8%). In the original study, Gonzales et al. (2018) found a performance of 14% on the *self* condition and 10% on the *other* condition for 2-year-olds. The sample ranged from 24-35 months and had a mean age of 31 months, which is several months older than the current sample. Considering the difference in age, the performance found in the current study is in line with previous work that found competence between 3-4 years of age (Ruffman & Olson, 1989; Woolley & Wellman, 1993).

Despite low performance, it was hypothesized that solving the EPT task might be easier for children who are already able to produce the respective mental state verbs, 'know' in particular. Harris, Yang, and Cui (2017) have also found that children frequently make affirmations of both their own knowledge and that of others, so it was expected that children would be able to report their own and others' knowledge better as long as they were competent in producing the relevant mental state verb.

Even though only a small percentage of children passed the *self* condition, it is still of interest to examine whether the relation between producing the mental state verb 'know' according to the MSLQ and passing the *self* condition of the EPT task is random. In the current study, children were significantly more likely to already produce the mental state verb 'know' without passing the task, than pass the task without being able to produce the verb 'know'. The same was the case with talking about the location of a toy by using 'know/don't know' according to the MSLQ (knowledge scale) and

[8] The VPT task will be discussed in the next section on concurrent relationships.

passing the *self* condition of the EPT task. Overall, this suggests that as above, the mere production of MSL does not necessarily involve the ability to use these verbs to report one's own knowledge. Interestingly, all of the children who passed the EPT *self* condition were already able to use 'know/don't know' to talk about the location of a toy. This might again suggest the necessity of experience with the mental state verbs in naturalistic contexts. As reported in the previous section, it may be that there is a developmental trajectory from identifying one's own ignorance and representing it, over using mental state verbs to denote it and then followed by the ability to report one's own knowledge when asked.

Judging by the percentage of children who passed the respective conditions, it appears that the *know* condition in the know-want task was easier for 27-month-olds compared to the *self* and *other* conditions of the EPT task. This might be due to the fact, that the know-want task did not require a verbal reply and could have been solved by handing the experimenter the correct box. The EPT task on the other hand, required children to assess their own (and another person's) knowledge about a box and answer accordingly. The fact that 20% of children were able to reply with 'I don't know' to spontaneously indicate their lack of knowledge for the MCI task, but struggled to do so when asked explicitly in the EPT task (13%) is in line with previous research by Kim et al. (2016), who showed that children were able to identify their own lack of knowledge, but overestimated it when asked explicitly. The same applies to 4- to 5-year-olds children claiming they have always known a new piece of information (Taylor, Esbensen, & Bennett, 1994). Thus, it may be that children are aware of their own ignorance and some of them can denote it spontaneously in the MCI task, but still a majority of children struggle to report their own knowledge when asked explicitly.

Overall, the evidence suggests that despite children's ability to produce the verb 'to know', it is difficult for them to answer questions about their own state of knowledge and that of another person. Thus, the production of MSL alone is not sufficient to conclude one's own knowledge state or that of another person from evidence provided. This is in line with 1) findings for the know-want task, which also suggest that despite producing the terms, children are not yet able to use them to draw pragmatic inferences, and 2) findings for the MCI task, which showed that a majority of children produce the verbs and seem to have some basic understanding of their own knowledge states, but do not yet show a relationship between the two.

Conclusions

The first objective of this study was to identify whether children who already produce mental state terms and appear to use them in appropriate contexts, also possess comprehension of the semantics and pragmatics of these terms. In particular, the question was whether children who produce mental state terms have created associations between MSU that they show preverbally and the terms that can be used to denote their MSU. Recent work by Harris, Bartz, and Rowe (2017); Harris, Ronfard, and Bartz (2017); Harris, Yang, and Cui (2017) in analyses of spontaneous speech data, suggests that children use mental state terms early and in appropriate contexts, in addition to possessing an early metacognitive awareness of their own ignorance. The evidence provided by the relationship between 27-month-olds' MSL production and their performance on various tasks assessing their comprehension of 'know', suggests that merely producing mental state terms in an experimental setting does not entail that children have elaborate understanding of mental state terms. Furthermore, the mere production of mental state terms does also not entail an association between children's MCI and the mental state verb that can be used to denote it. Children are able to draw correct pragmatic inferences for the mental state verb 'want', but struggle when it comes to the verb 'know'. In addition, despite using the formulation 'I don't know' to appropriately indicate their own ignorance, children are still unable to answer questions about their own knowledge states and those of others in an EPT task. Overall, even though a majority of children produced 'know', only a very small percentage associated their own ignorance with the mental state verb and was able to solve the comprehension tasks in this study.

Even though Harris, Ronfard, and Bartz (2017); Harris, Yang, and Cui (2017) have argued that children's early use of mental state terms is appropriate and might hint at genuine MSU before the commonly believed age of 4, it cannot be concluded that this is the case for a majority of children with the current evidence. Children produce mental state terms early on, however, it is difficult to draw conclusions about the association to their own MCI since only 20% of the 27-month-olds produced the mental state verb 'know' to talk about their own ignorance. Only a minority of children appear to produce MSL appropriately to denote their own mental states at the age of 27 months, whereas a majority is not yet able to associate their lack of knowledge with the corresponding mental state term. The same was the case for drawing correct inferences from a speaker not knowing the location of an object in the know-want task and talking about their own and others' knowledge states. Children struggled to infer a speaker's ignorance and need for information from the mental state verb 'know', despite a majority already producing the term. But if children can produce mental state terms without being able to solve comprehension tasks and without associating the terms with mental states, the question is what children require beyond the mere production of the terms.

It is important to note that Harris, Yang, and Cui (2017) have worked with spontaneous speech data, which include a variety of situations that are difficult to model in an experimental setting. Furthermore, naturalistic contexts involve a variety of cues that may be nonverbal and therefore aid children to apply mental state terms to situations. Possibly, children would have performed better on the know-want task if given a demonstration of the speaker looking inside one box, but not the other before giving their knowledge statement. Also in the EPT task children may have performed differently without explicit questions about their knowledge. Naturalistic situations are also more likely to involve explanations and additional feedback by the caregiver (Bartsch & Wellman, 1995) that is not present in experimental tasks and therefore makes solving them more difficult. Children's ability to use nonverbal cues around them at an early age might make an association between their mental states and verbs to denote them superfluous before the usually observed age at which they show competence (Bartsch & Wellman, 1995; Kristen-Antonow et al., 2019). One might even argue that the presence of concepts as hypothesized by the ST might not be necessary early on, as children can draw all the inferences they need from observing behavior and creating production rules.

However, it is also possible that children use mental states more commonly in situations that have not been investigated in the current study. For example, children are shown to ask information-seeking questions (Harris, Ronfard, & Bartz, 2017) and have a tendency to question their own knowledge, but not that of another person (Harris, Yang, & Cui, 2017). Additionally, almost all uses of 'know' that have been identified by Harris, Yang, and Cui (2017) were spontaneous and not repetitions of what another person said to them. This suggests that tasks in which spontaneous responses are coded, similarly to the MCI task, might be better ways to model children's natural speech.

Still, experimental paradigms allow us to control for various aids, for example preverbal cues, and thereby make sure that we assess children's cue-independent comprehension of mental state terms. The results of this study suggest that without additional support or aids, most children are not able to associate their preverbal metarepresentations of their own ignorance with the mental state verb 'know'. Children also fail to infer a speaker's ignorance from a statement involving 'know' and to conclude that the speaker requires information. If one stays with the idea that children possess early concepts of mental states, one possibility is that children at the age of 27 months do not have sufficient social and communicative experience to be able to associate the MSL verbs that they use with their representations of mental states that they show early on (O'Neill, 1996). It is much more difficult to talk about something that the child knows as it is not easily observable and has less behavioral consequences compared to something that the child wants. Therefore, it might be necessary for parents to use the terms in several contexts to give a richer view of what the term can denote and that it refers to a mental state and is not used in order to bring about an effect (i.e. to not have to reply anymore as with 'I don't know') (Montgomery, 2002; Rubio-Fernández, 2019). Additionally, awareness of one's own ignorance might be the first step before children are able to denote more complex knowledge states and especially the knowledge states of others. A few children already appear to have sufficient experience and use the term 'know' in situations that indicate they are referring to their own ignorance, but a majority of children still lack associations between their mental representations and verbs that can be used to denote them.

Thus, a preverbal metarepresentation of own ignorance appears to be present in most 27-month-olds, but the association to MSL and pragmatic inferences that can be drawn from it only develops later on. Whether it occurs at the previously hypothesized age of 4 according to the traditional view (Bartsch & Wellman, 1995; Bretherton & Beeghly, 1982; Shatz et al., 1983) or before that is difficult to answer with the current evidence. Future studies should investigate children's performance at later time points and examine the effect of repeated exposure to mental state verb usage in different contexts. Furthermore, the comprehension of children's own ignorance and the ignorance of another person based on a mental state verb should be examined to determine whether children first develop an association between a mental state verb and their own mental states, before they are able to make the same associations for the mental states of others.

Even though only a few children showed associations between their metarepresentation of ignorance and the mental state verb 'know', it appears that these associations begin to develop in the third year of life and most likely gain importance with age. Since mental state verbs are the first step to verbally express a mental state (in this particular case one's own ignorance), the next question is what role MSL plays for the transition from early preverbal mental state understanding to later mastery of explicit ToM tasks.

3.4.2 Concurrent relationship between mental state language and preverbal and verbal mental state understanding

A second objective of the current study was to investigate the role of MSL for the transition between early preverbal skills and later mastery of explicit ToM tasks. Thus, concurrent relationships between MSL and preverbal and verbal MSU were examined by focusing on abilities that develop throughout the third year of life, in particular verbal and action-based perspective-taking and implicit FBU. FBU is considered to be the gold-standard for measuring ToM, however, as implicit FB has shown low reliability, perspective-taking was taken as a simpler model with the potential for both, preverbal and verbal competence. Specifically, the question was whether MSL relates to both nonverbal and verbal perspective-taking and thereby argues for continuity in concepts (continuity account) and at the same time presence of mental state concepts preverbally (ST) or whether it only shows a relation to verbal MSU and thus, suggest preverbal understanding to be only due to production rules (TT).

To this date, only one study has investigated the relationship between level-1 verbal perspective-taking and MSL (Chiarella et al., 2013), finding a significant relationship independent of general language at the age of 30 months. For preverbal perspective-taking no study has been done as of yet. Concerning FB understanding, only the relationship between explicit FB understanding and MSL has been investigated until now, with a strong relation between early MSL and later performance on explicit FB tasks. Significant correlations between both preverbal and verbal perspective-taking and MSL would speak for continuity in the development of MSU during the third year of life. In the following, concurrent relationships between MSL production and 1) perspective-taking and 2) implicit FBU are discussed separately.

Mental state language production and level-1 visual perspective-taking

In order to investigate concurrent relationships between MSL production and level-1 VPT, two types of tasks were used. The first one was a verbal task developed by Gonzales et al. (2018) that also included EPT, and the second one was an action-based perspective-taking task developed by Moll and Tomasello (2006). It was hypothesized that children should perform better on the action-based perspective-taking task given that Moll and Tomasello (2006) have found competence in 24-month-olds.

In the verbal perspective-taking tasks, 27-month-olds did not perform significantly above chance in any of the tasks or conditions. A GEE model revealed a significant effect of task type (visual vs. epistemic) and condition (self vs. other), but no interaction. This indicates that it is possible to arrange the conditions in order of visual before epistemic and self before other, but it cannot be determined whether children master visual other before they master epistemic self as children performed similarly on the *other* conditions of both task types. This is mostly in line with the model computed by Gonzales et al. (2018) as the authors found an effect of task type, condition but also sex, suggesting that girls were more likely to pass the task than boys. In the original study, the 2-year-olds were already able to pass the visual *self* condition, whereas this was not the case for our sample of 27-month-olds. Theoretically, the observed pattern is in line with a hybrid approach that combines the idea of concepts being acquired hierarchically as predicted by the TT (belief before desire, perception before knowledge) and the idea of ST that the self is more accessible than the other and thus acquired earlier.

An additional finding is that most children in the current sample overestimated what they could see and replied with ‘yes’ towards the animal that was hidden behind an occluder. This could not be due to priming through the experimenter as the order of ‘yes’ and ‘no’ during prompts was randomized over children. A McNemar test revealed that more children had a bias towards answering with ‘yes’ than towards answering with ‘no’, suggesting a tendency to overestimate what they could see. Gonzales et al. (2018) predicted that approximately 50% of children should be able to pass the visual *self* condition around the age of 31 months. For 27 months the prediction was at about 35%, which is several percent off from the observed value of 30% in this study. However, it is important to note that the sample of 2-year-olds comprised 50 children in total and had a very diverse age group (24-35 months). When only looking at younger children between 24 and 30 months ($N = 18$), performance was at about 33%, which is comparable considering the small sample.

Concerning the effect of sex found in the previous study (Gonzales et al., 2018), but not in this one, it is possible that differences are more obvious in older children and less so at the age of 27 months. In adults, women have been argued to have heightened social motivation and empathy (Baron-Cohen, 2010), whereas men appear to perform better on spatial perspective-taking tasks (Meneghetti, Pazzaglia, & De Beni, 2012). But even these differences seem to depend on several external factors such as fear of low performance and sex-associated expectations and can be reduced (Tarampi, Heydari, & Hegarty, 2016). An additional possibility is that differences in general language and comprehension might have been the source of sex differences in the original study. Such differences were not prominent in the current sample, but were not examined in the original one.

In the action-based perspective-taking task, 27-month-olds gave the experimenter the occluded object significantly more often than the visible object, which implies that they had a grasp of what the experimenter was able to see and what not. The original study was performed with 18- and 24-month-olds, out of which only the 24-month-olds consistently gave the experimenter the occluded toy. The performance measured for the 24-month-olds was at 67% which is considerably higher than the 57% observed in the current study. This difference may be due to two factors. First of all, the sample in the study by Moll and Tomasello (2006) comprised only 18 children for the 24-month-olds and was therefore considerably smaller than the 60 children in the current study. Second, the original study does not report on excluding children who did not have a minimal amount of trials, leaving the question open whether children with only a single trial were kept in the sample. The current study only considered children who had at least 3 valid trials for the final sample. However, also when including children with only 1 trial, performance remained the same and far from the 67% observed by Moll and Tomasello (2006) in 24-month-olds. A possibility is that 24-month-olds had less reasons for exclusion of trials as children are less mobile at that age and therefore less likely to pick up both objects instead of just one. At the age of 27 months, it is conceivable that children were mobile enough to speed up the process of getting two new toys to play with by handing the experimenter both old toys at once. Such trials were excluded in both studies, but might not have been as common in the work done by Moll and Tomasello (2006).

Overall, when looking at the performance of 27-month-olds on VPT tasks, they seem to struggle with the most basic questions on their own perspective in a verbal task, but are able to pass an action-based perspective-taking task that assesses their ability to conceptualize what another person can or cannot see. Thus, they are not able to converse about perspectives, not even their own, but seem to be competent in judging what another person can see nonverbally.

Even though the tasks were different in difficulty level, they should correlate if they measure the same underlying construct. The results of the current study support this conclusion, given that there was a correlation between children's performance on the *self* condition of the verbal perspective-taking task and their performance on the pointing measure of the action-based perspective-taking task. This relationship remained significant when controlling for general language, supporting that it is not comprehension of the task instructions and the ability to reply that drives the correlation between the two tasks. It is important to keep in mind that due to floor effects in the *other* condition, the correlation is only present for the *self* condition of the verbal perspective-taking task, however, also awareness of one's own visual perspective is an important part of the concept of perspective-taking and requires mental representation (Flavell, 1974). Another interesting possibility is that the verbal concept of one's own perspective thrives on the preverbal concept of the perspective of another. However, also a production-rule-like association between the two tasks might be possible with children using verbs to express the rules that they have acquired preverbally, especially since children show neither above chance competence on their own verbal perspective, nor on another's.

If both tasks are measuring a construct that is inherent to perspective-taking and requires a mental representation (Flavell, 1974), be it in the form of production rules or a concept, the question remains what role MSL production plays for the relationship between these two types of tasks. As depicted in the literature review, MSL is relevant for both, precursors of ToM (Kristen et al., 2011) but also later performance on explicit FB tasks (Brooks & Meltzoff, 2015). However, its role for the relationship between preverbal and verbal tasks concurrently has not been investigated yet. Based on the only study that examined how MSL production relates to verbal perspective-taking (Chiarella et al., 2013), a positive relationship between children's verbal perspective-taking abilities and their MSL production was expected. As the relationship was shown to be independent of general language at the age of 30 months, the same was hypothesized for the 27-month-olds in this study. The expectation was partially confirmed as there was a significant relationship between the *self* condition of the verbal perspective-taking task and children's score on the MSLQ. In particular, children's own verbal perspective-taking correlated with the knowledge scale, even after controlling for children's general language at 24 months. An advantage of the current study is the knowledge scale as it assessed whether children already used the formulation 'I don't know' in three naturalistic contexts according to their parents. When examining the relationship between children's ability to report their own perspective (30%) and their ability to use 'I don't know' in various naturalistic contexts (53%), there was a significant relationship independent of general language. This suggests that using mental state terms in different contexts (i.e. 'know/don't know') is related to the ability to describe one's own perceptions correctly. One may thus argue that MSL is related to early, verbal representations of mental states, more specifically those of one's own perception. In the current study, the early onset of this relation could be shown with more specificity than in the study by Chiarella et al. (2013). Taken together, with the findings by Chiarella et al. (2013), there is support for continuity in the mental domain and thus, the view that MSL is mapped onto one's own (and others') visual perspectives quite early in development. This supports the idea of highlighting perspectives in discourse that has been identified as a potential mechanism for ToM development (Harris et al., 2005; Lohmann & Tomasello, 2003) as caregivers provide context for visual perspectives by emphasizing them verbally through MSL. Still, the causal role that MSL production in naturalistic contexts might play for the acquisition of explicit visual perspectives and whether MSL production builds on a preverbal representation of visual perspectives remains to be investigated.

An additional hypothesis was that the relationship between perspective-taking and MSL would not depend on particular subscales of the MSLQ, as Chiarella et al. (2013) have found correlations with every subscale. This was partially the case in the current study, however, not independent of general language, and only applied to the *self* condition of the VPT task. Specifically, the relationship between MSL and perspective-taking, independent of general language, only applied to the knowledge scale. The language-independent correlation with the knowledge scale implies that it may be situational experience with the terms in different contexts that fuels the relationship between perspective-taking and MSL. This is plausible as independent of whether one uses mental state terms like 'want', 'see' or 'know', all of them include applying these terms to different people and differentiating between perspectives. But what enables an association between a representation of one's own perspective and a mental state verb might be experience with the situations in which a mental state verb applies to one's own perceptual state, independent of the type of verb that is used. This finding is of particular importance as it shows that there might not yet be a relationship between perspective-taking and the mere production of mental state terms, independent of general language, at the age of 27 months (which appears to change at the age of 30 months according to Chiarella et al. (2013)), but there is already a language-independent impact of children being able to use mental state terms in naturalistic contexts.

There were several differences between the design employed in the current study and the one employed by Chiarella et al. (2013) and arguably such differences, especially in the task used, might have contributed to incongruent results. In the following I will address these differences.

The first difference is that the current study only found a relationship to the child's perspective-taking ability in relation to their own perspective, whereas Chiarella et al. (2013) found a general relationship to VPT. This difference is partially due to the types of tasks used and the way they were coded. The VPT task used by Chiarella et al. (2013) was adapted from McGuigan and Doherty (2002) and consisted of children being shown a card with a picture on one side and none on the other. The child was asked whether she could see the picture and whether the experimenter could see the picture. The questions were repeated after the card was turned. In total, children were able to score 4 points, one point for each correct answer. In contrast, in the current study children had to answer two questions for each condition in order to be considered competent. Furthermore, they were confronted with two percepts at once, namely one object they could see and one they could not see while in the task by Chiarella et al. (2013) there was only one percept at a time. On the one hand, if coded similarly, the current task might have resulted in higher performance while still including children's ability to judge someone else's perspective, but on the other hand, the measure would have been less conservative and children might have replied correctly by chance. A better and more robust measure is to take both *self* questions and evaluate them together as has been done by Gonzales et al. (2018) and in the current design. Additionally, it is important to consider that the children in the current study were younger and therefore might have shown less competence. Overall, these task differences may have contributed to the difference in children's performance, however, the observed relationships with MSL across two different perspective-taking tasks and with three months age difference between the samples, support the robustness of the current finding.

