Children's Mental State Reasoning Abilities in the Third Year of Life and Relations with Complement Syntax, Mental State Vocabulary, and Perspective-Shifting Discourse

Inaugural-Dissertation zur Erlangung des Doktorgrades der Philosophie an der Ludwig-Maximilians-Universität München

vorgelegt von

Larissa Jasmin Kaltefleiter

aus München

München, 2022

Erstgutachter: Prof. Dr. Beate Sodian

Zweitgutachter: Prof. Dr. Markus Paulus

Drittes Mitglied der Prüfungskommission: PD Dr. Tobias Schuwerk

Tag der mündlichen Prüfung (Disputation): 26. Juli 2022

Acknowledgments

The research presented here was supported by the Elitenetzwerk Bayern (ENB) and the Deutsche Forschungsgemeinschaft (DFG). I would like to extend my sincere gratitude to the opportunities provided by both the ENB and the DFG.

I would like to thank my supervisor Prof. Dr. Beate Sodian for her support and guidance during the past years, for providing valuable feedback, and for allowing me to benefit from her experience and expertise. I would like to thank Prof. Dr. Markus Paulus for being my second supervisor and for insightful comments on my work during research colloquia and research group meetings.

I am thankful for the interdisciplinary feedback and perspective on my work supplied by colleagues and professors of the REASON program supervised by Prof. Dr. Frank Fischer as well as the members of the DFG research group "Crossing the Borders between Language, Cognition, and the Brain" led by Prof. Dr. Barbara Höhle. I enjoyed the experience provided by being part of these two groups and getting to know their research from a little further afield.

I would like to express my gratitude to PD Dr. Tobias Schuwerk and PD Dr. Susanne Kristen-Antonow for their helpful comments, support, and for always being reachable for questions. Special thanks go to PD Dr. Tobias Schuwerk for supporting me with his knowledge, skills, and expertise from the creation of my Zulassungsarbeit under his supervision up to now. I am really thankful for all the exciting research projects I was able to assist with or peek at during my time as his student and research assistant. I very much enjoyed sharing an office with Dr. Özgün Köksal. Even though Covid-19 did not allow us to spend much of the past years together there, I always loved and appreciated our chats, lunchtimes, and status updates. Thank you for being such a kind person and a great office mate.

Moreover, I would like to thank Prof. Dr. Daniel Sommerhoff and Prof. Dr. Stefan Ufer for everything I have learned during my time as a student assistant at the chair of mathematics education and during my work for the COSIMA research group. During this time, I have acquired important new skills that I feel I was able to later apply also in my work in developmental psychology.

I am very grateful towards all of my student assistants for their assistance in data collection and data coding and their continued enthusiasm for the research project.

Thank you to my parents and my sister for their love and encouragement. I am very grateful for your lasting interest in my work. Thanks to all of my friends for the lovely conversations, dinner parties, and game nights during the past years. Max, thank you for being such a wonderful, loving, and caring partner. Thank you for always being there for me, for listening, providing feedback and helpful support, and for brightening up my days.

Elaborate Summary

Children's development of a Theory of Mind has probably been one of the most well-researched topics in childhood developmental psychology in the past 40 years. The aspect of a Theory of Mind in childhood that has received most attention is the understanding of false beliefs, this means, the comprehension that somebody can have a belief about something in the world which contradicts the actual reality. Numerous primary studies and a meta-analysis have revealed that children begin to succeed in attributing such false beliefs to themselves and others from age three to five (Wellman et al., 2001; Wimmer & Perner, 1983). Children's comprehension of other mental states (such as desires, perception, or knowledge) develops earlier (Masangkay et al., 1974; Pratt & Bryant, 1990; Ruffman & Olson, 1989; Wellman & Woolley, 1990) but has received considerably less attention in Theory of Mind research (Beaudoin et al., 2020; Bloom & German, 2000). Around the turn of the century, evidence started to accumulate which demonstrated that children understand false beliefs much earlier than assumed (for an overview, see Scott & Baillargeon, 2017). Researchers employed novel, spontaneous-response paradigms to assess infants' and young toddlers' implicit reasoning about false beliefs in a nonverbal way. These studies showed that infants, children, and adults implicitly attribute false beliefs in several different scenarios (Buttelmann et al., 2009; Clements & Perner, 1994; Onishi & Baillargeon, 2005; Schneider, Bayliss et al., 2012; Southgate et al., 2007; Wang et al., 2012). Currently, however, unsuccessful replication attempts of these studies have started dominating the field (Kampis et al., 2021; Kulke & Rakoczy, 2018; Paulus & Sabbagh, 2018), leading to what some researchers consider a replication crisis of implicit Theory of Mind (Poulin-Dubois et al., 2018).