Differences in strength of the correlation between MSL production and VPT may have also been due to age. In the study by Chiarella et al. (2013) the correlation had an r of .50 (.40 when controlling for general language). In the current study, the r of the correlation was at .38, reducing to non-significance and .19 when controlling for general language. The children in the current study were 27-month-olds, whereas the children in the study by Chiarella et al. (2013) were 30-month-olds. It is likely that the relationship between perspective-taking and MSL production is present during the second year of life, but strengthens with age. Especially considering that perspective-taking performance was relatively low, it is likely that only when children are able to perform above chance on perspective-taking measures, the relationship to MSL is independent of their general language skills.

Still, the current study was able to identify a relationship between children's ability to describe their own perspective and their usage of mental state verbs and a language-independent relationship between their ability to describe their own perspective and their usage of epistemic mental state verbs in different contexts. Similar to what I have outlined above, it appears to be important that children learn to use mental state terms in the right scenarios and the mere production of the terms is not enough. Possibly the importance of these scenarios becomes less prominent with age as more children are able to make these experiences, but at the age of 27 months it is only usage across various contexts that shows a general language independent relationship with perspective-taking. Thus, one can argue that there is an early association between children's representations of their own perception and the MSL terms necessary to report it, but that this relation is only general language independent for children who apply mental state terms (e.g. 'know') across contexts.

This hints at MSL playing a role for verbal perceptive-taking, now the question remains whether there is a relationship between MSL and a preverbal measure of perspective-taking. In the case of a relationship between preverbal perspective-taking and children's production of MSL, continuity between children's preverbal comprehension and their MSL may be assumed, which supports the idea of ST that children already possess concepts of mental states early on. However, in the case of no relationship between the two, the idea that children create production rules based on their experiences and that each production rule needs to be associated with a mental state verb seperately appears to provide a better explanation. ST would have to explain why children's preverbal concepts are not associated with mental state terms if a concepts is already present and should only need a few instances to be associated with a verb. In the current study, despite children performing above chance in handing the experimenter the toy she could not see, there was no significant relationship to children's MSL production according to the MSLQ. This suggests that children showed compre-

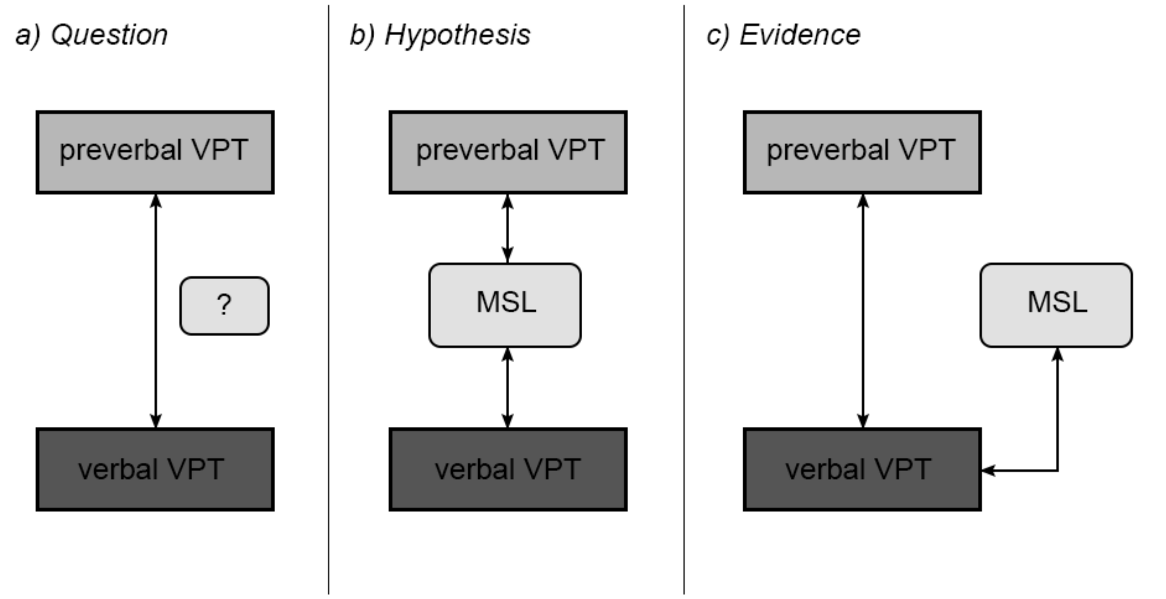

Fig. 3.10: Overview of how the present evidence fits to the hypothesis on the relationship between mental state language and visual perspective-taking. *a)* a relationship between preverbal and verbal visual perspective-taking is expected if both tasks measure an underlying concept, but the nature of that relationship and what fuels the transition is unknown. *b)* mental state language was hypothesized to play a role for the transition between preverbal to verbal visual perspective-taking, similarly to the role it plays for precursors and later explicit ToM mastery. If preverbal concepts are present, mental state language might make them explicit by associating them with a mental state verb that can be used to express them. *c)* evidence of the present study suggests that there is no concurrent relationship between preverbal visual perspective-taking and mental state language production, but that mental state language production matters for verbal visual perspective-taking. The current results suggest that it is more plausible for preverbal perspective-taking competence to result from production rules that children create early on and not from a concept of mental states.

hension of perspectives non-verbally and were able to produce mental state terms in naturalistic contexts according to their parents, but their production of MSL did not indicate that they associate mental state terms with what they are able to do preverbally. Thus, MSL production does not relate to nonverbal VPT concurrently, but contextual experience with mental state verbs is relevant for the verbal representation of children's own perspective. See Figure 3.10 for a visual depiction of the identified relationships.

Due to the correlational nature of the study, it is difficult to conclude whether it is MSL that enables children to talk about perspectives, or whether it is awareness of perspectives that leads to associations with the words. Future studies would have to examine this relationship in detail and administer several perspective-taking tasks with varying difficulty to determine which tasks children are already able to solve verbally at a young age. Additionally, a longitudinal study may investigate the predictive relationship between preverbal perspective-taking, MSL and verbal perspective-taking further and determine whether preverbal perspective-taking predicts later MSL production, but is not related to it concurrently. If children's preverbal perspective-taking is indeed only based on production rules, the amount of these production rules should not matter for MSL and subsequent verbal perspective-taking as each production rule needs to be associated seperately. However, if it is indeed a concept that children possess preverbally, even if that concept is immature, then these children should have stronger predictions from their preverbal competence to their MSL and verbal competence. Concurrently, it appears like MSL production in naturalistic contexts is what matters for verbal perspective-taking, whereas no such associations are present for preverbal perspective-taking.

Because both perspective-taking tasks appear to measure a similar underlying concept, but only the verbal task relates to MSL, it can be considered additional evidence towards the importance of MSL for the verbal denotation of children's mental states. However, the exact process by which children learn to associate their preverbal comprehension of perceptual states (be it concepts or production rules) with mental state terms remains unknown and should be examined in future work. It is plausible that children build their comprehension on production rules which help them to behave accordingly, but that only through experience with MSL and the highlighting of perspectives they establish a conceptual understanding of mental state terms. The current study suggests that production of mental state terms alone is not the main driving force for MSU, but rather the varying contexts in which children learn to apply these terms.

Overall, there appears to be a relationship between verbal VPT and MSL production at the age of 27 months, however, this relationship only applies for children's own perspective and is related to the contextually appropriate production of epistemic terms as reported by their parents independent of general language. In the case of mere MSL production, the relationship is still strongly dependent on general language abilities. Since MSL production did not relate to action-based perspective-taking, one may argue that children do not build their early mental state vocabulary on their preverbal implicit understanding of perspectives. Possibly, children are able to produce mental state verbs before they are associated with their preverbal mental understanding. However, children's early mental state vocabulary matters for verbal expression of their own perspective. One way to create an association between early preverbal understanding of perspectives and mental state terms may be by providing experience with naturalistic contexts in which these terms can be applied. If naturalistic contexts are indeed what is required for children to associate their preverbal competence with mental state terms, then it seems more plausible that their competence is build on behavioral production rules, than on a preverbal concept of mental states. Future studies are needed to compare children's performance on perspective-taking tasks of various difficulties that are both preverbal and verbal at different ages. By comparing performance one can determine the transitional period between nonverbal and verbal competence and identify the exact role that MSL plays at different stages. Since the current study examined the relationship concurrently, it is not possible to draw any inferences about predictions. Thus, additional longitudinal studies are necessary to determine if there is indeed no predictive relationship from preverbal VPT to MSL production and finally to verbal VPT. In a training design the causal relationship between nonverbal perspective-taking and verbal perspective-taking could be examined in detail by training nonverbal perspective-taking and investigating which effects it has on subsequent verbal perspective-taking. It may also be of interest to examine the reverse relationship and see if improving verbal perspective-taking might help children with their nonverbal perspective-taking skills. Finally, the nonverbal perspective-taking task could be used in order to prompt children's verbal denotations.

Mental state language production and implicit false-belief understanding

Implicit FBU is a concept that has been a topic of debate in recent years, especially due to the rising amount of non-replication studies (Kulke, Reiß, et al., 2018; Kulke, von Duhn, et al., 2018; Poulin-Dubois & Yott, 2018). Still, several authors postulate that under certain circumstances, children are able to solve these tasks and seem to use some type of FB reasoning to do so (Grosse Wiesmann et al., 2017). If we assume the presence of implicit FBU, the question is how it relates to explicit FBU. Concurrent studies have shown no relationship between explicit and implicit (Grosse Wiesmann et al., 2017; Ruffman et al., 2001), but longitudinal work suggests developmental continuity (Sodian, 2016; Thoermer et al., 2012). The current study intended on addressing children's preverbal MSU and whether it is build up on preverbal concepts or production rules by examining how it relates to MSL production. Previous studies have shown significant relationships between MSL and explicit FBU (Brooks & Meltzoff, 2015; Olineck & Poulin-Dubois, 2007), but how implicit FBU is related to MSL remains unclear.

Since MSL production relates to nonverbal precursors of ToM (Kristen et al., 2011), but also to subsequent explicit ToM performance (Olineck & Poulin-Dubois, 2007), it was hypothesized that it might provide insight into the relationship between the two. As mentioned in the previous section, it was expected that if children possess a preverbal concept of FB and there is continuity between preverbal and verbal FB, it should be easier for children to associate that concept with MSL, than if they do not.

As in the original study by Grosse Wiesmann et al. (2017), children anticipated correctly in the familiarization trials, suggesting that they understood the task. However, opposed to the original results in 3- and 4-year-olds, the 27-month-old children in this study did not perform equally on the FB1 and the FB2 trials, showing above chance performance on the FB1 trials, but below chance performance on the FB2 trials. Because performance on both types of FB trials differed significantly they were not collapsed into one. In the original study, performance was collapsed and 3-year-olds performed one-tailed significantly above chance and 4-year-olds two tailed significantly.

In general, children performing differently on FB1 and FB2 trials is not particularly surprising and has been found in several studies previously (Grosse Wiesmann et al., 2018; Kulke, Reiß, et al., 2018). One of the reasons is considered to be the additional complexity of the FB2 trial, as it cannot be solved with a simple 'last location' strategy and is therefore a proper measure of FBU. Grosse Wiesmann et al. (2018) argued that a last location strategy is unlikely if children do not perform significantly below chance on the FB2 condition. This was the case in the current sample with 27-month-olds performing significantly below chance for FB2, while performing significantly above chance on FB1. This pattern of results makes a last location strategy particularly likely, especially considering that it was significantly more likely for children to pass the FB1 trials, but not the FB2 trials, than the other way around.

It is possible that the percentage of children who are competent and pass both types trials increases with age and was far higher in the sample of 3-year-olds investigated by Grosse Wiesmann et al. (2017), but at the age of 27 months only a minority of children is able to infer the FB of an agent implicitly, while a majority mostly pays attention to the last location.

Given children's performance on the task, there are several ways of interpreting the results. First, the implicit FB task used in the current study is not a reliable measure of implicit FBU and therefore does not give us any insight on the relationship between children's implicit MSU and their production of mental state terms. Second, it does measure FB accurately, but at the age of 27 months only a minority of children are able to solve the task and a majority use a 'last location' strategy. Alternatively, one may argue that FB1 does already measure some competence, but that the additional length of FB2 and the necessity to keep in mind several more actions while the chaser is gone make it difficult for children to respond accurately.

In order to dissociate the interpretations mentioned above, it is crucial to consider which theoretical perspective is adopted in order to discuss these options. The continuity account proposes that there is a fluent transition between children's early implicit skills and later explicit skills, but that explicit tasks have additional demands that make it impossible to identify children's early competence (Baillargeon et al., 2010). Thus, by using implicit tasks, one is able to show children's early competence by reducing task demands. The expectation would therefore be that the current implicit FB task is a functioning task for as long as the demands posed by the task are not too high to reveal competence. The question is therefore whether children are struggling because the task demands are still too high in order to show competence. This can also be applied to the discrepancy between the FB1 and the FB2 trials. Possibly the FB1 condition is showing children's early implicit competence, but the FB2 condition has too high demands to reveal that competence. Two argument speaks against this interpretation, (1) the lack of a relationship between children's performance on FB2 and the inhibition or cognition measure at 24 months and (2) the fact that the two FB measures do not correlate with each other. If children required better executive functions or cognition in order to keep track of the agent's belief throughout the FB2 trials, one would expect children with better inhibition skills to perform better on the FB2 condition. Interestingly, this was not the case, but there was a trend towards the opposite relationship for learning and the FB1 condition. Children who performed better on the FB1 condition appeared to be worse at learning the association between a cue and the appearance of an object at the age of 24 months. It might therefore be possible that children with worse learning skills were more likely to perform better on the FB1 trials as they adopted a simple strategy in order to pass it. Lack of learning skills for the FB task may mean that children were not able to associate the agent with the respective goal (the chasee) and therefore focused on the last location of the goal (the last box the chasee was at). This implies that the implicit FB task might therefore not be an accurate measure of implicit FBU, at least not when the two types of FB trials are considered separately. If FB2 is used as a control condition for FB1 and only those children who were correct in both FB1 and FB2 are considered passers, then the second interpretations of the results appears to be the most reliable one. Namely, that the task measures implicit FBU, but that children at the age of 27 months are not yet able to implicitly track an agent's beliefs. Furthermore, this inability to track beliefs does not depend on their inhibition or cognition skills. This conclusion argues against the continuity account, at least in relation to implicit FBU[9], as one would expect children at the age of 2 to be able to pass implicit FB tasks that do not require inhibitory or cognitive skills. Additionally, if the current task (in particular the FB1 condition) does not address implicit FBU, but a more rudimentary tracking of goal-directed actions, then one would not necessarily expect a correlation with MSL production. In order to address the relationship further one might have to look at older children and investigate whether they perform better on the implicit measure used in the current study and whether their performance relates to MSL production.

If we adopt the theoretical perspective of the dual system (or two systems) account, one would expect children to perform better on an implicit FB task, as implicit tasks can be solved by using system 1, whereas later explicit tasks require system 2 (Apperly, 2010, pp.172-189). System 1 mostly works by tracking an agent's belief-like states and also enables children to solve level-1 perspective-taking tasks. Perner and Roessler (2012) proposed that children create 'experiential records' for agents that help them to solve implicit tasks early on. Evidence that did not find a relationship between implicit and explicit measures was one way to argue towards a separation between the two systems (Grosse Wiesmann et al., 2017). The simplicity of system 1 would presuppose that children are able to use it much earlier, but only a small percentage of children in the current study was able to solve both trial types. If we assume that FB1 is a valid measure of implicit FB according to system 1, then one would expect children's performance on the FB1 condition to relate to their performance on level-1 perspective-taking (Apperly, 2010, pp.77-95). Verbal perspective-taking tasks have not been found to correlate with FBU (Sodian & Kristen-Antonow, 2015), but a nonverbal task has (Yeung et al., 2019). Also Thoermer et al. (2012) have found a correlation between implicit FBU and implicit visual-perspective taking, which is in line with the dual systems account. However, in the current study neither the

[9] This does not mean that children do not show preverbal MSU and only applies to implicit FB reasoning.

action-based perspective-taking measure, nor the verbal perspective-taking measure were related to children's performance on the implicit FB task. This suggests that independent of task format, there is no evidence for a relationship between implicit FBU and level-1 perspective-taking in the current sample. This suggests two options: Either, 1) children's performance was not due to them using system 1 in order to answer correctly and therefore the task might be more difficult than what system 1 is capable of, or 2) the task is not necessarily difficult but distractors in the design keep children from focusing on relevant cues. If children used a simple strategy like the 'last location' strategy in order to solve the task without focusing on an agent's experiences and behavior, then there is no reason to expect that children will require system 1 in order to solve it. Rubio-Fernández and Geurts (2013, 2016) emphasized that focusing on the agent is one of the crucial factors that enable children to pass implicit and explicit FB tasks. Thus, it may be possible that the absence of the agent from the scene make it more difficult for children to focus on the agent and thereby keep track of the agent's 'experiential record' to solve it (Perner & Roessler, 2012). Overall, if one adopts the stance of the dual systems account, the task appears to have been too difficult for system 1, either due to additional demands or due to a focus on the goal instead of the agent.

After examining the position of the continuity account and the dual systems account, it seems like there is little evidence that speaks for the FB1 condition as a valid measure of implicit FBU on its own. Several other studies have observed differing performance on FB1 and FB2 (Grosse Wiesmann et al., 2018; Kulke, Reiß, et al., 2018) and argued for a necessity to combine the two in order to control for other factors influencing children's performance (i.e. a 'last location' strategy as in the current study). One might argue that children performed below their actual abilities because of a focus on the goal instead of the agent (Rubio-Fernández & Geurts, 2013). However, the more likely explanation appears to be that 27-month-olds are not yet able to track an agent's implicit FBs to the degree that has been found in prior studies and prefer to use a 'last location' strategy to anticipate an agent's action.

Since low performance on the task makes it difficult to interpret null correlations, there are still several open possibilities that may explain how MSL may relate to implicit FBU. On the one hand, there may be a significant relationship between implicit FBU and MSL in older children that already show competence in the implicit task. On the other hand, one may argue that implicit FBU does not require MSL for the basic representation of beliefs and that this representation only requires MSL much later when it becomes verbal. All these options would have to be investigated further in future studies.

Overall, only few children are consistently able to anticipate an agent's actions based on his false belief and the ability to do so is neither related to general language, cognition or executive function skills, nor to a precursor of explicit ToM like MSL production. The low performance on the task makes it difficult to conclude that 27-month-olds are capable of implicit FBU and it remains to be investigated whether this skill improves with age. A study by Grosse Wiesmann et al. (2018) found similar performance on FB1 and FB2 and also only for older children around the age of 4. It might be possible that despite popular belief, implicit FBU develops mostly in parallel to explicit FBU and is therefore too difficult for children at the age of 2. Future studies should focus on confirming children's performance on the new task developed by Grosse Wiesmann et al. (2017) over time in a similarly large sample and investigate further whether later competence is predictive of explicit ToM. For this purpose, it is important to keep in mind that children's implicit performance at a later age might already be influenced by their explicit performance. Finally, it is especially important to replicate implicit FB tasks in young children as recent work suggests that children might not be as competent as earlier studies concluded (Kulke, Reiß, et al., 2018; Kulke, von Duhn, et al., 2018).

Additional task relations

One important additional finding that should be addressed is the relationship between the action-based perspective-taking task and 1) the verbal perspective-taking task and

2) the *want* condition of the know-want task. As mentioned above, there was a significant relationship between the pointing score of the action-based perspective-taking task and the *self* condition of the verbal perspective-taking task. Interestingly, this did not apply for the giving score of the action-based perspective-taking task. However, when examining the relationship between the *want* condition of the know-want task and the action-based perspective-taking task, it is only the giving score that correlates significantly. This is an important finding as the pointing and giving score differ on how much effort the child has to invest in order to receive a high score. For the pointing score, it was sufficient to point towards the correct toy, whereas for the giving score it was necessary to not only identify what the speaker was looking for (namely the hidden object), but also to register that an action was required, namely to give the experimenter the hidden toy. When comparing the know-want task and the verbal perspective-taking task, the same difference in behavior can be observed. The verbal perspective-taking task only requires a brief response to the speaker, whereas the *want* condition in the know-want task involves additional inferences and behavior, namely that the speaker wants to receive a particular toy that the child is asked to provide.

One may argue that these relationships are due to general language skills and the children in the current study were less likely to give a toy or to pass the *want* condition if they were not able to comprehend the request, however, the correlations between the task types remained significant when controlling for general language. It was also not the additional motor effort as the relationship between the *want* condition and the giving score remained significant when controlling for children's fine motor skills at 24 months. Thus, it may be that children who gave the speaker the toy she couldn't see were inferring her desire, namely to receive the toy similarly to the way they did for the *want* condition of the know-want task. Possibly the pointing score only requires children to identify what the speaker was unable to see, whereas the giving score has the additional requirement of inferring the speaker's desire for the object. Furthermore, the action-based perspective-taking task can be argued to measure children's understanding of the subjectivity of desires, which is a skill that children appear to be able to solve at the age of 24 months (Carlson et al., 2004; Repacholi & Gopnik, 1997).

These relations between the action-based perspective-taking task, the verbal perspective-taking task and the *want* condition of the know-want task, provide additional evidence that there is an association between children's early representations of mental states and mental state terms that can be used to denote them, however, this relationship is only present for scenarios in which children demonstrate the ability to use mental state terms appropriately which was not a requirement in the mental state language parental questionnaire. Since desire terms are acquired before cognitive terms, one might argue that children already had enough experience with the verb 'want' in discourse and have therefore been able to associate what they can do preverbally with the terms. Whether their preverbal competence is based on production rules or concepts still remains to be answered, but the specificity of the situation lets production rules appear more plausible.

Conclusions

The second objective of study 1 was concerned with bridging the gap between early nonverbal competence in MSU and later mastery of explicit ToM tasks. The third year of life was recognized as a transitional period and therefore used in order to assess how MSL production, which is considered to be an early sign of explicit MSU, concurrently relates to preverbal and verbal perspective-taking and implicit FBU. The results of the current study support the idea that MSL production and verbal VPT are related, but that this relationship is mostly driven by usage of mental state verbs (e.g. 'know') in various contexts and otherwise still strongly influenced by general language abilities. No relation between MSL production and action-based perspective-taking was found, which makes the view plausible that children's preverbal competence is not based on concepts as the ST suggests, but rather on production rules that need to be associated with a verb that denotes them one by one (theory theiry). However, as both the verbal and action-based perspective-taking task were related and the action-based perspective-taking task showed a relation to children's ability to infer desires from the mental state verb ‘want’, one may argue that it is not mere production that relates to early representations of perspectives, but production in appropriate contexts. Children may need to experience a variety of contexts in which mental state verbs are used in order to create conceptual understanding. Finally, the interpretability of implicit FBU appears to be questionable at the age of 27 months as only a limited percentage of children showed competence and most used a ‘last location’ strategy. Thus, children at the age of 27-months do not seem to be competent in implicit FB.

Overall, the current results argue for continuity between verbal perspective-taking and MSL production in the domain of mental state understanding and support a relationship between preverbal perspective-taking and MSL for tasks that show appropriate use of mental state terms. This combination of results makes a TT-based interpretation of children’s preverbal competence more likely than a ST-based interpretation as limited contextual associations resemble specific production rules more than a general concept. The current results do not confirm this view conclusively, but ST has some explaining to do. Furthermore, the reliability and interpretability of implicit FB tasks remains questionable. In the future, it would be of particular interest to investigate these relationships further by validating implicit FB tasks with older children. Additionally, one needs to determine how children connect their preverbal representations of perspectives with MSL terms if preverbal representations are not what drives children’s production of mental state vocabulary. A longitudinal study that begins preverbally might be able to identify predictive relationships between children’s early mental representations of perspectives, their later MSL and subsequent verbal expression of their mental representations of perspectives.

3.4.3 Basic cognitive and linguistic predictors of mental state language production

The third and last objective of study 1 was to investigate which basic cognitive and linguistic skills at the age of 24 months are significant predictors of MSL production at the age of 27 months. To date, no study has investigated the relationship in a longitudinal fashion and attempted to answer which skills are particularly critical for children developing an appropriate mental state vocabulary. It was hypothesized that inhibitory control would play a major role as it has been found to correlate with MSL production (Bellagamba et al., 2014; Carlson et al., 2004). Moreover, language was expected to play a major role as general language skills have commonly been found to predict later MSL acquisition (Kristen et al., 2011; Ruffman et al., 2002). When looking at important predictors of explicit FBU, language and executive functions have been identified in meta-analyses (Devine & Hughes, 2014; Milligan et al., 2007), but the spectrum of relevant abilities in those domains was broad. In the current study, significant correlations suggested basic cognitive and motor skills as additional entries for a predictive model. As expected, inhibition was a significant predictor of later MSL production, however, unexpectedly it was not all aspects of general language, but only sentence production that predicted later MSL production significantly. In addition, general fine motor skills were also found to be a significant contributor. In total the model explained 45% of the variance in MSL production at 27 months.

The importance of language is straight-forward as MSL is a subcategory of language. However, the fact that it is sentence production in particular that plays a role implies that children need to reach a certain complexity in their verbal utterances before they are able to produce mental state terms. One possibility is that 2-word utterances and more enable children to have an understanding of 'aboutness', of the idea that a mental state is about something, possibly an object or events in the world or even non-existent things (Astington, 1993, pp.67-70). It is difficult to talk about desires or epistemic states without the ability to refer to what exactly one desires or knows. This is of particular interest as complement syntax is thought to be important for the development of FBU (J. G. De Villiers & Pyers, 2002). The importance of sentence production does not imply the need for syntax yet, but it emphasizes the need for more complex language output before children are able to take the next step and use mental state terms. Since all subscales of the MSLQ correlated with sentence production, it seems like not only desire terms and epistemic terms require a basic sentence length, but also physiological and emotional terms. One possibility may be that children who are quicker to use longer sentences, are more likely to receive more complex input from their caregivers. In the future it would be interesting to investigate whether the role of sentence production is reduced if parental input is included in the model.

In addition to sentence production, inhibitory control was an important predictor for children's MSL production. Inhibitory control can be considered crucial as it might aid children in communicating about their own and other's mental states and keep the distinction in mind, particularly to suppress their own mental states as is commonly the case for FBU. As the study by Bellagamba et al. (2014) was only an exploratory study with a small sample, the current work extends the findings on the relationship between inhibition and MSL production by confirming the importance of inhibition for later production of mental state terms. The importance of inhibition for MSL also further supports its close relationship to later ToM abilities and its role as early ToM.

A particularly interesting finding is the importance of fine motor skills as a predictor of MSL production. Several studies have already identified that language abilities and fine motor abilities are closely related early in development and that impairments in one area commonly co-occur with another (Iverson, 2010). One of the main reasons for this relationship is that producing language, in particular, is a motor act and involves a movement (Thelen, 1991). Iverson (2010) argues that children's developing ability to manipulate objects gives them the possibility to pay attention to details of these objects which then again goes hand in hand with learning words that relate to particular parts and could not have been identified before. Additionally, words for actions are learned in the process and while a child is performing them. It is very well possible that a mother would be more likely to use the term 'want' with a child that is reaching towards an object, than with a child that does not exhibit such reaching motions. Therefore, it is comprehensible that aside from sentence production and inhibition skills, also fine motor skills play a crucial role for the acquisition and production of MSL. When considering fine motor skills it seems feasible that children who move more and interact more with their surroundings, are more likely to motivate parents to give them feedback and to elaborate on what happens around them. In particular, it would be interesting to investigate the difference between mother's interactions with children who differ in their motor development. Possibly mental state terms are less commonly used with children with lower scores for motor development than for children with higher scores.

Since fine motor skills are important for MSL production and might influence language acquisition and inhibition in return, it was tested whether the relationship between MSL and sentence production and MSL and inhibition remains significant when controlling for children's fine motor skills. This was the case, suggesting that the two predictors are relevant independent of children's fine motor skills.

Overall, the third objective of study 1 aimed to answer the question which early abilities are relevant for the development of MSL production and identified sentence production, inhibition and fine motor skills as the main contributors. To my knowledge, this is the only study to date that attempted to answer this question and did so in a longitudinal design. In the future it might be interesting to add other known predictors like mothers' mental state talk and see what role it plays in relation to already identified predictors. Early identification of predictors enables interventions that might aid children in their acquisition of mental state terms. For example, if motor skills are relevant for the mental state talk input that children receive from their mothers, it may be important to aid children's motor development or instruct mothers to provide the same input in order to avoid setbacks in MSL production.

3.4.4 Conclusions

Overall, study 1 addressed three major questions. The first question was how MSL production relates to children's MSL comprehension. Results suggest that even though children already produce the respective mental state terms and seem to have basic understanding of their own knowledge and ignorance (enough to show signs of uncertainty), they are unable to associate the two consistently and thereby show comprehension of the mental state terms they produce. Young children struggle with drawing the correct pragmatic inferences from terms like 'know' and are unable to report their own knowledge or that of another person. It was hypothesized that experience with verb usage in various contexts is what enables children to connect their representations of mental states with the respective terms for it and that identifying another person's ignorance from a mental state verb and acting accordingly may be more difficult than doing the same for oneself.

The second question was how MSL fits into the transitional period between preverbal MSU and subsequent mastery of explicit ToM tasks. For this purpose the relationship to other preverbal and verbal measures during the third year of life was investigated, in particular visual level-1 perspective-taking and implicit FBU. As expected there was a significant relationship between verbal level-1 VPT and MSL production, however, that relationship was only independent of general language for children's ability to use mental state terms in various contexts and applied to children's own perspective, not perspectives of others. This provides additional insight into the mechanisms of how MSL may aid in associating children's mental state representations with the respective terms, namely by learning to use mental state terms in various contexts and gaining experience. Since there was no relationship to preverbal perspective-taking, it is not possible to argue that early mental state representations are related to children's mental state vocabulary production. An exception are scenarios in which children already use mental state terms appropriately (in particular the verbal perspective-taking task and the *want* condition of the know-want task). Theoretically, it was argued that it is likely that children build production rules for their preverbal competence instead of preverbal concepts as production rules would need for each rule to be associated separately whereas a concept, no matter how simple, should be easier to associate. In the current sample a majority of children were able to produce mental state verbs and showed competence in preverbal perspective-taking, however, associations were only present for a limited set of situations. Since the current study investigated the relationships concurrently, a longitudinal study would be necessary to examine whether early preverbal perspective-taking predicts later MSL, even if there is no concurrent relationship at the age of 27 months. Concerning implicit FBU there were no significant relations and children's performance suggests a preference for the last location strategy which suggests that children at the age of 27-months are not capable of implicit FB.

The third and last question examined which early skills at the age of 24 months predict subsequent MSL production at the age of 27 months. Sentence production, inhibition and fine motor skills were identified as the major contributors to children's later production of mental state terms. Especially motor skills were argued to have a role for the linguistic input that children receive from their caregivers as parents explain more when children are physically active, than when they are not. Considering how important children's ability to use mental state verbs in various contexts appears to be, additional support may be crucial for children to be able to receive sufficient linguistic input and to be able to associate their preverbal mental state representations with the mental state verbs that they acquire later on.

If one adopts the TT and assumes that children's preverbal mental state understanding is not conceptual, the question remains at which stage their understanding becomes conceptual. If it is only when children draw pragmatic inferences and have encountered verbs sufficiently in naturalistic contexts, then a task is necessary to test when they are able to draw such pragmatic inferences. One possibility is to test for the inferences from the verbs 'want' and 'know' at different ages. Thus, the second study in this thesis investigated at which age children are able to solve the know-want task implemented in study 1.

4. STUDY 2: THE DEVELOPMENT OF PRAGMATIC COMPETENCE IN RESPONSE TO MENTAL STATE VERBS IN 2- TO 5-YEAR-OLD CHILDREN

4.1 Objectives and Hypotheses

The main objective of study 2 was to identify at which age children are able to draw correct pragmatic inferences from the mental state terms 'know' and 'want'. As I have been able to show in study 1, 27-month-olds were not able to infer the appropriate inferences from a speaker who states that she does not know the location of an object. The main reason for expecting competence at the age of 2 was that children have been shown to draw similar inferences about a caregiver's knowledge state during preverbal communication (O'Neill, 1996). Furthermore, their early production of mental state terms concerning their own knowledge and that of another person according to Harris, Yang, and Cui (2017) and children's tendency to ask many information-seeking questions between 2- and 3 years of age (Harris, Ronfard, & Bartz, 2017) was an additional argument for expecting 2-year-olds to be able to draw the same inferences verbally. However, as mentioned above, this was not the case for 27-month-old children.

In study 1 it was proposed that children's difficulties with drawing the correct inference about a speaker's need for information may be due to children understanding their own ignorance before they are aware of sources of knowledge. Thus, children are not able to respond appropriately to the task as they are not aware that the experimenter lacked knowledge about the location of a toy and therefore expected to be informed. However, if children show signs of uncertainty and are able to ask information-seeking questions, it is very well possible that children already make this connection, at least for themselves. The alternative option is that making the same inference for another person is more difficult than doing it for oneself, thus, children may be able to recognize their own ignorance (even if they do not express it verbally) and infer the need to ask for information, but are not able to do the same for an interaction partner in the know-want task.

Aside from preverbal communication, children are rather late in their ability to report the knowledge states of another person and only do so at the age of 4 (Gonzales et al., 2018). Verbally reporting on another person's knowledge state might even be considered easier than passing the know-want task, as it only involves identifying the knowledge state of another person, but not inferring the need for information and acting accordingly. In both cases children have to draw conclusions from the mental state verb 'know', but in the pragmatic task they are expected to help and identify the speaker's need for information about the location of a toy. An ability that develops at a similar age to children verbally reporting knowledge states is understanding factivity. Children are only able to explicitly differentiate between a speaker who 'knows' the location of an object or 'thinks' an object to be in a particular location between the age of 4 and 5 years (Kristen-Antonow et al., 2019; Moore et al., 1989). Both factivity and drawing pragmatically correct inferences from the verb 'know' require an understanding of knowledge sources. For drawing inferences children need to understand that a speaker requires an additional source of knowledge, whereas for factivity children require an understanding of knowledge sources in order to judge which one is more reliable. Despite factivity requiring a more complex judgment, both have the common requirement for children to understand knowledge sources.

As both, talking about others' knowledge states and factivity, develop rather late, it was hypothesized that drawing the right pragmatic inferences from the verbs (i.e. that ignorance can be solved by providing information, for example through pointing) might also develop later than expected in study 1. However, identifying the ignorance of a communication partner and inferring the need for information only involves a differentiation between knowing and not knowing and does not require children to differentiate between different levels of certainty. Thus, drawing the right pragmatic inferences from mental state verbs should be easier than explicitly differentiating between a factive and non factive mental state verb. In consequence, it was expected that despite not mastering the know-want task at the age of 2, children might be able to show competence before the age at which they master factivity judgments.

As demonstrated in study 1, children already produce 'know' and even do so in various contexts, but nevertheless the mere production of the verb is not sufficient for comprehension. The rather specific relations between task performance and mental state verb production in particular contexts in study 1 hint at associations that are too specific to be due to a preverbal concept of mental states (ST). If we assume that children do not possess a preverbal concept, but only production rules that help them act appropriately preverbally (TT), a question is at what age children move away from production rules and show more complex comprehension of mental states. The TT postulates that around the age children solve explicit FB tasks, so around 4 years of age, they acquire metarepresentations (or concepts). If we accept the premise that having conceptual understanding is demonstrated through pragmatically adequate inferences from mental state verbs, then children showing competence in such a pragmatic task may be one way to investigate the age at which children acquire elaborate concepts of mental states. Even without the premise, determining the age at which children are able to draw appropriate inferences from mental state verbs, provides insight[1].

For this purpose children were investigated in a cross-sectional study at 2, 3, and 5 years of age. For the 2-year-olds, performance in accordance with the children in study 1 was expected. Based on how early children are able to infer the knowledge state of another person preverbally and infer a speaker's need for information, it was expected that children may already show competence with 3 years of age. However, if recognizing a speaker's need for information from the mental state verb 'know' alone does not develop until children are able to explicitly report knowledge states, it was

[1] The age at which pragmatically adequate inferences are drawn can be compared to the age at which children are able to report knowledge verbally or solve factivity tasks. The comparison enables us to determine which tasks may be easier to solve and why.

hypothesized that performance on the task would emerge similarly to competence in EPT and factivity judgments (Gonzales et al., 2018; Kristen-Antonow et al., 2019). Delayed competence on the implemented task in comparison to competence shown preverbally, may be an additional argument against children possessing conceptual understanding preverbally and instead creating production rules based on their behavioral observations.

4.2 Methods

4.2.1 Participants

Eighty-five monolingual children took part in this study. Out of these, 34 children were 2 years of age (16 female; M_{age} = 27;11 months, range = 24;11 – 28;23), 26 were 3 years of age (17 female, M_{age} = 39;16 months, range = 38;17 – 40;27) and 25 were 5 years of age (16 female, M_{age} = 63;27 months, range = 54;03 – 73;16). Fourteen additional 2-year-old children and sixteen 3-year-old children were tested but had to be excluded as they were bilingual and did not fulfil the inclusion criteria. Additionally, three 2-year-olds and six 3-year-olds were excluded due to refusal to participate in the task. Lastly, one additional 2-year-old and two 3-year-olds were excluded from analysis as they provided only one ‘want’ and only one ‘know’ trial.

Information on socioeconomic status was obtained from the 2- and 3-year-olds but not for the 5-year-olds as they were recruited from kindergartens and the mother was not present during the testing. The maternal educational level was as follows for the 2- and 3-year-olds: 15 mothers (23,8%) had a lower secondary level degree comprising 10 years, 15 mothers (23.8%) had passed the Abitur examination (equivalent to A levels), 33 mothers (52.4%) had a university degree. On average, children had 1 sibling with the number of siblings ranging from 0 to 3.

The study followed the ethical standards for conducting experiments involving humans in accordance with the 1964 Helsinki declaration and its later amendments or comparable ethical standards and was approved by the local ethics committee. Participants were given monetary compensation for travel costs and children received a book as a gift for participation. Following best practice recommendations, I report how the sample size was determined, all data exclusions, all manipulations and all measures in the study (Simmons et al., 2012). An apriori power analysis using G*Power (Faul et al., 2007) showed that a sample size of $n = 54$ would be sufficient to detect a significant effect within an ANOVA-like F test statistic ($r = 0.4$; $\alpha = .05$; $(1 - \beta) = .8$).

4.2.2 Apparatus/Stimuli

The task consisted of a familiarization and a test phase. In the test phase, two conditions were administered, the *want* condition, and the *know* condition. Half the children completed the *want* condition first, the other half started with the *know* condition. Two trials were administered in each task condition. In each trial, four small cardboard boxes of different colors (red, green, white and blue) were arranged on the table between the experimenter and the child. The task required four boxes to enable counterbalancing of object location and distance. The boxes were the same in all trials. They were arranged at equal distance from each other in the following order: red box, green box, white box, blue box. Each box was 8 cm x 15 cm in size. Two small objects were hidden in one box each, while the child was watching. The pairs of objects were a ball and a chair, a bird and a cup, a dog and a car and a toy figure and a spoon. Object location and object colors were counterbalanced over the boxes. No object was ever placed in a box that had the same color in order to avoid salience. Furthermore, the distance between objects was varied over trials with no box, one box or two boxes in between the locations. The order of appearance was the same for all objects (as above), however, the allocation to condition was counterbalanced over trials. See Table 3.2 from study 1 for order of appearance and box allocation of the objects.

4.2.3 Procedure and Design

The 2- and 3-year-olds were invited into a child-friendly research laboratory and each session began with a warm-up in which the experimenters and the child read a short picture book together. The current task was part of a larger task battery which included several tasks on perspective-taking and implicit ToM. The tasks were substantially different in design and material and did not involve feedback. The 5-year-olds were tested individually in a separate room in their kindergarten and performed the task after a short warm-up.

In the familiarization phase, Experimenter 1 (E1) and the child were seated across from each other at a small table, while an Experimenter 2 (E2) sat on the other side of the table between them. In order to familiarize children with the act of pointing and choosing an object without being shy, two items were presented, a book and a plate. E1 asked E2 to indicate the location of the book by saying "Show me the book. Can you help me?" E2 answered with "yes" and pointed to the book. This was repeated with the plate. The same procedure was then repeated with E1 addressing the child to give children the opportunity to apply what they have observed. The familiarization did not involve hiding elements and was therefore not meant to familiarize children with the hiding game itself. Through replying to the question by E1 and subsequently pointing to the respective object, E2 showed children how to respond to the specific requests, namely identification of an object and helping another person. Thus, E2 served as a model for the children. All children pointed at the book and the plate correctly and did not require any prompts to do so. Once the child had pointed at both familiarization objects correctly the test phase began.

In the first test trial, E1 hid two objects in two of four boxes while E2 was absent. The boxes remained on the table throughout all four trials. Before hiding the objects, the child was asked to identify them, e.g. a ball and a chair. This was done to ensure that the child knew the identity of the objects and would recognize them later on. Whenever a child used a different description for one of the objects, such as 'duck' for the bird figure, the child's descriptor was adopted for the ongoing trial. During the introduction of the objects E2 remained seated at the table with E1 and the child. Subsequently, it was explained that E2 would leave the room and upon return would require help from the child ("Let us play the hiding game! E2 will leave and you will help her search."). As soon as E2 had left the room, both objects were hidden in two out of the four boxes in front of the child. Once both objects were hidden, E1 called in E2, who announced the object that she desired to have (*want* condition) or did not know the location of (*know* condition).