While children's language acquisition is well-known to play a crucial role in Theory of Mind development between ages three and six, the evidence of an implicit Theory of Mind raises the question of which role language plays for the acquisition of these earlier emerging Theory of Mind abilities. Moreover, language skills may play a crucial role also in the transition from an early implicit to a later explicit understanding of mental states. For Theory of Mind development from age three on, different researchers have proposed that either knowledge about complement syntax, the acquisition of mental state vocabulary, or participation in perspective-shifting conversations with adults is most relevant for acquiring a Theory of Mind. Olson (1988) argued that children need to acquire a metalanguage suitable for talking about mental states to come to understand others' mental states. However, many mental state

verbs require complement syntax constructions to refer to a mental state content. Thus, de Villiers (2005) emphasized that only the acquisition of complement syntax equips children with the appropriate representational format for understanding and representing false beliefs. Complement sentences under non-factive matrix verbs also have the feature that they can be false while the whole sentence remains true. This feature makes complement syntax particularly suitable for representing beliefs that contrast with reality (de Villiers & de Villiers, 2009). However, Harris (1996, 1999) assumed that children's repeated involvement in information-exchanging, perspective-shifting conversations with their parents confronts them with situations in which two people have different beliefs about the same thing or in which the belief of a person conflicts with reality. From these instances, children learn about differences in epistemic states and about counterfactual states.

Empirically, cross-sectional, longitudinal, and training studies provided supporting evidence for the role of each of these aspects of language for children's false belief understanding from age three to four on. In contrast, the influence of different aspects of language on the acquisition of other mental states in the third year of life has rarely been investigated. However, research on children's language acquisition demonstrated that children use a variety of different mental state terms, including cognitive terms (Harris et al., 2017c; Kauschke, 2012; Kristen et al., 2012, 2014), and produce complement syntax constructions with a number of complement-taking verbs (Bloom et al., 1989; Diessel & Tomasello, 2001) already in the third year of life. Moreover, children show considerable proficiency for engaging in conversations with adults in the third year of life (Pan & Snow, 1999; Rohlfing, 2019; Snow et al., 1996). Against this background, this dissertation contributes to answering the following three overarching research questions: 1) What early Theory of Mind abilities do children possess in the third year of life, and how are they interrelated? 2) What role does children's sensitivity to the syntax of complementation play in their acquisition of early Theory of Mind abilities in the third year of life? 3) What role do children's emerging mental state vocabulary and their engagement in perspective-shifting discourse play in their acquisition of early Theory of Mind abilities in the third year of life? Three empirical studies were conducted with children between 24 and 52 months, and their results are reported in the three enclosed articles.

Article 1 attempted to replicate the original finding of above-chance false belief-congruent action anticipation in a multi-trial anticipatory-looking false belief task (Grosse Wiesmann et al., 2017). Children longitudinally participated in the task at 27, 36, and 52 months.