In the *want* condition, E2 told the child: "I want the chair. I do not want the ball. Can you help me?" In the *know* condition, E2 said: "I know where the ball is. I do not know where the chair is. Can you help me?" The child was supposed to respond by pointing to or grabbing one of the boxes. If the child did not respond spontaneously, E1 prompted by asking "Can you help E2?" Overall, six of the 2-year-olds and four of the 3-year-olds required a prompt for *want*. Twelve 2-year-olds and six 3-year-olds required a prompt for *know* trials. Two 2-year-olds did not respond to the prompt for *want* and eight 2-year-olds did not respond to the prompt for *know* on one of the two trials and were therefore counted as missing for the respective trial. None of the 5-year-olds required a prompt. After the child had responded, E2 either let the child open the chosen box or opened the box for the child in case the child struggled to do so. After thanking the child, E1 placed the object back on the table. E1 then emptied the other box, placed the two objects in a basket outside the child's view, and brought two new objects which were introduced. Subsequently, E2 left the room and a new trial started.

4.2.4 Coding

Responses were coded from video recordings of the test sessions by two independent coders, indicating the box that the child chose. There were no ambiguous cases and raters agreed in 100% of the cases.

Children's performance was dichotomous and rated with 0 or 1 in both conditions.

A response was rated as correct in the *want* trials if children indicated or gave the box that contained the object E2 'wanted'. In the *know* trials, a response was correct if the child indicated or opened the box that contained the object for which E2 'did not know' the location. All responses in which an empty box was chosen were excluded from analysis. This was done as choosing an empty box is no indicator for children's understanding of want vs. not want and know vs. not know, but rather for their lacking understanding of the task. In order to be rated as competent in the *want* and *know* conditions respectively, children had to pass both trials in each condition. Two trials were chosen to have a more conservative measure that avoids high performance due to false positives.

Statistical analyses were performed with SPSS 25 (IBM Corp. Released 2019. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.). Preliminary analyses have shown no effect of sex, position of assertion and negation, or trial order, therefore these factors have not been considered further. In order to correct for multiple comparisons where appropriate the false discovery rate (FDR) (Benjamini & Hochberg, 1995) was used. The statistical significance level was $p < .05$ for all analyses.

4.3 Results

We used a generalized estimating equations model (GEE; with an unstructured working correlation matrix, a binomial distribution and a logit link function; (Zeger & Liang, 1986; Zeger, Liang, & Albert, 1988)) to assess the effect of condition (know, want), group (2-year-olds, 3-year-olds, 5-year-olds), as well as the interaction between group and condition on task performance. Detailed results are provided in Table 4.1. For the GEE a subsample was used as not all children had both *want* and *know* trials. Ten 2-year-olds were excluded as they refused to participate in one of the trials and seven 2-year-olds and six 3-year-olds were excluded as they chose an empty box in one of the trials. Group was found to be the only significant predictor of performance. Further analyses revealed that the younger the children were, the lower were the odds (.51) for giving a correct answer. Neither the interaction between group and condition, nor the factor condition were significant predictors of task performance.

In a follow-up analysis with the full sample, we looked more closely at performance concerning the two conditions for each age group separately. Table 4.2 shows the proportion of competent children in the *know* and *want* conditions. As children were required to pass both independent trials per condition to be considered a passer and receive a score of 1, chance level was .25 instead of .50. Both 2-year-olds and 3-year-olds were above chance in the *want* condition, but not in the *know* condition, and 5-year-olds were above chance in both conditions. Figure 4.1 depicts the changes in performance over 2-, 3- and 5-year-olds in both conditions. Only at 3 years of age the difference between performance on the *know* and *want* conditions was significant.

Tab. 4.1: Generalized estimating equations results assessing the influence of condition and age group on performance on the know-want task.

Parameter	*B*	*SE*	Wald	df	p value	Exp(*B*)	95% CI for Exp(*B*)	
							Lower	Upper
Group	-.67	.32	4.42	1	.036	.51	.27	.96
Condition	1.42	.82	2.97	1	.085	4.12	.82	20.57
Group by Condition	-.06	.40	.02	1	.876	.94	.43	2.04

Tab. 4.2: Proportion of children (in percent) scored as passers in the know-want task.

	Want			Know		
Age	M (SD)	*N*	*p*-value (**q*)	M (SD)	*N*	*p* value
2-year-olds	.44 (.51)	25	.030 (.045)	.23 (.43)	26	.515
3-year-olds	.71 (.46)	24	<.001 (.006)	.27 (.46)	22	.483
5-year-olds	.76 (.44)	25	<.001 (.003)	.52 (.51)	25	.003 (.006)

Binominal test against chance performance (.25). FDR-corrected p-values are denoted with **q*.

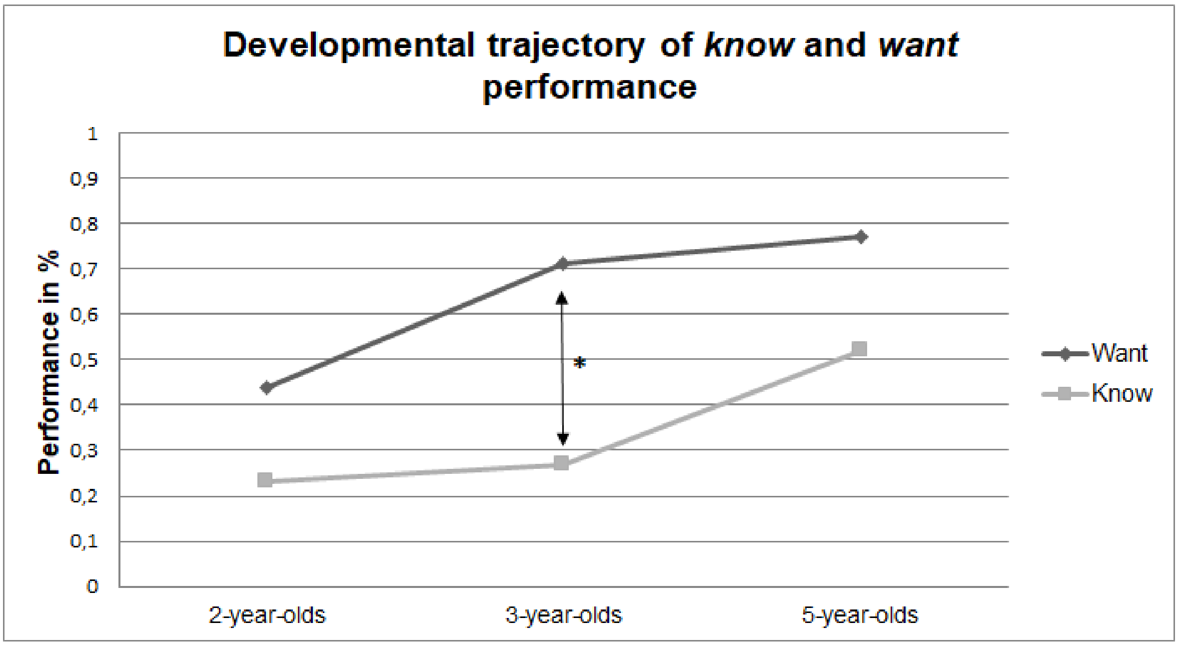

Fig. 4.1: Performance on know and want trials depicted as a timeline over three age groups.

To further investigate the developmental pattern for understanding ‘know’ and ‘want’, cross tables were generated to examine whether passing the *want* condition can be considered significantly easier than passing the *know* condition. A difference in difficulty might suggest that acquiring pragmatic competence for ‘want’ is easier than it is for ‘know’, as has been identified in study 1. Out of the 17 2-year-olds, five showed competence in the *want*, but no competence in the *know* condition, whereas only two children showed the opposite pattern. In the 20 3-year-olds, 12 children showed competence in the *want*, but no competence in the *know* condition, compared to only three children who showed the opposite pattern. Out of the 25 5-year-olds, seven passed the *want* condition, but not the *know* condition, compared to only one child that passed the *know* condition without passing the *want* condition. Table 4.3 shows the results of a McNemar test on the difference in passing the *want* condition vs. passing the *know* condition over all age groups. Similarly to study 1, passing the *want* condition was significantly easier than passing the *know* condition.

Tab. 4.3: McNemar test results on the developmental trajectories of children who passed want before know versus children who passed know before want.

	Know = 0	Know = 1	N	p-value
Want = 0	15	6		
Want = 1	24	17	62	$p < .01$

McNemar's Chi-squared test with continuity correction.

4.4 Discussion

The second study of the current thesis explored 2- to 5-year-old children's understanding of pragmatic implications of a speaker's knowledge and ignorance claims. As study 1 suggested that children at the age of 27 months were not yet able to draw the appropriate inferences despite producing the terms and showing some understanding of their own ignorance, study 2 aimed at investigating the age at which children are competent in inferring the need for information from a speaker's knowledge statement. To this end, children were confronted with desire and knowledge statements in the context of a hiding game. Results indicate that competence in the desire condition (want) preceded competence in the epistemic condition (know), similarly to study 1. While 2- and 3-year-olds reliably handed the speaker the box that contained the object she wanted, only 5-year-olds showed competence in responding to the speaker's statements about what she knew and did not know. As was the case for the previous study involving only 27-month-olds, this pattern of results is compatible with the idea of a developmental sequence from understanding desire terms to epistemic mental state terms (Kristen et al., 2011; Wellman et al., 2001).

In order to solve the task successfully, children were expected to draw correct inferences from statements such as "I want x, I don't want y" and "I know where x is, I don't know where y is". Both statements suggest different inferences, namely handing the speaker object x in the first, while pointing out the location of/giving the speaker object y in the second. Children needed to be aware of the implied meaning of the verbs to be able to act accordingly. The results of the second study suggest that drawing these inferences for the *know* condition was particularly difficult for our sample of 2- and 3-year-olds. Thus, the results that have been obtained in study 1 for 27-month-olds were mostly replicated in this sample. One exception are sex differences that have been identified in study 1, with girls outperforming boys on the *want* condition. It is very well possible that the current sample was too small to determine differences in performance between the two sexes and that a similarly large sample as in study 1 might have illuminated these differences.

Since both 2-year-olds and also 3-year-olds performed low on the *know* condition, the question remains whether their performance was really due to struggles with drawing the right inferences or something else that was inherent to the task. Several alternative explanations have been addressed in study 1 and will be briefly discussed here.

One alternative explanation of the present findings could be that the general complexity of the task, brought about by four boxes instead of two and the necessity to memorize the location of objects, made 2- and 3-year-olds struggle to solve the *know* condition. However, it is important to note that these task demands did not hinder the 2-year-olds of study 1 and 2 to pass the *want* condition. It is thus very unlikely that the observed poor performance can be attributed to overburdening unspecific task demands. An additional hindrance, with respect to the language demands of the task, might be that the processing of a negated sentence posed a problem for most 2-year-olds and some 3-year-olds in both the *want* and the *know* conditions (Nordmeyer & Frank, 2014). In addition, the *know* condition involved additional inhibitory demands since the correct response may have required the children to inhibit a spontaneous tendency to respond by gesturing towards the object the speaker "knew", namely the object without negation, rather than the one she did "not know". It should be noted, however, that even infants seem to be able to inhibit such response tendencies, especially when gesturing towards an absent, rather than a present object or event (Liszkowski et al., 2007). Moreover, a study by Austin et al. (2014) has identified that 27-month-olds were able to comprehend negation in a search task for as long as they were given both affirmative and negated cues (see also Reuter et al., 2018). Furthermore, if children had a preference for 'know' over 'not know', then they should have performed significantly below chance on the *know* condition, which would indicate a tendency to give the speaker the distractor object for which the location was known. Similarly, performance on the *want* condition should have been higher if children interpreted the object with an affirmative statement as the goal object. However, as that was not the case it seems to be unlikely that the negation of the verbs posed a problem for 2- and 3-year-olds.

Aside from addressed differences in affirmation and negation, one possible limitation of the design may have been the absence of the speaker during the hiding of the objects in the *know* condition. Because the speaker claimed to have knowledge about the location of one of the objects, children may have been confused by the source of that knowledge. Although it was not explicitly mentioned how the speaker was able to acquire knowledge about the location of one of the objects, children could make a plausible inference such as "The speaker was told before she reentered". In many game-like social interactions young children are presented with play partners who have some prior information, and in many everyday interactions children are confronted with adults having prior knowledge from unknown sources. There is reason to believe that children accepted the information about the speaker's prior knowledge and did not question its source due to pragmatic constraints.

Independent research shows that young children begin to question a speaker's entitlement to make a knowledge claim around the age of 4 to 5 years, but not earlier (Fedra & Schmidt, 2019). Had this led to confusion or general reluctance to answer the test questions in the task, then the 5-year-old group's performance should have been affected. This was not the case in the current sample. Therefore, it is unlikely that general task demands or linguistic specifications of the task were the reason for 2- and 3-year-olds' poor performance on the *know* condition. Instead, it seems like inferring the appropriate response from linguistic information alone (namely the verb 'know' vs 'not know') is much more difficult than inferring it from cues about another person's knowledge state.

The fact that the current study did not present children with visual information on the speaker's knowledge in the *know* condition can be considered an advantage, as most studies that investigate children's understanding of knowledge states always involve situational cues. In early studies by Liszkowski et al. (2007) and O'Neill (1996) children are given cues on what a speaker has or has not seen by providing them with visual evidence. Also naturalistic contexts that have been investigated by Harris, Yang, and Cui (2017) are not controlled for the influence of prior conversation, gestures and other indicators by parents that could make the interpretation of mental state verbs much easier. The only studies that have controlled for the influence of situational cues were studies on factivity (Kristen-Antonow et al., 2019; Moore et al., 1989). Thus, it is crucial to have experimental studies without the involvement of situational cues that investigate how much information children are able to infer from the mental state verbs alone. The know-want task can therefore be considered a difficult task for the comprehension of mental state terms as children fail the task unless they have the ability to infer the knowledge state of a speaker from the contrast of 'know' and 'not know' alone.

Although previous research indicates that even 12-month-old infants tailor their informing behaviors to a speaker's knowledge (Liszkowski et al., 2007), the choice between an informative and a redundant response when responding to a verbal statement about the speaker's knowledge without situational cues was difficult for the 2- and 3-year-olds in the sample. The current pattern of results indicates that children might be able to use the words appropriately in speech, as suggested by Harris, Ronfard, and Bartz (2017); Harris, Yang, and Cui (2017), but are not able to draw the appropriate inference from the verbs the way they do from social and situational cues. Even the 5-year-olds performed far from ceiling in the *know* condition which further supports the impression that identifying a request for information in the contrast between 'know' and 'not know' is much more difficult than expected.

As mentioned in the objectives of the study, it may be that children are aware of their own ignorance, but do not associate knowledge with sources of that knowledge. Furthermore, it may be that lack of awareness of knowledge sources makes it difficult for young children to infer a speaker's need for information.

Considering that children only question a knowledge claim around the age of 4 to 5 years (but not earlier) is of particular interest if we consider the interpretation that children are only able to create an association between 'know' and sources of knowledge later on. If the mental state verb is indeed only used for identifying ignorance, but does not enable children to conclude that information is needed to circumvent the ignorance, it is very well possible that performance on the *know* condition only develops around the same age that children question a knowledge claim (Fedra & Schmidt, 2019), report another person's knowledge state (Gonzales et al., 2018) or even differentiate between factive verbs (Kristen-Antonow et al., 2019).

The traditional view supported by C. N. Johnson and Wellman (1980) is that young children learn to distinguish mental state verbs relatively late, namely around the age of 4 to 5 years. Harris, Ronfard, and Bartz (2017); Harris, Yang, and Cui (2017) have argued that children at the age of 2 already use the mental state verb 'know' appropriately and should therefore already possess a rudimentary understanding of knowledge and be able represent whether a speaker knows something or not. The results of the current study speak in favor of the traditional view, showing that young children struggle to infer a person's knowledge state from their use of the verb 'know' alone. Even though children were able to make inferences about knowledge from non-linguistic cues (Behne et al., 2012), it may be the case that learning inferences from a mental state verb is much less intuitive. Two- and 3-year-old children may have learned the word 'know' and use it in appropriately in some contexts as suggested by Harris, Ronfard, and Bartz (2017) and partially shown in study 1, but nevertheless fail to grasp how it relates to sources of knowledge and may indicate the need for information, which would enable them to draw appropriate inferences. The in-depth understanding of 'know' that includes the association between knowledge and a knowledge source, might not be necessary for children to use the verb in their daily life. Thus, children's understanding at 2 and 3 years of age seems limited even if they can display competence in naturalistic contexts with additional cues.

Since 2-year-old children ask information-seeking questions, one may argue that they are aware of sources of knowledge, at least to the degree that it can be used as a way to bridge their own ignorance. However, it is not possible to determine how much of children's asking is due to a behavioral strategy (similarly to what has been proposed by Rubio-Fernández (2019)) and how much is awareness that asking helps them overcome their ignorance. Future studies are necessary to determine the degree to which children associate knowledge with sources of that knowledge. For example it may be possible to prompt children to ask information-seeking questions by providing them with incomplete information. The same could be repeated for the knowledge of a third person to determine whether children have more difficulties with determining the need for information for another person than they do for themselves.

Another important point mentioned in the objectives of study 2 is that children's preverbal competence appears to be rather limited and strongly supported by cues and absence of such cues makes it difficult for children to solve tasks. If it is indeed mostly cues that aid children in drawing inferences, then it seems much more likely that their preverbal understanding is based on production rules and behavioral observation instead of preverbal concepts of mental states. The results of study 2 add to this observation as even when one investigates children's performance at a later age, only 5-year-olds are able to draw the right inferences and even then their performance is far from ceiling, indicating that the associations are much more difficult than previously assumed. In the case of a preverbal concept of mental states, associations to mental state verbs should be drawn much more easily than if each cue and behavior has to be internalized separately in the form of production rules. Even though one cannot conclude that children build production rules and develop conceptual understanding only later, the results of study 2 make it apparent that showing understanding by drawing the right inferences only develops around the age of 5 and leaves much room for improvement.

An important limitation of the current study is that it has only investigated children's ability to draw pragmatic inferences in a single, very limited context, namely in a hiding game. Therefore, it is possible that in other situations, 2- and 3-year-olds would be more competent in showing their comprehension of mental state verbs and their implications. In the future it would be important to investigate a variety of situation in which children use mental state verbs according to Harris, Ronfard, and Bartz (2017) and use well-controlled experiments to determine whether their comprehension of the verbs differs in these situations.

Overall, while mental state verbs appear to be used mostly appropriately by children in the third year of life in everyday contexts, the present findings indicate that most 2-year-olds, 3-year-olds, and even some 5-year-olds fail to understand the utterance "I do not know x" as a request for information about x, when they cannot rely on situational cues. One possibility is that children have not yet learned to associate mental state terms with sources of knowledge and are therefore not able to infer a speaker's need for information from a statement of ignorance alone. When comparing the present findings with those in naturalistic contexts, it seems feasible that children's performance on mental state verbs is overestimated as some of the pragmatic implications that they are able to draw preverbally, are not possible when presented with the mental state verb 'know' alone. The present findings converge with earlier research indicating that there is protracted development of the mastery of mental state verbs, with basic distinctions emerging only around the age of 4 to 5 years (C. N. Johnson & Maratsos, 1977; C. N. Johnson & Wellman, 1980). Still, it is important to investigate children's performance in other scenarios to determine whether the overestimation of mental state verb comprehension is present over various contexts.

5. GENERAL DISCUSSION

The ability to assign mental states to ourselves and the people around us - known as *Theory of Mind* (ToM) (Premack & Woodruff, 1978) - is crucial for social interaction and later positive relationships in life (Fink et al., 2015; Slaughter et al., 2015). Thus, it is crucial to investigate when exactly children acquire this ability and what is necessary for its development. Research over the past few decades showed that children seem to posses an early grasp of what an agent knows or does not know in preverbal interaction (Moll et al., 2007; O'Neill, 1996), whereas traditional, explicit tasks for ToM are only mastered around the age of 4 (Wellman & Liu, 2004; Wimmer & Perner, 1983). Children's early grasp of knowledge states can be argued to indicate an early understanding of mental states (MSU) that is not yet associated with a means to express these mental states explicitly. One possibility is that an additional representational resource is necessary in order to make preverbal understanding explicit and thereby enable children to master explicit ToM tasks.