Using the original combined score of two false belief conditions, children's performance did not differ from chance at all three measurement points. Follow-up analyses yielded that children only performed above chance in one of the two false belief conditions (FB1) but below chance in the other (FB2), providing mixed evidence of two- to four-year-olds' belief-based action anticipations. In combination with the pattern of findings from replication attempts of other anticipatory-looking studies, the results from Article 1 indicate that higher working memory and attention demands in the FB2 condition compared to the FB1 condition might negatively affect performance. Thus, the mixed findings in Article 1 add to the accumulating evidence of unsuccessful replication attempts of anticipatory-looking paradigms and question whether anticipatory-looking paradigms are suitable to reliably assess children's and adults' implicit Theory of Mind abilities (Poulin-Dubois et al., 2018).

Article 2 investigated longitudinal relations between children's complement syntax competence assessed at 33 and 36 months and their mental state reasoning skills measured at 27 or 33 and at 36 months. While complement syntax at 33 months was predictive of children's visual and epistemic perspective-taking and metacognition at 36 months, even when controlling for children's general language abilities, earlier mental state reasoning was not predictive of later complement syntax skills. This pattern of findings suggests a causal role of children's early complement syntax skills for their understanding of mental states such as perception and knowledge in the third year of life. These findings extend the propositions of linguistic determinism theory by de Villiers (2005), which only argued for a causal role of complement syntax for understanding and representing false beliefs. Thus, children's knowledge about complement syntax seems to affect their reasoning about mental states already before children master the standard false belief task. Moreover, Article 2 found cross-sectional relations between early false belief understanding and complement syntax at 33 months. While this finding is in line with the assumption that complement syntax fosters false belief understanding, de Villiers (2005) argued that the complement syntax structure is only fully mastered around the age of 3.5 years, and only once this structure is completely understood can it support false belief reasoning. Article 2's findings, however, indicate that complement syntax is relevant for false belief understanding earlier than assumed.

Article 3 investigated the effectiveness of a microgenetic mental state training that repeatedly engaged children in perspective-shifting discourse enriched with mental state terms in contrast to a comparable complement syntax training and a time-lag control group.

Children's performance in the seeing-knowing task and the metacognition task improved more in the mental state training group than in the other two groups. However, surprisingly, the complement syntax training did not result in more substantial improvement in the complement syntax task than in the other two groups. Thus, the lack of significant effects of the complement syntax training on children's mental state reasoning abilities can only be interpreted cautiously. The mental state training's effects suggest that, as proposed by Harris (1996, 1999), children's repeated engagement in conversations in which differences between different people's knowledge states are discussed is beneficial for acquiring a Theory of Mind. Article 3 discovered that the mental state training fostered children's comprehension of epistemic states. However, when controlling for children's pre-training complement syntax skills in the analyses of the training effects, the mental state training's effect on the metacognition task disappeared, and complement syntax skills at 33 months turned out to be a predictor of metacognition across groups and measurement points. This finding adds to Article 2's result that complement syntax plays a role in children's Theory of Mind development in the third year of life. Overall, Article 3 presented one of the first Theory of Mind training studies conducted with children below age three. Moreover, unlike in other training studies, a linguistic intervention fostered comprehension of mental states other than false beliefs. Article 3 also found that mental state training children's pre-training productive vocabulary for cognitive particles predicted their training-related improvement in the seeing-knowing task. Thus, productive vocabulary for cognitive particles, as a particularly complex part of cognition terms, might indicate children's capability to profit from such a linguistic intervention.

Overall, this dissertation demonstrates children's ability to reason about mental states in the third year of life and elucidates how advancements in this ability go hand in hand with increasingly proficient language skills. Both, children's complement syntax knowledge and their participation in perspective-shifting discourse supported their reasoning about perception and knowledge already in the third year of life. These findings extend what numerous studies have revealed about the influence of complement syntax, mental state language, and participation in conversations on children's false belief understanding between ages three and five. The third year of life is a critical time period in children's socio-cognitive development in that they transition from a more implicit understanding to an explicit reasoning about mental states. Around the same time, children make considerable advances in their language development. However, this dissertation's finding that both developments are related in the third

year of life needs to be further investigated by studies focusing on this time period. Moreover, the extent of young children's implicit reasoning about false beliefs needs to be further assessed using novel paradigms and stimuli. This dissertation presented some additional evidence that anticipatory-looking paradigms may not reliably capture implicit false belief understanding and emphasized the relevance of large-scale studies to evaluate the replicability of theoretically meaningful effects. Finding reliable and replicable evidence of implicit Theory of Mind in infancy would inform theoretical ideas about how and at which age children come to understand others' minds.

Deutsche Zusammenfassung

Die Entstehung einer Theory of Mind im Kindesalter ist vermutlich eines der am besten untersuchten Themengebiete in der Entwicklungspsychologie des Kindesalters in den letzten 40 Jahren gewesen. Der Aspekt einer Theory of Mind im Kindesalter, der am meisten Aufmerksamkeit erhalten hat, ist das Verstehen von falschen Überzeugungen, das heißt, das Verständnis, dass jemand eine Überzeugung zu etwas in der Welt haben kann, die im Widerspruch zur tatsächlichen Wirklichkeit steht. Zahlreiche Primärstudien sowie eine Metaanalyse haben ergeben, dass Kinder ab dem Alter von drei bis fünf Jahren anfangen, erfolgreich anderen und sich selbst solche falschen Überzeugungen zuzuschreiben (Wellman et al., 2001; Wimmer & Perner, 1983). Das Verständnis von Kindern für andere mentaler Zustände (wie z.B. Wünsche, visuelle Wahrnehmung, Wissen) entwickelt sich bereits früher (Masangkay et al., 1974; Pratt & Bryant, 1990; Ruffman & Olson, 1989; Wellman & Woolley, 1990), wurde aber in der Theory of Mind-Forschung deutlich weniger beachtet (Beaudoin et al., 2020; Bloom et al., 2000). Um die Jahrhundertwende begann sich Evidenz anzuhäufen, dass Kinder falsche Überzeugungen viel früher als bisher angenommen verstehen (siehe Scott & Baillargeon, 2017 für eine Übersicht). Forscher verwendeten neuartige Paradigmen, die die spontane Reaktion der Teilnehmer messen, um das implizite Schlussfolgern über falsche Überzeugungen bei Säuglingen und jungen Kleinkindern nonverbal zu erfassen. Diese Studien zeigten, dass Säuglinge, Kinder und Erwachsene in verschiedenen Szenarien implizit falsche Überzeugungen zuschreiben (Buttelmann et al., 2009; Clements & Perner, 1994; Onishi & Baillargeon, 2005; Schneider, Bayliss et al., 2012; Southgate et al., 2007; Wang et al., 2012). Gegenwärtig haben jedoch nicht erfolgreiche Replikationsversuche dieser Studien angefangen, das Feld zu dominieren (Kampis et al., 2021; Kulke & Rakoczy, 2018; Paulus & Sabbagh, 2018), was laut manchen Forschern zu einer Replikationskrise der impliziten Theory of Mind geführt hat (Poulin-Dubois et al., 2018).

Während es wohlbekannt ist, dass der kindliche Spracherwerb eine entscheidende Rolle in ihrer Theory of Mind-Entwicklung im Alter zwischen drei und sechs Jahren spielt, wirft die Evidenz zu impliziter Theory of Mind die Frage auf, welche Rolle Sprache beim Erwerb dieser sich früher entwickelnden Theory of Mind-Fähigkeiten spielt. Darüber hinaus spielen Sprachfähigkeiten möglicherweise auch eine entscheidende Rolle beim Übergang von einem frühen impliziten zu einem späteren expliziten Verständnis mentaler Zustände. Für die Theory of Mind-Entwicklung ab dem Alter von drei Jahren haben verschiedene Forscher vorgeschlagen, dass entweder das Wissen über Komplementsyntax, der Erwerb mentalen Vokabulars