One such representational resource that was proposed in this thesis is mental state language (MSL), the acquisition of terms that denote mental states i.e. 'to see', 'to know', 'to want', but also emotional and physiological terms. The present thesis aimed to investigate the role that MSL plays for MSU, the impact it has on the transition between preverbal MSU and subsequent mastery of explicit ToM tasks and finally which abilities can be considered predictors for the development of MSL production. In particular, the questions asked were 1) Is the production of mental state terms a sign of children's MSU? 2) What is the role of MSL production for the transition between preverbal MSU and the mastery of explicit ToM tasks? and 3) Which language and cognitive abilities predict MSL production early in the third year?

In order to answer the first question, it was crucial to determine whether children already use MSL in naturalistic contexts as has been reported in earlier studies and to compare children's naturalistic usage with different experimental tasks that assess whether children use the terms appropriately and show comprehension of some of the inferences that can be drawn from these terms. One way to do so was by concurrently comparing children's production of mental state terms to their ability to use the same terms to denote their own ignorance (which can be considered an appropriate context), to draw the correct pragmatic inferences from mental state terms (which indicates in-depth comprehension) and finally to use epistemic terms to report their own and somebody else's knowledge state (an additional sign for appropriate context). In case of a relationship between children's production of mental state terms and their performance on the above mentioned tasks, it can be concluded that children's ability to use mental state terms in naturalistic contexts is related to their comprehension of these terms and might therefore speak for an association between the mental state verb and the mental state that it denotes early in development. Such an association is crucial as it indicates continuity from children's early preverbal sensitivity for mental states and their subsequent comprehension of mental state terms during production.

In order to answer the second question, it was important to investigate the relationship between MSL production and both preverbal and verbal abilities that develop around the third year of life. Identifying a relationship between preverbal and verbal abilities enables us to conclude the presence of a continuous transition from early sensitivity towards later explicit mastery on the one hand, while on the other hand examining how MSL production relates to preverbal and verbal abilities gives us insight into whether children's preverbal MSU relates to children acquiring MSL or whether only MSL relates to verbal MSU. In particular, examining the relations helps to understand whether children possess conceptual understanding preverbally and use their concepts for MSL (ST), or whether encoding of behaviors in the form of production rules is a more reasonable explanation for preverbal competence (TT). Without investigating directionality, these questions could be examined concurrently and required tasks in which children show beginning preverbal and verbal competence. Despite FBU being the gold standard for MSU, children do not show consisent explicit FBU competence in the third year of life. Therefore, only implicit FBU was included and level-1 VPT was chosen as a simpler model. VPT had the advantage that children were reported to show beginning competence in the third year of life for both preverbal and verbal tasks while also using mental state language and requiring a differentiation between self and other as is needed for FBU.

Finally, in order to answer the third question, it was necessary to examine the development of MSL production longitudinally by measuring possible predictors at an earlier time point and determining the role they play for children's subsequent production of mental state terms.

To address the above mentioned questions, two studies were conducted. In the first longitudinal study, children were assessed at two time-points, once at the age of 24 months with measures of general language, cognitive and motoric development and inhibition skills, and an additional time at 27 months in order to assess children's MSL production, their preverbal and verbal perspective-taking skills, meta-cognitive awareness of ignorance, implicit FBU and finally a task that measured their ability to infer a speaker's need for information from a statement about knowing or not knowing the location of an item or wanting or not wanting to receive a particular item. The goal was to determine children's performance on these tasks as prior literature on 2-year-olds remains inconsistent about the level of MSU that children are capable of and its relation to MSL. To address these inconsistencies several aspects were included. First, an implicit FB task was added as implicit FBU has been drawn into question recently and requires additional research (Kulke, von Duhn, et al., 2018). Furthermore, the parental report measure for MSL production that is most commonly used in studies on children's MSL was extended by a scale that determines the contexts in which children use the epistemic verb 'to know'. This was done in order to have a measure of appropriate contexts that goes beyond the mere production of the mental state verbs in naturalistic situations. These additional aspects were used for the three main aims, namely (1) to determine the concurrent relationship of children's MSL production and their ability to comprehend and use these terms on experimental tasks, (2) to identify the concurrent relationship between MSL and preverbal and verbal perspective-taking and implicit FBU and finally (3) to determine which developmental abilities at the age of 24 months are significant predictors of children's subsequent MSL production. As children did not demonstrate consistent relations between MSL and MSU at the age of 27-months, a second cross-sectional study was conducted. In study 2 children's ability to infer a speaker's need for information was investigated further by administering the same task to 2-, 3- and 5-year-olds and determining at which age children show competence. The second study was crucial in order to investigate further at which point in time children are able to create associations between a mental state verb and the respective inferences that can be drawn from it, if they are not able to do so at the age of 2.

The results of study 1 showed that 27-month-olds were already able to produce mental state terms and showed nonverbal understanding of their own knowledge and ignorance, but failed to show in-depth comprehension of mental state verbs on experimental tasks. Due to floor performance, the *other* condition of the verbal VPT task, the EPT task and the *know* condition of the know-want task were not considered for correlations with other tasks and will only be discussed descriptively.

Children's performance on the MSLQ, the verbal visual and epistemic perspective-taking tasks and on the MCI task was in line with prior findings (Bartz, 2017; Gonzales et al., 2018; Kristen et al., 2011). Also the additionally employed scale of the MSL production questionnaire was in line with data from analyses of spontaneous speech. In case of action-based perspective-taking, children were competent but performed lower than in the original study with 24-month-olds. Children's increasing mobility at the age of 27 months was proposed as a possible explanation for the difference. Opposed to findings with 3-year-olds with the same implicit FB task (Grosse Wiesmann et al., 2017) and even 2-year-olds with other tasks (Southgate et al., 2007), 27-month-olds did not perform above chance on the implicit FB measure. The pattern of above chance performance on the FB1 condition and below chance performance on the FB2 condition was argued to be a sign for children's preference for the last location of the chasee, thus indicating a simple strategy instead of implicit FB competence. In the know-want task, young children were able to draw the right pragmatic inferences for something a speaker wants or does not want, but failed to do so for a speaker's need for information in relation to what she knows or does not know. The performance in the *want* condition was in line with prior research on children's understanding of desires and desire terms developing early on (Bartsch & Wellman, 1995; Wellman et al., 2001). Performance on the *know* condition was inconsistent if one assumes children to be able to draw the same inferences verbally that they already draw preverbally (O'Neill, 1996), but consistent if one expects children to be competent at the same age as they are able to differentiate between mental state verbs experimentally (Kristen-Antonow et al., 2019; Moore et al., 1989). However, as judgements on factivity require comparisons between verbs that are more complex than comparing the asertion and negation of a verb, it is surprising that 27-month-olds were not able to solve the task. Thus, despite early competence in producing a variety of mental state terms, children's comprehension of epistemic terms appeared to be limited.

Concerning the main aims of this thesis, the results can be summarized as follows: (1) Children's MSL production in the knowledge scale correlated with their ability to denote their own ignorance by using 'I don't know' when confronted with an unfamiliar picture. However, there was no relation between children's nonverbal signs of ignorance and their MSL production. Furthermore, children's MSL production on the knowledge scale was related to their ability to hand a speaker the object she wanted instead of the object she did not want, even though the knowledge scale was concerned with the epistemic verb 'to know'. Thus, using MSL in various contexts according to the knowledge scale was more important than the mere production of MSL. The results of study 2 showed that despite being able to draw the right pragmatic inferences for the mental state verb 'want' at the age of 2, only at the age of 5 children were able to draw the right pragmatic inferences for the mental state verb 'know/don't know'.

(2) Children's own verbal level-1 VPT was related to MSL production, in particular to the knowledge scale, independent of general language. The action-based (preverbal) perspective-taking task was related to children's ability to infer a speaker's desire from the mental state verb 'want' and to children's ability to report their own visual perspective in the verbal perspective-taking task, indicating a relationship between appropriate use of MSL and a preverbal inference about a speaker's desire. However, there was no relationship between action-based (preverbal) level-1 perspective-taking and MSL production according to the parental questionnaire, neither for the general scales, nor for the specific knowledge scale, suggesting that mere production in naturalistic contexts is not sufficient for the relationship. Finally, (3) the results of study 1 identified sentence production, inhibition and fine motor skills at 24 months as major contributors for children's later production of MSL at 27 months.

In the following, I will discuss in detail how both studies together provide information on the research questions in the broader context of this thesis. Findings that are not related to the three main aims of the thesis were discussed in 3.4 and will not be addressed in this section.

5.1 Is the production of mental state terms a sign of children's comprehension of the semantics and pragmatics of these terms and thereby MSU?

Harris, Ronfard, and Bartz (2017); Harris, Yang, and Cui (2017) argued that by 2 years of age children show some basic understanding of ignorance and that in spontaneous speech data children show many more references towards mental states than previous research suggested. One conclusion could be that children who use mental state terms appropriately have already established an association between their preverbal grasp of ignorance and knowledge (Harris, Ronfard, & Bartz, 2017; O'Neill, 1996) and the respective mental state terms that can be used to denote it i.e. 'to know'. Study 1 was able to confirm that 2-year-old children show a variety of signs of uncertainty and appear to have basic sensitivity towards their own ignorance, however, only few of these children were using a mental state term to denote that ignorance. A majority of children were able to produce mental state terms in naturalistic situations according to their parents, but did not use these terms in order to admit to not knowing the name of an object when confronted with their own ignorance in a task.

There can be several reasons for children not using a mental state term in order to talk about their own knowledge, even if they are capable of producing it. Therefore, it is important to investigate children's comprehension in experimental tasks that require an appropriate response. In the EPT task used in study 1, children were not able to report on their own knowledge even though some of them already showed competence in using the verb 'know' to denote their own ignorance, however, as the task required children to make associations between seeing and knowing and to use the epistemic verb 'know' to describe the association, low performance is not surprising. The combined results on the know-want task in study 1 and 2, show that even when the response can be non-verbal (pointing at the right box) 2- and 3-year-olds were not able to infer a speaker's need for information from the mental state verb 'know' while even 2-year-olds were able to make the right pragmatic inference for 'want'. Thus, the evidence provided in this thesis suggests that even though children show early sensitivity towards knowledge states and are able to produce mental state terms in naturalistic contexts, the mere production of these terms does not entail that they are able to pass experimental tasks that test for their comprehension of mental state terms.

An important question is why children appear to be competent when looking at spontaneous speech data and preverbal communication, but experimental set ups show much lower performance. One crucial difference between naturalistic situations and experimentally controlled designs is the presence of situational and social cues. In preverbal communication studies, children are always provided with visual input on locations of objects and experience of what a caregiver has or has not seen by having them either be present or absent while a toy is hidden. In some cases caregivers were even explicitly asked to cover their eyes and ears in order to signal that they were not able to observe (O'Neill, 1996). Furthermore, children could be using information on what an adult has seen without having any elaborate comprehension of what the verb 'seeing' actually entails. Rubio-Fernández (2019) would argue that children do not need to understand mental state verbs in order to be able to act appropriately as most mental state verbs are used in specific contexts by caregivers and can therefore be imitated without comprehension. Since children begin to hear and to use mental state verbs in rich situational contexts, they do not have to rely on the verb meaning alone in order to infer meaning. Also in the MCI task children have less social cues than they might have in naturalistic contexts, simply because parents respond to their children's nonverbal signs of ignorance and might propose a solution by asking 'you don't know?'. Bartsch and Wellman (1995) have also argued that caregivers provide explanations and additional feedback in naturalistic situations and therefore influence children's responses in speech data, even if children's utterances are spontaneous. The know-want task can thus be considered a difficult test for children's comprehension of 'know' as it does not leave children the possibility to acquire additional cues. There is no visual demonstration of the speaker's knowledge by observing objects being hidden and there is no additional hint for one of the objects as the speaker was absent during the hiding of both. Thus, children had to infer the speaker's need for information from the mental state verb alone and for as long as children are not able to do so, they fail the task. This suggests that one of the main reasons for the difference between children's appropriate usage of terms in spontaneous situations and their inability to do the same on experimental tasks, may be situational cues that can be used to reply accordingly without necessarily having an association between a mental state and a mental state verb that can be used to denote it.

If situational cues are what constitutes children's preverbal comprehension of mental states, the theoretical question whether children possess concepts preverbally (as suggested by the ST) or merely create production rules (as suggested by the TT at this age) gains an additional argument for the second interpretation. As I have mentioned in earlier sections, possessing preverbal concepts suggests that as soon as children acquire mental state verbs that can be used to denote these concepts, the associations should be straight-forward. A single denotation of a concept should make it more likely for a whole concept to be activated, than just the activation of a small part of the concept as would be expected with production rules (see Figure 5.1). If children's preverbal MSU is only built on situational cues from which they conclude behaviors, then it seems feasible for children to use these situational cues in order to create behavioral rules than to develop a functional concept. Children's performance on more complex tasks in the current thesis suggests that without such cues, an association between mental state verbs and pragmatic inferences are not possible. Thus, production rules appear to be the more reasonable way in which children show preverbal MSU, namely by using situational cues to develop production rules which aid them with their behavioral predictions.

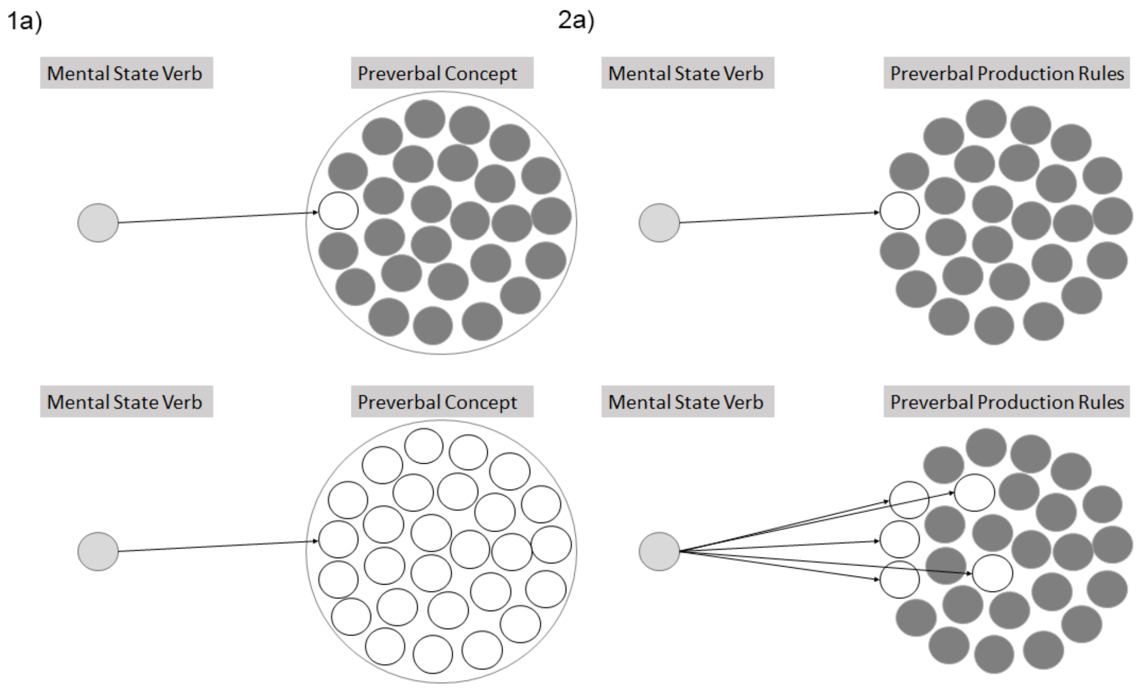

Fig. 5.1: Predictions for associations between mental state verbs and MSU for 1) a conceptual view in which a mental state verb is associated with one aspect of the concept, which results in an association to the whole concept and 2) a production rule based view in which a mental state verb requires to be associated with each production rule separately.

Corresponding to the presented view of the TT, Roberts (2016) has argued that children do not have a metarepresentational idea of 'mind' when they use expressions such as 'I know' or 'I think'. The author proposes that children only copy the expressions in a 'practical way' to perform illocutionary acts (bring about a behavior) without the ability to grasp the adult locutionary act (utterance and all its meanings) that hides behind it. Performing the illocutionary act is learned according to Wittgenstein's language games (cited in Roberts, 2016), namely by using 'I know' to bring about a behavioral effect in the real world. The idea is that children do not grasp the locutionary nature of the used mental state verbs but only use them in new situations to achieve something, for example not to be reprimanded when asked for an object they have hidden. Replying that they "don't know" the location could help to solve the conflict. Roberts (2016) called this process *practical abduction*. In practical abduction children acquire a verb through language games from which they are able to create practical rules to use the verb in new situations. However, the meaning of the verb remains the same as it had been during the language game in which the children acquired it. Since the meaning of a verb can vary dramatically, learning through practical abduction is not sufficient as the child will not be able to apply the rule it created to new situations. Through experiencing corrections and seeing other uses of the verb, children eventually learn various ways to use the verb and bring about an effect. However, the concept that is actually denoted by the mental state verb remains implicit and is never explicitly mentioned despite being necessary for MSU. Only through a second step that Roberts (2016) calls *practical-made-explicit abduction* children begin to construct hypotheses about explicit knowledge without the presence of practical knowledge. In this second step, children begin to use verbs in new situations with certain expectations without having practical rules to rely on. Overall, Roberts (2016) argues that children acquire practical rules for using mental state verbs first and only later begin to understand the appropriate implications and inferences that can be drawn from them. This view is in accordance with the TT which suggests that children merely create representations and only later on, around the age of 4 they are able to create metarepresentations and therefore acquire a conceptual understanding of MSU. The finding of the first and second study, namely that children at age of 2 and 3 do not differentiate between 'know' and 'not know' in a pragmatic task, but that 5-year-olds show beginning competence, is in line with this interpretation. Still, considering that even the 5-year-olds were not at ceiling, it is difficult to conclude that a conceptual understanding of mental state verbs has fully developed by then.

Considering that 2-year-olds fail to show comprehension of mental state terms when tested on experimental paradigms, the question remains what exactly is necessary for that competence to develop and why it only begins to show around the age of 5, similarly to children's ability to report on their own and other's knowledge states (Gonzales et al., 2018), their comprehension of factivity (Kristen-Antonow et al., 2019) and their ability to question a knowledge claim (Fedra & Schmidt, 2019). One possibility that can be inferred from Roberts (2016) is the need for additional experience with alternative uses of mental state verbs that go beyond what children have experienced during acquisition. Parents need to provide children with various contexts in which a mental state verb can be used and correct inappropriate uses to ensure that children learn to abstract from a particular context to the denotation of a mental state. The correlations with the knowledge scale of the MSLQ which assessed the number of situations in which children used the mental state verb 'know/don't know' is additional support towards this interpretation. Also Harris et al. (2005) have argued that children acquire an understanding of mental states through conversation and through caregivers emphasizing different perspectives. Since the perspectives of others are usually described in third person, it also seems likely that third-person input results in better acquisition of mental state terms (Gola, 2012). Maybe lack of third person input is also the reason for 4-year-olds struggling with changing a mental state verb from first person to third person (Brandt et al., 2011). Also, work on autism is in line with this interpretation as autistic adults appear to lack personal narratives despite similar mental state vocabulary, which appears to be due to a lack of qualitative conversational input (Bang et al., 2013). Thus, one of the reasons for children's late comprehension of mental state terms in experimental tasks and possible a later conceptual development of MSU may be due to lack of experience with various contexts in which these mental states can be used and minimal input with third person utterances about others' mental states.

Despite recent research suggesting that children may have an early understanding of mental state terms and what they entail, the results of the current thesis support the traditional view proposed by C. N. Johnson and Wellman (1980), namely that children learn to distinguish between mental state verbs relatively late, also when the distinction only applies to 'know' and 'not know'. I have proposed that experiencing mental state terms in various contexts may be what drives children's comprehension and that at the age of 2, most children have not been able to experience mental state terms in sufficiently many contexts to be able to generalize from situations to the denotation of a particular mental state. Most of their competence stems from situational cues and rules that they have developed preverbally. However, it is important to note that despite lacking the ability to infer a speaker's need for information and to respond to questions about their own knowledge state, 2-year-olds show beginning understanding of their own ignorance and the ability to use mental state verbs to denote it. Thus, it is possible that understanding their own ignorance is a first step in the comprehension of mental state terms and only later children acquire an understanding for sources of knowledge and that providing information is a way to solve one's own and someone else's state of ignorance. Through experience with mental state verbs, children acquire associations between the rules they have created preverbally and mental state verbs that can be used to express them. Additionally, 27-month-olds seem to be able to infer another person's desire not only in the know-want task, but also during the preverbal perspective-taking task. This implies that associating the verb 'want' with a mental representation is already possible earlier than comprehension of epistemic terms like 'know' and that desire terms may need less conversational input in order for children to be able to use them. Possibly, children's experience with the mental state verb 'want' is richer and might therefore result in earlier development of conceptual understanding.