oder die Teilnahme an perspektivenwechselndem Diskurs mit Erwachsenen am meisten relevant für den Erwerb einer Theory of Mind ist. Olson (1988) argumentierte, dass Kinder eine Metasprache erwerben müssen, die geeignet ist, um über mentale Zustände zu sprechen, damit sie anfangen können, die mentalen Zustände anderer zu verstehen. Jedoch benötigen viele mentale Verben Komplementsyntaxkonstruktionen, um den Inhalt eines mentalen Zustands auszudrücken. Daher betonte de Villiers (2005), dass erst der Erwerb der Komplementsyntax Kinder mit dem passenden Repräsentationsformat ausstattet, um falsche Überzeugungen zu verstehen und zu repräsentieren. Komplementsätze unter nicht-faktiven Matrixverben haben zudem die Eigenschaft, dass sie falsch sein können, während der Gesamtsatz wahr bleibt. Diese Eigenschaft macht Komplementsätze besonders geeignet, um Überzeugungen, die der Realität widersprechen, zu repräsentieren (de Villiers & de Villiers, 2009). Dagegen nahm Harris (1996, 1999) an, dass Kinder durch die wiederholte Beteiligung an informationsaustauschenden, perspektivenwechselnden Konversationen mit ihren Eltern mit Situationen konfrontiert werden, in denen zwei Personen verschiedene Überzeugungen zu der gleichen Sache haben, oder in denen die Überzeugung einer Person der Realität widerspricht. Durch diese Gelegenheiten erlangen Kinder Kenntnisse über Unterschiede in epistemischen Zuständen und über kontrafaktische Zustände.

Empirische Querschnitt-, Längsschnitt- und Trainingsstudien lieferten unterstützende Belege für die Rolle jedes dieser Aspekte von Sprache für das Verständnis von falschen Überzeugungen bei Kindern ab dem Alter von drei bis vier Jahren. Der Einfluss verschiedener Aspekte von Sprache für den Erwerb anderer mentaler Zustände im dritten Lebensjahr wurde dagegen selten untersucht. Forschung zum kindlichen Spracherwerb zeigte jedoch, dass Kinder bereits im dritten Lebensjahr eine Vielzahl an unterschiedlichen mentalen Begriffen, einschließlich kognitiver Begriffe, verwenden (Harris et al., 2017c; Kauschke, 2012; Kristen et al., 2012, 2014) und Komplementsyntaxkonstruktionen mit einer Reihe an Verben produzieren (Bloom et al., 1989; Diessel & Tomasello, 2001). Zudem zeigen Kinder im dritten Lebensjahr beträchtliche Fähigkeit, Gespräche mit Erwachsenen zu führen (Pan & Snow, 1999; Rohlfing, 2019; Snow et al., 1996). Vor diesem Hintergrund trägt diese Dissertation dazu bei, die folgenden drei übergreifenden Forschungsfragen zu beantworten: 1) Welche frühen Theory of Mind-Fähigkeiten besitzen Kinder im dritten Lebensjahr und wie hängen sie miteinander zusammen? 2) Welche Rolle spielt die Sensitivität der Kinder für Komplementsyntax für ihren Erwerb früher Theory of Mind-Fähigkeiten im dritten Lebensjahr? 3) Welche Rolle spielen das

aufkommende mentale Vokabular der Kinder und ihre Beteiligung an perspektivenwechselndem Diskurs für ihren Erwerb früher Theory of Mind-Fähigkeiten im dritten Lebensjahr? Drei empirische Studien wurden mit Kindern im Alter von 24 bis 52 Monaten durchgeführt und deren Ergebnisse werden in den drei beigefügten Artikeln berichtet.