It is important to keep in mind that the current thesis has only investigated children's comprehension of mental state verbs in several limited scenarios, namely reporting their own knowledge, showing awareness of their own ignorance and inferring a speaker's need for information from statements about her ignorance. Based on the spontaneous speech data by Harris, Yang, and Cui (2017), it is possible to identify additional scenarios in which children can demonstrate their comprehension of mental state terms. Five such scenarios are: a) associating seeing something with knowing it, b) associating being told something with knowing it, c) contrasting their own knowledge with someone else's, d) denying someone's knowledge claim and e) differentiating between previous and current stages of knowledge (Jarvers, Kristen-Antonow, & Sodian, 2017). All of these scenarios indicate an aspect of mental state verb comprehension and the current study has only investigated an association between lack of knowledge and a need for information. Possibly the scenarios mentioned above are simpler for young children and stages of comprehension can be identified in future research. It would be crucial to determine whether competence in all five scenarios enables children to generalize from the specific situations to an overall concept of knowledge.

In summary, the question dealt with the degree to which appropriate use of mental state terms in naturalistic contexts reflects children's comprehension of these terms and the associated mental states. The combined results of study 1 and 2 suggest that despite early usage of epistemic mental state terms in naturalistic contexts, 27-month-olds are not yet able to comprehend detailed implications of these terms and only few appear to associate the mental state term with the respective mental state. Since only 5-year-olds showed incipient competence at inferring a speaker's ignorance and providing information, one may argue that elaborate comprehension of epistemic mental state terms only develops much later than recent research suggests and closer to the traditionally proposed age of 4-5 (C. N. Johnson & Wellman, 1980). One possibility is that children require additional conversational input on others' mental states and various contexts in which mental state terms can be used, in order to separate the mental state verb from a particular situation and generalize it to the mental state it denotes. It is important to note that children already appear to have associations for mental state verbs like ‘want’, but associations for epistemic verbs like ‘know’ seem to develop much later.

5.2 What is the role of MSL production for the transition between preverbal MSU and the mastery of explicit ToM tasks?

One role that MSL could play in the transition from early preverbal MSU to subsequent mastery of explicit tasks is by making preverbal representations of mental states verbal and thereby enabling children to report on them in explicit tasks. I have mentioned previously that theoretical approaches differ in whether they would classify preverbal MSU as conceptual (ST) or rather as an accumulation of production rules (TT). Investigating the role that MSL plays for preverbal and verbal competence is another way to shed light on the discrepancy. In the present thesis, this relationship was investigated concurrently by examining MSL production and how it relates to verbal and preverbal (action-based) level-1 VPT. One of the reasons for choosing perspective-taking was children’s early competence in preverbal measures (Masangkay et al., 1974; Moll & Tomasello, 2006) and beginning competence in verbal measures (Gonzales et al., 2018). Thus, both measures together make it possible to determine the relationship between MSL and both preverbal and verbal VPT. VPT was used as a simplified model instead of FBU, as children do not show reliable performance on implicit and explicit FBU at the age of 2. The idea was that preverbal VPT may establish a conceptual base for children’s acquisition of mental state terms and subsequently mental state terms enable children to report on perspectives in a verbal task. By examining the relationship concurrently, it was possible to determine how the two perspective-taking tasks relate to each other and whether they measure the same concept in addition to determining whether MSL is important for both preverbal and verbal VPT at the age of 27 months. Preverbal VPT was related to children’s own verbal VPT and thus confirmed that both tasks were measuring a similar underlying concept independent of children’s general language abilities. As reported by Chiarella et al. (2013) in 30-month-olds, there was a significant relationship between children's own verbal perspective-taking and their MSL independent of general language. This relationship applied to children's usage of the epistemic verb 'know/don't know' in several naturalistic contexts according to their parents. This suggests that children's ability to report their own perspective when asked in a verbal VPT task is related to the amount of contexts in which they are able to use the epistemic verb 'know/don't know'. This is interesting as it refers back to the previous section in which I discussed the importance of conversational input for children's comprehension of mental state terms. It is possible that children who have experienced several scenarios in which they used the epistemic verb 'to know', were also children who received sufficient conversational input to relate their own visual perception to a mental state verb that can be used to denote it. Unfortunately, the nature of correlation does not allow for inferences about causation and therefore it is not possible to determine whether (1) naturalistic usage of epistemic verbs in several contexts results in children reporting their own visual perspective better, (2) whether reporting their own visual perspective well results in children being able to use epistemic terms in various naturalistic contexts or (3) whether both abilities were influenced by a third causal factor. The direction of this relationship would have to be examined further in future longitudinal work or in a training study that investigates the impact of conversational input on children's verbal perspective-taking abilities.

From the results it can be concluded that there is continuity between verbal perspective-taking and MSL production. However, when examining how MSL production (measured by the parental questionnaire) relates to the preverbal perspective-taking task there was no relationship. This suggests that children's preverbal representations of perspectives are not related to their ability to produce mental state verbs, independent of category. However, the relation between the two perspective-taking tasks and the additional correlation with the ability to infer desires from the mental state verb 'want', implies that an association between early representations and mental state verbs is present in specific cases, in the current sample for the mental state verb 'want'. However, as drawing inferences for 'want' was related to producing 'know' in several naturalistic contexts, it is plausible that this association is only present for mental state terms that were experienced in a variety of contexts. If Roberts (2016) is right about children acquiring mental state terms in particular contexts without insight into the deeper meaning behind these terms, it might explain why there is no relationship between mere MSL production and children's nonverbal sensitivity for perspectives. Children acquire these terms to using them in particular contexts and bring about an effect (e.g. divert attention towards an object) but these contexts are not necessarily related to implicit representations of perspectives. Rather, children have production rules which they create based on the behavior they observe and associate with a mental state verb the more commonly it is used in their presence. However, as mentioned above, in cases where children show comprehension of MSU (e.g. the verbal perspective-taking task) there seems to be continuity from children's preverbal representations of mental states to their denotation of these mental states through MSL. Thus, there appears to be a relationship between preverbal and verbal VPT and MSL seems to be important, but the exact role that MSL plays for the transition and whether experience may aid the transition remains to be investigated.

An important aspect is that children were competent in the preverbal VPT task, but struggled with the verbal one. This implies that similarly to tasks about knowledge, children were able to show implicit competence, but commonly failed as soon as they were supposed to infer how to answer correctly based on a mental state verb, in this case 'to see'. According to Gonzales et al. (2018), children were competent at reporting their own perspective and those of others around 3 years of age which is older than the current sample and much later than children show competence in the preverbal task (Moll & Tomasello, 2006). One may argue mental state verbs like 'seeing' are easier to acquire as situations in which they apply are simpler to identify, but still children need to have a sufficient amount of experience with and conversational input involving these terms to be able to use them in experimental scenarios. Therefore, some children show competence for the verb 'seeing' with their own perspective before they are able to do the same with the perspective of others, even though they seem to draw inferences preverbally for both, their own perspective and that of others. If children's preverbal competence is present in the form of production rules, it is very well possible that they had the time and experience to associate rules for their own perspesctive with a mental state verb, but not yet rules about the perspective of another person. Furthermore, if preverbal competence is based on cues (i.e. objects being in front of or behind a bucket in addition to searching behavior) then children would not need to develop conceptual understanding to behave appropriately in their social interactions.

Thus, on the one hand, children are able to develop an early sensitivity for perspectives, but only associate this sensitivity with mental state terms in a scenario in which they report their own perspective. On the other hand, their ability to explicitly report on their own perspective is related to the naturalistic usage of epistemic terms in various contexts. There are several ways in which MSL might connect early preverbal VPT and subsequent verbal VPT. One possibility is that early preverbal competence in VPT enables the development of mental state vocabulary as conversations about what is visible and what is not motivate caregivers to use terms for representations. Another possibility is that children develop early sensitivity for perspectives and acquire mental state terms independently, but that through conversational input and discourse about perspectives (Harris et al., 2005; Lohmann & Tomasello, 2003) they learn to associate these terms with their early sensitivity for perspectives and thus learn to use the terms in verbal tasks. Future research would have to investigate both options further and determine whether it is indeed conversational input and discourse about mental states that enables children to associate preverbal abilities with subsequent verbal mastery.

Overall, the question was what role MSL plays for the transition between early preverbal abilities and later explicit abilities. The current thesis aimed at answering this question by examining preverbal and verbal VPT in order to determine whether MSL plays a role for both and might therefore be considered important for the transition. In summary, there was evidence for continuity between preverbal and verbal perspective-taking with preverbal perspective-taking relating to children's ability to report their own perspective through the mental state verb 'see', but MSL production was only directly related to the verbal perspective-taking task, but not the preverbal perspective-taking task. It was hypothesized that children might not associate their early sensitivity for perspectives with the mere production of mental state verbs yet as relations were only present for cases in which children showed use of mental state terms in various contexts. Furthermore, the theoretical view that children's preverbal competence is build on production rules instead of conceptual understanding appears more likely given the specificity of the relationships found. Future studies would have to investigate whether preverbal perspective-taking relates to MSL production in cases where children have more situational experience and whether one might even fuel the transition from preverbal to verbal by providing conversational input.

5.3 What are the basic linguistic and cognitive abilities that contribute to the development of mental state language?

Since MSL production appears to play a role as a representational resource, it is of major importance to identify which abilities may be important for its emergence. So far there are no studies that have investigated the relevant predictors of later MSL production longitudinally. The current thesis identified sentence production, inhibitory control and fine motor skills at 24 months to be predictive of children's MSL production at 27 months. Sentence production has been argued to be relevant as children may need to speak at least 2-word sentences before they are able to acquire mental state terms and talk about what another person feels, wants or knows. Mental states are always 'about' something and this aspect of aboutness (Astington, 1993) may require 2-word sentences in order for children to be able to express it. However, as sentence production mattered for all types of mental state verbs, including emotional and physiological terms, it is also possible that caregivers are more likely to use mental state terms with children that produce more complex linguistic output.

Inhibitory control was identified in an earlier concurrent study (Bellagamba et al., 2014) and confirmed in the current longitudinal design. The idea is that inhibitory control has a similar role for MSL as it does for ToM, namely to help children differentiate between their own mental states and those of others in addition to between mental states and reality.

Finally, fine motor skills have been identified as a crucial predictor for children's later MSL production. Iverson (2010) has argued that children's developing motor skills give them more opportunities to learn words by manipulating objects and have caregivers react to these manipulations by naming them. Also verbs are most commonly learned by observing an action or performing it oneself, thus, it is possible that mothers are more likely to use the word 'want' with a child who grasps towards an object, than a child who does not show such motions.

In the context of the previous questions, these predictors for MSL are particularly interesting as they once again emphasize the importance of conversational input. Good motor skills and sentence production enable children to be more interactive with their caregivers and thereby increase the linguistic input that they receive. The relevance of using mental state terms in various contexts, as has been identified in study 1, also hints at the importance of receiving conversational input and learning about mental state terms through explanations and feedback. In the future it would be of great interest to investigate which predictors remain important when maternal mental state talk is included as a predictor. If conversational input is one of the main contributors to both MSL production and the association between MSL terms and mental states, it would be crucial to support children by providing additional input and training their comprehension of mental state terms.

5.4 Limitations and Future Research

First of all, the current work was mostly correlational and therefore does not allow for conclusions about causality of the identified relationships. For this purpose a training design could be beneficial as it would enable us to relate performance changes to the intervention used. A future study could examine the effect of conversational input on children's performance on MSL comprehension tasks and thereby determine if it is indeed the additional context and discourse about mental states that helps children to associate their mental states with the respective verbs. An additional limitation is the concurrent nature of the study as predictive relationships could not be examined. In the case of implicit FBU an additional longitudinal study might be helpful to determine whether children's performance improves over time and if the preference for a 'last location' strategy reduces with age. A longitudinal study could also be used to investigate the impact of preverbal perspective-taking on MSL and whether acquisition is aided by children's early sensitivity. Furthermore, one may be able to investigate the importance of MSL as a mediator in the transition from preverbal perspective-taking to verbal perspective-taking.

An additional limitation is that this study has only investigated VPT, but there are also other abilities that develop around the same time and might have influenced the concurrent relationships with MSL production. For example precursors like joint attention have been shown to relate to MSL and later explicit ToM performance. Possibly, preverbal competence is only associated with mental state production for children that fulfill certain prerequisites. In particular Brooks and Meltzoff (2015) have shown that gaze-following predicted children's later MSL, which may be an indicator of preverbal awareness enabling children to acquire mental state terms. Moreover, as children already show competence in preverbal communication, it might have been beneficial to start out earlier and thereby investigate the predictive relationship between children's early preverbal competence and later usage of mental state terms.

The only measure in this study that assessed whether children show early comprehension of knowledge independent of mental state terms was the MCI task as children were required to indicate their own ignorance without using mental state terms. However, using one of the earlier tasks by O'Neill (1996) and Behne et al. (2012) may provide an additional measure for children's early sensitivity for others' knowledge states in addition to sensitivity for their own knowledge state. However, since 27-month-olds struggled to use mental state terms for their own knowledge, it is unlikely that they would be able to use them for another person, especially if conversational input is as crucial as has been argued above.

Another important limitation is that despite mother's mental state talk predicting both children's MSL production and their later ToM, it has not been included as a possible predictor of later MSL production. In a future study it would be important to identify the impact that mother's mental state talk has and whether fine motor skills and sentence production will still matter as soon as a measure for conversational input is included. Possibly sentence production and fine motor skills are ways to modulate how mothers talk to their children and by introducing interventions one is able to manipulate linguistic input.

The current thesis identified that naturalistic usage in various contexts matters, but there was no way to measure it directly with the exception of the knowledge scale. Furthermore, the knowledge scale of the MSLQ only included three scenarios and in the future it might be better to create a measure that includes several contexts for different mental state verbs. This would make it possible to determine the impact of mental state verb contexts on children's MSL comprehension in more detail (in particular differentiate for desire and epistemic terms).

Finally, the current study has only investigated children's comprehension of mental state terms in very limited contexts, namely in relation to an epistemic verb and more specifically how children are able to associate their own ignorance with it and infer another person's need for information from it. It would be important to consider broader aspects of comprehension in the future and to determine whether children might be able to show competence in other measures of comprehension. The following scenarios may be considered, namely a) associating seeing something with knowing it, b) associating being told something with knowing it, c) contrasting their own knowledge with someone else's, d) denying someone's knowledge claim and e) differentiating between previous and current stages of knowledge (Jarvers et al., 2017). It is possible that children are familiar with some of these situations early on and therefore show competence before they are able to do so in the comprehension tasks used in the current study.

Future studies might investigate children's competence in each of these scenarios in addition to examining how conversational input may improve their performance.

6. CONCLUSION

In summary, the current thesis investigated children's MSL production in the third year of life and identified that despite early usage of epistemic mental state terms in naturalistic contexts, 27-month-olds are not yet able to comprehend the detailed implications of these terms. Only at the age of 5 children were able to show competence in inferring a speaker's need for information, which supports the view that elaborate comprehension of mental state terms only develops later. Furthermore, this thesis showed that MSL production appears to be relevant for children's later reports of their own VPT, but shows no relation to children's nonverbal sensitivity for the perspectives of others. Only experimental tasks that required children to use mental state terms appropriately showed a relationship to preverbal perspective-taking, thereby suggesting that production alone is not enough for an association. Finally, a longitudinal design was used to identify sentence production, inhibitory control and fine motor skills at 24 months as relevant precursors of MSL production at 27 months. Overall, it was argued that the crucial aspect of MSL is not mere production of the words, but the number of contexts in which children are able to use these terms. The number of naturalistic contexts appeared to relate to children's mental representations of their own ignorance, to their ability to infer a speaker's desire and their ability to verbalize their own perspective, thus suggesting the conversational input that children receive as one of the main ways in which mental representations become explicit. The current work is the first attempt to identify important precursors of MSL production in a longitudinal design and is the first large sample study that investigated the relationship between MSL production and comprehension of mental state terms experimentally. Future research is needed to investigate the importance of conversational input further and to determine whether it is indeed the driving force for an association between early mental representations and mental state terms that can be used to denote them.

References

Amsterlaw, J., & Wellman, H. M. (2006). Theories of mind in transition: A microgenetic study of the development of false belief understanding. *Journal of Cognition and Development*, *7*(2), 139–172.

Apperly, I. A. (2010). *Mindreaders: the cognitive basis of "theory of mind"*. Psychology Press.

Apperly, I. A. (2012). What is "theory of mind"? concepts, cognitive processes and individual differences. *The Quarterly Journal of Experimental Psychology*, *65*(5), 825–839.

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, *116*(4), 953–970.

Arslan, B., Verbrugge, R., Taatgen, N., & Hollebrandse, B. (2018). Accelerating the development of second-order false belief reasoning: A training study with different feedback methods. *Child Development*, *91*(1), 249–270.

Aschersleben, G., Hofer, T., & Jovanovic, B. (2008). The link between infant attention to goal-directed action and later theory of mind abilities. *Developmental Science*, *11*(6), 862–868.

Astington, J. W. (1993). *The child's discovery of the mind* (Vol. 31). Harvard University Press.

Astington, J. W., & Gopnik, A. (1991). Theoretical explanations of children's understanding of the mind. *British Journal of Developmental Psychology*, *9*(1), 7–31.

Austin, K., Theakston, A., Lieven, E., & Tomasello, M. (2014). Young children's understanding of denial. *Developmental Psychology*, *50*(8), 2061–2070.

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

Bang, J., Burns, J., & Nadig, A. (2013). Brief report: conveying subjective experience in conversation: production of mental state terms and personal narratives in individuals with high functioning autism. *Journal of Autism and Developmental Disorders*, *43*(7), 1732–1740.

Baron-Cohen, S. (2010). Empathizing, systemizing, and the extreme male brain theory of autism. In *Progress in brain research* (Vol. 186, pp. 167–175). Elsevier.

Bartsch, K., & Wellman, H. M. (1995). *Children talk about the mind.* Oxford University Press.

Bartz, D. T. (2017). *Young children's meta-ignorance.* (Unpublished doctoral dissertation). Harvard Univeristy, Cambridge, MA.

Behne, T., Liszkowski, U., Carpenter, M., & Tomasello, M. (2012). Twelve-month-olds' comprehension and production of pointing. *British Journal of Developmental Psychology*, *30*(3), 359–375.

Bellagamba, F., Laghi, F., Lonigro, A., Pace, C. S., & Longobardi, E. (2014). Concurrent relations between inhibitory control, vocabulary and internal state language in 18-and 24-month-old italian-speaking infants. *European Journal of Developmental Psychology*, *11*(4), 420–432.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300.

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, *29*(4), 1165–1188.

Bigelow, A. E., & Dugas, K. (2009). Relations among preschool children's understanding of visual perspective taking, false belief, and lying. *Journal of Cognition and Development*, *9*(4), 411–433.

Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition*, *77*(1), B25–B31.

Brandone, A. C., Stout, W., & Moty, K. (2019). Triadic interactions support infants' emerging understanding of intentional actions. *Developmental Science*, *23*(2), e12880.

Brandt, S., Buttelmann, D., Lieven, E., & Tomasello, M. (2016). Children's understanding of first-and third-person perspectives in complement clauses and false-belief tasks. *Journal of Experimental Child Psychology*, *151*, 131–143.

Brandt, S., Verhagen, A., Lieven, E., & Tomasello, M. (2011). German children's productivity with simple transitive and complement-clause constructions: Testing the effects of frequency and variability. *Cognitive Linguistics*, *22*(2), 325-357.

Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology*, *18*(6), 906–921.

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.

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.

Cannon, E. N., & Woodward, A. L. (2012). Infants generate goal-based action predictions. *Developmental Science*, *15*(2), 292–298.

Carlson, S. M., Mandell, D. J., & Williams, L. (2004). Executive function and theory of mind: stability and prediction from ages 2 to 3. *Developmental Psychology*, *40*(6), 1105-1122.

Carpendale, J. I., & Lewis, C. (2004). Constructing an understanding of mind: The development of children's social understanding within social interaction. *Behavioral and Brain Sciences*, *27*(1), 79–96.

Chiarella, S. S., Kristen, S., Poulin-Dubois, D., & Sodian, B. (2013). Concurrent relations between perspective-taking skills, desire understanding, and internal-state vocabulary. *Journal of Cognition and Development*, *14*(3), 480–498.

Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, *9*(4), 377–395.

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

Dennett, D. C. (1989). *The intentional stance.* MIT Press.

De Villiers, J. (2007). The interface of language and theory of mind. *Lingua*, *117*(11), 1858–1878.

De Villiers, J. G., & De Villiers, P. A. (2000). Linguistic determinism and the understanding of false beliefs. *Children's Reasoning and the Mind*, *14*, 191–228.

De Villiers, J. G., & Pyers, J. E. (2002). Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cognitive Development*, *17*(1), 1037–1060.

Devine, R. T., & Hughes, C. (2014). Relations between false belief understanding and executive function in early childhood: A meta-analysis. *Child Development*, *85*(5), 1777–1794.

Devine, R. T., & Hughes, C. (2019). Let's talk: Parents' mental talk (not mind-mindedness or mindreading capacity) predicts children's false belief understanding. *Child Development*, *90*(4), 1236–1253.

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*, *46*, 12–30.

Dretske, F. (1990). Seeing, believing, and knowing. In *Visual cognition and action: An invitation to cognitive science, vol 2* (pp. 129–148). MIT Press.

Dudley, R., Orita, N., Hacquard, V., & Lidz, J. (2015). Three-year-olds' understanding of know and think. In *Experimental perspectives on presuppositions* (pp. 241–262). Springer.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Fedra, E., & Schmidt, M. F. (2019). Older (but not younger) preschoolers reject incorrect knowledge claims. *British Journal of Developmental Psychology*, *37*(1), 130–145.

Fenici, M. (2017). What is the role of experience in children's success in the false belief test: Maturation, facilitation, attunement or induction? *Mind & Language*, *32*(3), 308–337.

Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the macarthur communicative development inventories. *Applied Psycholinguistics*, *21*(1), 95–116.

Fink, E., Begeer, S., Peterson, C. C., Slaughter, V., & de Rosnay, M. (2015). Friendlessness and theory of mind: A prospective longitudinal study. *British Journal of Developmental Psychology*, *33*(1), 1–17.

Flavell, J. H. (1974). The development of inferences about others. In *Understanding other persons.* Rowman and Littlefield.

Flavell, J. H. (1999). Cognitive development: Children's knowledge about the mind. *Annual Review of Psychology*, *50*(1), 21–45.

Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the level 1–level 2 distinction. *Developmental Psychology*, *17*(1), 99–103.

Fodor, J. A. (1983). *The modularity of mind.* MIT ress.

Gagne, J. R., & Saudino, K. J. (2016). The development of inhibitory control in early childhood: A twin study from 2–3 years. *Developmental Psychology*, *52*(3), 391–399.

German, T. P., & Leslie, A. M. (2001). Children's inferences from 'knowing'to 'pretending'and 'believing'. *British Journal of Developmental Psychology*, *19*(1), 59–83.

Gola, A. A. H. (2012). Mental verb input for promoting children's theory of mind: A training study. *Cognitive Development*, *27*(1), 64–76.

Gonzales, C. R., Fabricius, W. V., & Kupfer, A. S. (2018). Introspection plays an early role in children's explicit theory of mind development. *Child Development*, *89*(5), 1545–1552.

Gopnik, A., & Meltzoff, A. N. (1998). *Words, thoughts, and theories.* Mit Press.

Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind & Language*, *7*(1-2), 145–171.

Gordon, R. M. (1986). Folk psychology as simulation. *Mind & Language*, *1*(2), 158–171.

Grazzani, I., & Ornaghi, V. (2012). How do use and comprehension of mental-state language relate to theory of mind in middle childhood? *Cognitive Development*, *27*(2), 99–111.

Grimm, H., & Aktaş, M. (2000). *Setk-2: Sprachentwicklungstest für zweijährige kinder: Diagnose rezeptiver und produktiver sprachverarbeitungsfähigkeiten: Manual.* Hogrefe, Verlag für Psychologie.

Grosse Wiesmann, C., Friederici, A. D., Disla, D., Steinbeis, N., & Singer, T. (2018). Longitudinal evidence for 4-year-olds' but not 2-and 3-year-olds' false belief-related action anticipation. *Cognitive Development*, *46*, 58–68.

Grosse Wiesmann, C., Friederici, A. D., Singer, T., & Steinbeis, N. (2017). Implicit and explicit false belief development in preschool children. *Developmental Science*, *20*(5), e12445.

Harris, P. L. (1992). From simulation to folk psychology: The case for development. *Mind & Language*, *7*(1-2), 120–144.

Harris, P. L., Bartz, D. T., & Rowe, M. L. (2017). Young children communicate their ignorance and ask questions. *Proceedings of the National Academy of Sciences*, *114*(30), 7884–7891.

Harris, P. L., de Rosnay, M., & Pons, F. (2005). Language and children's understanding of mental states. *Current Directions in Psychological Science*, *14*(2), 69–73.

Harris, P. L., German, T., & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition*, *61*(3), 233–259.

Harris, P. L., & Jones, D. P. (1997). Commentary on" talking about feelings"(aldridge & wood, 1997). *Child Abuse and Neglect*, *21*, 1217–1220.

Harris, P. L., Kavanaugh, R. D., Wellman, H. M., & Hickling, A. K. (1993). Young children's understanding of pretense. *Monographs of the Society for Research in Child Development*, i–107.

Harris, P. L., Ronfard, S., & Bartz, D. (2017). Young children's developing conception of knowledge and ignorance: Work in progress. *European Journal of Developmental Psychology*, *14*(2), 221–232.

Harris, P. L., Yang, B., & Cui, Y. (2017). 'i don't know': Children's early talk about knowledge. *Mind & Language*, *32*(3), 283–307.

He, Z., Bolz, M., & Baillargeon, R. (2012). 2.5-year-olds succeed at a verbal anticipatory-looking false-belief task. *British Journal of Developmental Psychology*, *30*(1), 14–29.

Heal, J. (1986). Replication and functionalism. In *Language, mind, and logic* (pp. 135–150). Cambridge University Press.

Helming, K. A., Strickland, B., & Jacob, P. (2014). Making sense of early false-belief understanding. *Trends in Cognitive Sciences*, *18*(4), 167–170.

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

Heron-Delaney, M., Wirth, S., & Pascalis, O. (2011). Infants' knowledge of their own species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1571), 1753–1763.

Heyes, C. (2014). False belief in infancy: a fresh look. *Developmental Science*, *17*(5), 647–659.

Iverson, J. M. (2010). Developing language in a developing body: The relationship between motor development and language development. *Journal of Child Language*, *37*(2), 229–261.

Jarvers, I., Kristen-Antonow, S., & Sodian, B. (2017). *Scaffolding talk about the mind in 33 month-olds in a laboratory-based setting.* (Poster presented at the Joint Conference of the Sections Developmental Psychology and Educational Psychology (PAEPSY), Münster, Germany)

Johnson, C. J. (1992). Cognitive components of naming in children: Effects of referential uncertainty and stimulus realism. *Journal of Experimental Child Psychology*, *53*(1), 24–44.

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

Johnson, C. N., & Wellman, H. M. (1980). Children's developing understanding of mental verbs: Remember, know, and guess. *Child Development*, 1095–1102.

Kammermeier, M., & Paulus, M. (2018). Do action-based tasks evidence false-belief understanding in young children? *Cognitive Development*, *46*, 31–39.

Kim, S., Paulus, M., Sodian, B., & Proust, J. (2016). Young children's sensitivity to their own ignorance in informing others. *PloS One*, *11*(3), e0152595.

Klann-Delius, G. (1998). *Manual for coding utterances of internal states.* Berlin, Germany: Free University of Berlin.

Knudsen, B., & Liszkowski, U. (2012). 18-month-olds predict specific action mistakes through attribution of false belief, not ignorance, and intervene accordingly. *Infancy*, *17*(6), 672–691.

Kovács, Á. M., & Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences*, *106*(16), 6556–6560.

Kristen, S., Sodian, B., Licata, M., Thoermer, C., & Poulin-Dubois, D. (2012). The development of internal state language during the third year of life: A longitudinal parent report study. *Infant and Child Development*, *21*(6), 634–645.

Kristen, S., Sodian, B., Thoermer, C., & Perst, H. (2011). Infants' joint attention skills predict toddlers' emerging mental state language. *Developmental Psychology*, *47*(5), 1207–1219.

Kristen, S., Thoermer, C., Hofer, T., Aschersleben, G., & Sodian, B. (2006). Skalierung von "theory of mind"-aufgaben. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *38*(4), 186–195.

Kristen-Antonow, S., Jarvers, I., & Sodian, B. (2019). Preschoolers' developing understanding of factivity in mental verb comprehension and its relation to first-and second-order false belief understanding: A longitudinal study. *Journal of Cognition and Development*, *20*(3), 354–369.

Kristen-Antonow, S., Licata-Dandel, M., Müller, M., & Sodian, B. (2018). Maternal cognition talk in the mother–toddler dyad mediates the influence of early maternal emotional availability on preschoolers' belief reasoning. *Social Development*, *27*(4), 841–857.

Kulke, L., Johannsen, J., & Rakoczy, H. (2019). Why can some implicit theory of mind tasks be replicated and others cannot? a test of mentalizing versus submentalizing accounts. *PloS One*, *14*(3), e0213772.

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

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

Lecce, S., Ceccato, I., Bianco, F., Rosi, A., Bottiroli, S., & Cavallini, E. (2017). Theory of mind and social relationships in older adults: the role of social motivation. *Aging & Mental Health*, *21*(3), 253–258.

Léger, C. (2007). The acquisition of two types of factive complements. In *Language acquisition and development: Proceedings of gala* (pp. 337–347). Cambridge Scholars Publishing.

Legerstee, M., Pomerleau, A., Malcuit, G., & Feider, H. (1987). The development of infants' responses to people and a doll: Implications for research in communication. *Infant Behavior and Development*, *10*(1), 81–95.

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in 'theory of mind'. *Trends in Cognitive Sciences*, *8*(12), 528–533.

Leslie, A. M., & Polizzi, P. (1998). Inhibitory processing in the false belief task: Two conjectures. *Developmental Science*, *1*(2), 247–253.

Liszkowski, U., Carpenter, M., & Tomasello, M. (2007). Pointing out new news, old news, and absent referents at 12 months of age. *Developmental Science*, *10*(2), F1–F7.

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

Lohmann, H., & Tomasello, M. (2003). The role of language in the development of false belief understanding: A training study. *Child Development*, *74*(4), 1130–1144.

Low, J. (2010). Preschoolers' implicit and explicit false-belief understanding: Relations with complex syntactical mastery. *Child Development*, *81*(2), 597–615.

Masangkay, Z. S., McCluskey, K. A., McIntyre, C. W., Sims-Knight, J., Vaughn, B. E., & Flavell, J. H. (1974). The early development of inferences about the visual percepts of others. *Child Development*, *45*(2), 357–366.

McGuigan, N., & Doherty, M. J. (2002). The relation between hiding skill and judgment of eye direction in preschool children. *Developmental Psychology*, *38*(3), 418–427.

Meltzoff, A. N., & Brooks, R. (2007). Eyes wide shut: The importance of eyes in infant gaze following and understanding other minds. In . D. M. E. R. Flom K. Lee (Ed.), *Gaze following: Its development and significance* (pp. 217–241). Mahwah, NJ: Erlbaum.

Meneghetti, C., Pazzaglia, F., & De Beni, R. (2012). Which spatial abilities and strategies predict males' and females' performance in the object perspective test? *Cognitive Processing*, *13*(1), 267–270.

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development*, *78*(2), 622–646.

Moll, H., Carpenter, M., & Tomasello, M. (2007). Fourteen-month-olds know what others experience only in joint engagement. *Developmental Science*, *10*(6), 826–835.

Moll, H., & Meltzoff, A. N. (2011). How does it look? level 2 perspective-taking at 36 months of age. *Child Development*, *82*(2), 661–673.

Moll, H., & Meltzoff, A. N. (2012). Joint attention as the fundamental basis of understanding perspectives. In *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience* (pp. 393–413). MIT Press.

Moll, H., Meltzoff, A. N., Merzsch, K., & Tomasello, M. (2013). Taking versus confronting visual perspectives in preschool children. *Developmental Psychology*, *49*(4), 646–654.

Moll, H., & Tomasello, M. (2006). Level 1 perspective-taking at 24 months of age. *British Journal of Developmental Psychology*, *24*(3), 603–613.

Montgomery, D. E. (2002). Mental verbs and semantic development. *Journal of Cognition and Development*, *3*(4), 357–384.

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

Morton, A. (1980). *Frames of mind: constraints on the common-sense conception of the mental.* Oxford University Press.

Nagel, J. (2017). Factive and nonfactive mental state attribution. *Mind & Language*, *32*(5), 525–544.

Nelson, K., & Shaw, L. K. (2002). Developing a socially shared symbolic system. In *Language, literacy, and cognitive development* (pp. 39–72). Psychology Press.

Nichols, S., & Stich, S. P. (2003). *Mindreading: an integrated account of pretence, self-awareness, and understanding other minds.* Clarendon Press/Oxford University Press.

Nordmeyer, A. E., & Frank, M. C. (2014). The role of context in young children's comprehension of negation. *Journal of Memory and Language*, *77*, 25–39.

Olineck, K. M., & Poulin-Dubois, D. (2005). Infants' ability to distinguish between intentional and accidental actions and its relation to internal state language. *Infancy*, *8*(1), 91–100.

Olineck, K. M., & Poulin-Dubois, D. (2007). Imitation of intentional actions and internal state language in infancy predict preschool theory of mind skills. *European Journal of Developmental Psychology*, *4*(1), 14–30.

O'Neill, D. K. (1996). Two-year-old children's sensitivity to a parent's knowledge state when making requests. *Child Development*, *67*(2), 659–677.

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

Pascual, B., Aguado, G., Sotillo, M., & Masdeu, J. C. (2008). Acquisition of mental state language in spanish children: a longitudinal study of the relationship between the production of mental verbs and linguistic development. *Developmental Science*, *11*(4), 454–466.

Perner, J. (1991). *Understanding the representational mind.* MIT Press.

Perner, J. (1995). The many faces of belief: Reflections on fodor's and the child's theory of mind. *Cognition*, *57*(3), 241–269.

Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, *5*(2), 125–137.

Perner, J., et al. (2010). Who took the cog out of cognitive science? mentalism in an era of anti-cognitivism. In *Cognition and neuropsychology* (pp. 255–276). Psychology Press.

Perner, J., & Roessler, J. (2012). From infants' to children's appreciation of belief. *Trends in Cognitive Sciences*, *16*(10), 519–525.

Perner, J., Stummer, S., Sprung, M., & Doherty, M. (2002). Theory of mind finds its piagetian perspective: Why alternative naming comes with understanding belief. *Cognitive Development*, *17*(3-4), 1451–1472.

Peterson, C., Slaughter, V., Moore, C., & Wellman, H. M. (2016). Peer social skills and theory of mind in children with autism, deafness, or typical development. *Developmental Psychology*, *52*(1), 46–57.

Peterson, C. C., & Slaughter, V. (2016). Culture and the sequence of developmental milestones toward theory of mind mastery. In *Theory of mind development in context* (pp. 47–62). Routledge.

Peterson, C. C., Wellman, H. M., & Liu, D. (2005). Steps in theory-of-mind development for children with deafness or autism. *Child Development*, *76*(2), 502–517.

Phillips, J., Ong, D. C., Surtees, A. D., 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.

Pons, F., Lawson, J., Harris, P. L., & De Rosnay, M. (2003). Individual differences in children's emotion understanding: Effects of age and language. *Scandinavian Journal of Psychology*, *44*(4), 347–353.

Poulin-Dubois, D., Rakoczy, H., Burnside, K., Crivello, C., Dörrenberg, S., Edwards, K., ... others (2018). Do infants understand false beliefs? we don't know yet–a commentary on baillargeon, buttelmann and southgate's commentary. *Cognitive Development*, *48*, 302–315.

Poulin-Dubois, D., & Yott, J. (2018). Probing the depth of infants' theory of mind: Disunity in performance across paradigms. *Developmental Science*, *21*(4), e12600.

Powell, L. J., Hobbs, K., Bardis, A., Carey, S., & Saxe, R. (2018). Replications of implicit theory of mind tasks with varying representational demands. *Cognitive Development*, *46*, 40–50.

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

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526.

Priewasser, B., Rafetseder, E., Gargitter, C., & Perner, J. (2018). Helping as an early indicator of a theory of mind: Mentalism or teleology? *Cognitive Development*, *46*, 69–78.

R Core Team. (2017). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: evidence from 14-and 18-month-olds. *Developmental Psychology*, *33*(1), 12–21.

Reuner, G., & Rosenkranz, J. (2014). *Bayley scales of infant and toddler developent, thirs edition - deutsche version [german version].* Frankfurt am Main: Pearson Assessment.

Reuter, T., Feiman, R., & Snedeker, J. (2018). Getting to no: Pragmatic and semantic factors in two-and three-year-olds' understanding of negation. *Child Development*, *89*(4), e364–e381.

Roberts, L. D. (2016). Children's early non-referential uses of mental verbs, practical knowledge, and abduction. In *Model-based reasoning in science and technology* (pp. 291–308). Springer.

Roby, E., & Scott, R. M. (2018). The relationship between parental mental-state language and 2.5-year-olds' performance on a nontraditional false-belief task. *Cognition*, *180*, 10–23.

Ruberman, W., Weinblatt, E., Goldberg, J. D., & Chaudhary, B. S. (1984). Psychosocial influences on mortality after myocardial infarction. *New England Journal of Medicine*, *311*(9), 552–559.

Rubio-Fernández, P., & Geurts, B. (2013). How to pass the false-belief task before your fourth birthday. *Psychological Science*, *24*(1), 27–33.

Rubio-Fernández, P., & Geurts, B. (2016). Don't mention the marble! the role of attentional processes in false-belief tasks. *Review of Philosophy and Psychology*, *7*(4), 835–850.

Rubio-Fernández, P., Jara-Ettinger, J., & Gibson, E. (2017). Can processing demands explain toddlers' performance in false-belief tasks? *Proceedings of the National Academy of Sciences*, *114*(19), E3750–E3750.

Rubio-Fernández, P. (2019). Theory of mind. In *The oxford handbook of experimental semiantics and pragmatics* (pp. 524–537). Oxford University Press.

Ruffman, T. (1996). Do children understand the mind by means of simulation or a theory? evidence from their understanding of inference. *Mind & Language*, *11*(4), 388–414.

Ruffman, T. (2014). To belief or not belief: Children's theory of mind. *Developmental*

Review, *34*(3), 265–293.

Ruffman, T., Garnham, W., Import, A., & Connolly, D. (2001). Does eye gaze indicate implicit knowledge of false belief? charting transitions in knowledge. *Journal of Experimental Child Psychology*, *80*(3), 201–224.

Ruffman, T., & Olson, D. R. (1989). Children's ascriptions of knowledge to others. *Developmental Psychology*, *25*(4), 601–606.

Ruffman, T., Slade, L., & Crowe, E. (2002). The relation between children's and mothers' mental state language and theory-of-mind understanding. *Child Development*, *73*(3), 734–751.

Ruffman, T., Slade, L., Devitt, K., & Crowe, E. (2006). What mothers say and what they do: The relation between parenting, theory of mind, language and conflict/cooperation. *British Journal of Developmental Psychology*, *24*(1), 105–124.

Ruffman, T., Slade, L., Rowlandson, K., Rumsey, C., & Garnham, A. (2003). How language relates to belief, desire, and emotion understanding. *Cognitive Development*, *18*(2), 139–158.

Ruffman, T., & Taumoepeau, M. (2016). Behaviour to beliefs. In *Theory of mind development in context* (pp. 93–108). Routledge.

Sabbagh, M. A., & Callanan, M. A. (1998). Metarepresentation in action: 3-, 4-, and 5-year-olds' developing theories of mind in parent–child conversations. *Developmental Psychology*, *34*(3), 491–502.

Sanefuji, W., Wada, K., Yamamoto, T., Mohri, I., & Taniike, M. (2014). Development of preference for conspecific faces in human infants. *Developmental Psychology*, *50*(4), 979–985.

Scholl, B. J., & Leslie, A. M. (1999). Modularity, development and 'theory of mind'. *Mind & Language*, *14*(1), 131–153.

Scott, R. M., & Roby, E. (2015). Processing demands impact 3-year-olds' performance in a spontaneous-response task: New evidence for the processing-load account of early false-belief understanding. *PloS One*, *10*(11), e0142405.

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, *18*(9), 668–671.

Setoh, P., Scott, R. M., & Baillargeon, R. (2016). Two-and-a-half-year-olds succeed at a traditional false-belief task with reduced processing demands. *Proceedings of the National Academy of Sciences*, *113*(47), 13360–13365.

Shahaeian, A., Peterson, C. C., Slaughter, V., & Wellman, H. M. (2011). Culture and the sequence of steps in theory of mind development. *Developmental Psychology*, *47*(5), 1239-1247.

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

Sifre, R., Olson, L., Gillespie, S., Klin, A., Jones, W., & Shultz, S. (2018). A longitudinal investigation of preferential attention to biological motion in 2-to 24-month-old infants. *Scientific Reports*, *8*(1), 2527.

Siller, M., Swanson, M. R., Serlin, G., & Teachworth, A. G. (2014). Internal state language in the storybook narratives of children with and without autism spectrum disorder: Investigating relations to theory of mind abilities. *Research in Autism Spectrum Disorders*, *8*(5), 589–596.

Simion, F., Regolin, L., & Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proceedings of the National Academy of Sciences*, *105*(2), 809–813.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). A 21 word solution. *Available at SSRN 2160588*.

Slaughter, V., Imuta, K., Peterson, C. C., & Henry, J. D. (2015). Meta-analysis of theory of mind and peer popularity in the preschool and early school years. *Child Development*, *86*(4), 1159–1174.

Slaughter, V., Peterson, C. C., & Carpenter, M. (2008). Maternal talk about mental states and the emergence of joint visual attention. *Infancy*, *13*(6), 640–659.

Sodian, B. (2011). Theory of mind in infancy. *Child Development Perspectives*, *5*(1), 39–43.

Sodian, B. (2016). Is false belief understanding continuous from infancy to preschool age. In *Core knowledge and conceptual change* (pp. 297–316). Oxford University Press.

Sodian, B., & Kristen-Antonow, S. (2015). Declarative joint attention as a foundation of theory of mind. *Developmental Psychology*, *51*(9), 1190–1200.

Sodian, B., Licata, M., Kristen-Antonow, S., Paulus, M., Killen, M., & Woodward, A. (2016). Understanding of goals, beliefs, and desires predicts morally relevant theory of mind: A longitudinal investigation. *Child Development*, *87*(4), 1221–1232.

Sodian, B., Thoermer, C., & Metz, U. (2007). Now i see it but you don't: 14-month-olds can represent another person's visual perspective. *Developmental Science*, *10*(2), 199–204.

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

Symons, D. K., Peterson, C. C., Slaughter, V., Roche, J., & Doyle, E. (2005). Theory of mind and mental state discourse during book reading and story-telling tasks. *British Journal of Developmental Psychology*, *23*(1), 81–102.

Tarampi, M. R., Heydari, N., & Hegarty, M. (2016). A tale of two types of perspective taking: Sex differences in spatial ability. *Psychological Science*, *27*(11), 1507–1516.

Tarchi, C., Bigozzi, L., & Pinto, G. (2019). The influence of narrative competence on mental state talk in kindergarten and primary school children. *The British Journal of Developmental Psychology*, *37*(4), 535–550.

Tardif, T., & Wellman, H. M. (2000). Acquisition of mental state language in mandarin- and cantonese-speaking children. *Developmental Psychology*, *36*(1), 25–43.

Taumoepeau, M., & Ruffman, T. (2008). Stepping stones to others' minds: Maternal talk relates to child mental state language and emotion understanding at 15, 24, and 33 months. *Child Development*, *79*(2), 284–302.

Taylor, M., Esbensen, B. M., & Bennett, R. T. (1994). Children's understanding of knowledge acquisition: The tendency for children to report that they have always known what they have just learned. *Child Development*, *65*(6), 1581–1604.

Thelen, E. (1991). Motor aspects of emergent speech: A dynamic approach. In *Biological and behavioral determinants of language development* (pp. 339–362). Psychology Press.

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

Tomasello, M. (1995). Joint attention as social cognition. In *Joint attention: Its origins and role in development* (pp. 103–130). Psychology Press.

Umberson, D., & Karas Montez, J. (2010). Social relationships and health: A flashpoint for health policy. *Journal of Health and Social Behavior*, *51*(1), S54–S66.

Wellman, H. M. (2016). Learning a theory of mind. In *Theory of mind development in context* (pp. 23–44). Routledge.

Wellman, H. M. (2018). Theory of mind: The state of the art. *European Journal of Developmental Psychology*, *15*(6), 728–755.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, *72*(3), 655–684.

Wellman, H. M., Fang, F., & Peterson, C. C. (2011). Sequential progressions in a theory-of-mind scale: Longitudinal perspectives. *Child Development*, *82*(3), 780–792.

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, *75*(2), 523–541.

Wellman, H. M., Lopez-Duran, S., LaBounty, J., & Hamilton, B. (2008). Infant attention to intentional action predicts preschool theory of mind. *Developmental Psychology*, *44*(2), 618–623.

Wellman, H. M., Phillips, A. T., Dunphy-Lelii, S., & LaLonde, N. (2004). Infant social attention predicts preschool social cognition. *Developmental Science*, *7*(3), 283–

288.

Wellman, H. M., Phillips, A. T., & Rodriguez, T. (2000). Young children's understanding of perception, desire, and emotion. *Child Development*, *71*(4), 895–912.

Wellman, H. M., & Woolley, J. D. (1990). From simple desires to ordinary beliefs: The early development of everyday psychology. *Cognition*, *35*(3), 245–275.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103–128.

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

Woolley, J. D., & Wellman, H. M. (1993). Origin and truth: Young children's understanding of imaginary mental representations. *Child Development*, *64*(1), 1–17.

Yeung, E., Müller, U., & Carpendale, J. I. (2019). Developmental continuity between social-cognitive skills at age 2 and false belief understanding at age 4. *Cognitive Development*, *50*, 157–166.

Yott, J., & Poulin-Dubois, D. (2016). Are infants' theory-of-mind abilities well integrated? implicit understanding of intentions, desires, and beliefs. *Journal of Cognition and Development*, *17*(5), 683–698.

Zeger, S. L., & Liang, K.-Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*, *42*(1), 121–130.

Zeger, S. L., Liang, K.-Y., & Albert, P. S. (1988). Models for longitudinal data: a generalized estimating equation approach. *Biometrics*, *44*(4), 1049–1060.

APPENDIX

A. MSLQ QUESTIONNAIRE

Code	Initialen		Datum

Fragebogen zum aktiven, mentalistischen Wortschatz von Kindern

Deutsche Version

Cognitive Development Laboratory
Centre for Research in Human Development
Concordia University

Bitte lesen Sie die Instruktion aufmerksam durch. Rückfragen beantwortet Ihnen gerne die Versuchsleiterin!

Kinder verstehen wesentlich mehr Wörter, als sie tatsächlich **selbst verwenden.** Wir sind besonders an den Wörtern interessiert, die ihr Kind **aktiv** benutzt.

Bitte lesen sie die umseitige Liste aufmerksam durch und markieren Sie die Wörter, die Sie bei Ihrem Kind bereits gehört haben. Die Worte also, die Ihr Kind **aktiv selbst** verwendet.

Falls Ihr Kind eine andere Aussprache für eines der Wörter gebraucht (z.B. „raffe" statt Giraffe oder „ghetti" statt „Spaghetti"), markieren Sie das Wort trotzdem. Bitte beachten Sie, dass dies eine Zusammenstellung mit allen Wörtern darstellt, die von vielen verschiedenen Kindern häufig verwendet werden. Machen Sie sich deshalb keine Sorgen, wenn Ihr Kind zu diesem Zeitpunkt nur einige dieser Wörter kennt.

Kreuzen Sie bitte **ja** an bei jedem Wort, das ihr Kind **verwendet** und **nein** bei jedem Wort, dass ihr Kind **noch nicht verwendet**!

Sollte Ihr Kind Wörter sagen, die einer der Kategorien zugeordnet werden können und die nicht im Fragebogen aufgeführt werden, ergänzen Sie diesen bitte!

Emotion (Gefühlszustände, Ausdruckverhalten, affektive Qualifizierung)				
toll ja □ nein □	**schön** ja □ nein □	**besser** ja □ nein □	**gut (Kind sagt z.B.: Mir geht es gut!)** ja □ nein □	**in Ordnung sein** ja □ nein □
ok ja □ nein □	**nett** ja □ nein □	**lieben** ja □ nein □	**mögen/gerne haben** ja □ nein □	**gern** ja □ nein □
lieb ja □ nein □	**lieber** ja □ nein □	**lustig** ja □ nein □	**stolz** ja □ nein □	**glücklich** ja □ nein □
überrascht ja □ nein □	**sich amüsieren** ja □ nein □	**Spaß haben** ja □ nein □	**küssen** ja □ nein □	**lachen** ja □ nein □
lächeln ja □ nein □	**umarmen** ja □ nein □	**Schätzchen/ Kindchen** ja □ nein □	**Küsschen** ja □ nein □	**sich fühlen (sich selbst gut, schlecht fühlen)** ja □ nein □
schlecht (etwas fühlt sich schlecht an) ja □ nein □	**schwer zu tun (hart, mühsam)** ja □ nein □	**eklig/scheußlich (etwas eklig finden)** ja □ nein □	**traurig** ja □ nein □	**weinen/weint** ja □ nein □
schade ja □ nein □	**scheiße (blöd, doof)** ja □ nein □	**streiten** ja □ nein □	**mag/mag nicht** ja □ nein □	**komisch/das sieht aber komisch aus** ja □ nein □
wütend ja □ nein □	**verrückt** ja □ nein □	**müssen** ja □ nein □	**gruselig** ja □ nein □	**können** ja □ nein □
wollen ja □ nein □	**brauchen** ja □ nein □	**Angst /Angst haben** ja □ nein □	**interessiert** ja □ nein □	**trösten** ja □ nein □
helfen ja □ nein □	**kuscheln** ja □ nein □	**gefällt Dir nicht** ja □ nein □	**das nervt** ja □ nein □	**hässlich** ja □ nein □
nicht erholsam ja □ nein □	**Sonstige Worte:** Kategorie „Emotion“			

Wollen (volitionale Zustände)				
brauchen (benötigen) ja □ nein □	**will/will nicht/ wollte** ja □ nein □	**möchte (was möchtest Du spielen?)** ja □ nein □	**hoffen** ja □ nein □	**wünschen** ja □ nein □
sehnen ja □ nein □	**hätte gerne** ja □ nein □	**würde gerne** ja □ nein □		
Sonstige Worte: Kategorie „Wollen“				

Physiologie (subjektive, körperliche Empfindung)				
sich anfühlen (weich, warm) ja □ nein □	**wach/ aufwachen/ aufgewacht** ja □ nein □	**müde** ja □ nein □	**Hunger haben/ ganz doll Hunger haben/ Hunger/hungrig** ja □ nein □	**Durst haben/durstig** ja □ nein □
schlecht (Kind sagt z.B. „Mir ist schlecht!“) ja □ nein □	**kalt (Kind sagt z.B. „Mir ist kalt!“)/frieren** ja □ nein □	**heiß (Kind sagt z.B. „Mir ist heiß“)** ja □ nein □	**warm (Kind sagt z.B. „Mir ist warm!“)/wärmt** ja □ nein □	**schlafen/ schläfrig/ schlaflos/ heia machen** ja □ nein □
schwer, mühsam (z.B. es geht so schwer!) ja □ nein □	**weh tun** ja □ nein □	**schmutzig** ja □ nein □	**unordentlich** ja □ nein □	**eklig** ja □ nein □
schlecht (sich schlecht fühlen) ja □ nein □	**sehen** ja □ nein □	**anschauen** ja □ nein □	**beobachten** ja □ nein □	**hören/zuhören** ja □ nein □
schmecken ja □ nein □	**riechen** ja □ nein □	**eklig (sich eklig fühlen)** ja □ nein □	**schwindlig sein** ja □ nein □	**verschlucken** ja □ nein □
„kaka/pipi machen“ „pullern“ „auf die Toilette gehen“ ja □ nein □	**Sonstige Worte:** Kategorie „Physiologie“			

Fähigkeiten / Obligation (Modalitäten)/Moral				
können **(Erlaubnis**) ja □ nein □	**soll/soll nicht/sollten** ja □ nein □	**kann/kann nicht „etwas schaffen“** ja □ nein □	**darf /dürfen** ja □ nein □	**lassen** ja □ nein □
müssen/muss/ muss nicht/ unbedingt müssen ja □ nein □	**schlecht(im moralischen Sinn)** ja □ nein □	**gut (im moralischen Sinn)** ja □ nein □	**gemein (im moralischen Sinn)** ja □ nein □	**versuchen, probieren** ja □ nein □
richtig/falsch (etwas richtig machen) ja □ nein □	**Recht/Unrecht** ja □ nein □	**gut/schlecht (z.B. etwas gut machen)** ja □ nein □	**„Du Frechdachs!“** **„Gutes Kind !“** **„Mein Schatz!“ „Eifrig!“** **(bewertende Ausdrücke)** ja □ nein □	
Sonstige Worte: Kategorie „Fähigkeiten / Obligation“				

Kognition (mentale Zustände)				
wissen/weiß ich nicht ja □ nein □	**vielleicht** ja □ nein □	**denken** ja □ nein □	**glauben** ja □ nein □	**erinnern** ja □ nein □
vergessen ja □ nein □	**dürfen** ja □ nein □	**verstehen** ja □ nein □	**so tun als ob** ja □ nein □	**träumen** ja □ nein □
echt ja □ nein □	**raten** ja □ nein □	**meinen** ja □ nein □	**vermuten** ja □ nein □	**abgelenkt sein** ja □ nein □
sehen/hören (im epistemischen Sinn) ja □ nein □	**Sonstige Worte:** Kategorie „Kognition“			

Kognitive Partikel				
eigentlich ja ☐ nein ☐	**tatsächlich** ja ☐ nein ☐	**wirklich** ja ☐ nein ☐	**offensichtlich** ja ☐ nein ☐	**doch** ja ☐ nein ☐
denn (was hast Du denn?) ja ☐ nein ☐	**aber (der ist aber groß)** ja ☐ nein ☐	**ja (das ist ja blöd)** ja ☐ nein ☐	**wohl (du bist wohl müde)** ja ☐ nein ☐	**richtig (eine richtige Eisenbahn)** ja ☐ nein ☐
wenigstens ja ☐ nein ☐	**schön als Adverb (machen, aussehen etc.)** ja ☐ nein ☐			

Modulierende Partikel				
bisschen ja ☐ nein ☐	**nur** ja ☐ nein ☐	**sehr** ja ☐ nein ☐	**ganz** ja ☐ nein ☐	**wenig** ja ☐ nein ☐
gerne/lieber/am liebsten (etwas tun) ja ☐ nein ☐	**schön (z.B. spielen)** ja ☐ nein ☐	**ruhig (z.B. spielen)** ja ☐ nein ☐	**wild(z.B. spielen)** ja ☐ nein ☐	**toll (z.B. gemacht)** ja ☐ nein ☐
gut/besser/am besten (da geht's gut) ja ☐ nein ☐	**gar, überhaupt (überhaupt nicht spielen)** ja ☐ nein ☐			

Frage zum Verb **„wissen“**

Zutreffendes bitte ankreuzen (mehrfache Kreuze möglich)

Verwendet ihr Kind den Begriff
„ich weiß“ oder „ich weiß nicht“
um zu sagen:

dass es weiß oder nicht weiß, wo z.B. ein Stofftier ist (auch spontan, also ohne dass es wortwörtlich danach gefragt wurde)	☐
wenn es schüchtern ist und etwas nicht sagen will	☐
wenn es wortwörtlich danach gefragt wurde, ob es etwas weiß oder nicht	☐
sie konnten noch nie beobachten, dass ihr Kind das Verb wissen gebraucht	☐

B. OTHER TASKS

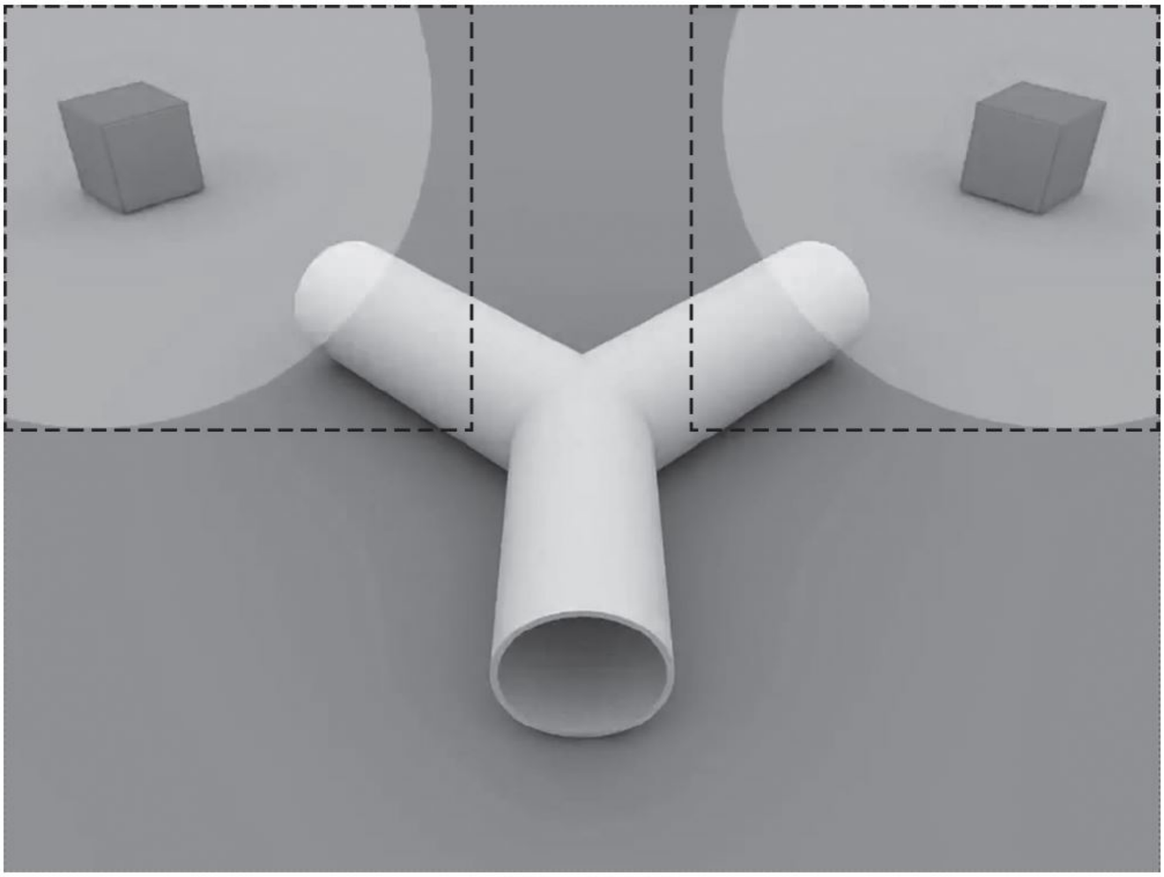

Fig. B.1: AOIs for the implicit false-belief task. Adapted from Grosse Wiesmann et al. (2017)

1 AXYB	1	*Know/don't know*	I **know** where the ball is. I **don't know** where the chair is. Can you help me?
	2	*Want/don't want*	I **want** the cup. I **don't want** the bird. Can you help me?
	3	*Don't want /want*	I **don't want** the car. I **want** the dog. Can you help me?
	4	*Don't know /know*	I **don't know** where the girl is. I **know** where the spoon is. Can you help me?

2 YABX	1	*Don't want /want*	I **don't want** the ball. I **want** the chair. Can you help me?
	2	*Know/don't know*	I **know** where the cup is. I **don't know** where the bird is. Can you help me?
	3	*Don't know /know*	I **don't know** where the car is. I **know** where the dog is. Can you help me?
	4	*Want/don't want*	I **want** the girl. I **don't want** the spoon. Can you help me?

3 XBAY	1	*Want/don't want*	I **want** the ball. I **don't want** the chair. Can you help me?
	2	*Don't know /know*	I **don't know** where the cup is. I **know** where the bird is. Can you help me?
	3	*Know/don't know*	I **know** where the car is. I **don't know** where the dog is. Can you help me?
	4	*Don't want /want*	I **don't want** the girl. I **want** the spoon. Can you help me?

4 BYXA	1	*Don't know /know*	I **don't know** where the ball is. I **know** where the chair is. Can you help me?
	2	*Don't want /want*	I **don't want** the cup. I **want** the bird. Can you help me?
	3	*Want/don't want*	I **want** the car. I **don't want** the dog. Can you help me?
	4	*Know/don't know*	I **know** where the girl is. I **don't know** where the spoon is. Can you help me?

Fig. B.2: Randomized trial order used in the know-want task.