Artikel 1 versuchte das ursprüngliche Resultat von überzufälliger, auf falschen Überzeugungen basierender Handlungsvorhersage in einer antizipatorischen Blickzeitaufgabe zu falschen Überzeugungen, bestehend aus mehreren Durchgängen (Grosse Wiesmann et al., 2017), zu replizieren. Die Kinder nahmen längsschnittlich im Alter von 27, 36 und 52 Monaten an der Aufgabe teil. Unter Verwendung des originalen Testwertes, der sich aus zwei Bedingungen zu falschen Überzeugungen zusammensetzt, unterschied sich die Leistung der Kinder an allen drei Messzeitpunkten nicht vom Zufall. Weiterführende Analysen ergaben, dass die Kinder nur in einer der beiden falsche Überzeugungs-Bedingungen (FB1) überzufällig abschnitten, jedoch unterzufällig in der anderen (FB2), was gemischte Evidenz hinsichtlich der überzeugungsbasierten Handlungsvorhersage zwei- bis vierjähriger Kinder darstellt. Kombiniert mit dem Ergebnismuster von Replikationsversuchen anderer antizipatorischer Blickzeitstudien deuten die Ergebnisse von Artikel 1 darauf hin, dass höhere Anforderungen an Arbeitsgedächtnis und Aufmerksamkeit in der FB2-Bedingung verglichen mit der FB1-Bedingung möglicherweise die Leistung negativ beeinflussen. Die gemischten Ergebnisse in Artikel 1 ergänzen die sich ansammelnde Evidenz nicht erfolgreicher Replikationsversuche von antizipatorischen Blickzeitparadigmen und stellen in Frage, ob antizipatorische Blickzeitparadigmen geeignet sind, um implizite Theory of Mind-Fähigkeiten bei Kindern und Erwachsenen reliabel zu erfassen (Poulin-Dubois et al., 2018).

Artikel 2 untersuchte längsschnittliche Zusammenhänge zwischen der Komplementsyntaxkompetenz der Kinder im Alter von 33 und 36 Monaten und ihren Fähigkeiten über mentale Zustände nachzudenken im Alter von 27 bzw. 33 und 36 Monaten. Während Komplementsyntax mit 33 Monaten die visuellen und epistemischen Perspektivübernahmefähigkeiten und die Metakognition der Kinder im Alter von 36 Monaten vorhersagte, selbst wenn für allgemeine Sprachfähigkeiten kontrolliert wurde, war das Schlussfolgern über mentale Zustände zum früheren Zeitpunkt nicht prädiktiv für spätere Komplementsyntaxfähigkeiten. Dieses Ergebnismuster deutet auf eine kausale Rolle früher Komplementsyntaxfähigkeiten von Kindern für ihr Verständnis mentaler Zustände wie visueller Wahrnehmung und Wissen im dritten Lebensjahr hin. Diese Ergebnisse erweitern die Aussagen der Theorie des

linguistischen Determinismus von de Villiers (2005), welche nur eine kausale Rolle von Komplementsyntax für das Verstehen und Repräsentieren falscher Überzeugungen postuliert haben. Das Wissen der Kinder über Komplementsyntax scheint also ihr Nachdenken über mentale Zustände zu beeinflussen schon bevor sie die Standardaufgabe zu falschen Überzeugungen beherrschen. Darüber hinaus fand Artikel 2 querschnittliche Zusammenhänge zwischen frühem Verständnis falscher Überzeugungen und Komplementsyntax im Alter von 33 Monaten. Während dieses Resultat mit der Annahme übereinstimmt, dass Komplementsyntax das Verstehen falscher Überzeugungen unterstützt, argumentierte de Villiers (2005), dass die Komplementsyntaxstruktur erst im Alter von 3,5 Jahren vollständig beherrscht wird und diese Struktur erst sobald sie ganz verstanden ist das Nachdenken über falsche Überzeugungen unterstützen kann. Die Ergebnisse von Artikel 2 deuten jedoch darauf hin, dass Komplementsyntax früher als angenommen für das Verstehen falscher Überzeugungen relevant ist.

Artikel 3 untersuchte die Effektivität eines mikrogenetischen mentale Zustände-Trainings, welches die Kinder wiederholt in perspektivenwechselndem Diskurs angereichert mit mentalen Begriffen involvierte, im Gegensatz zu einem vergleichbaren Komplementsyntaxtraining und einer Kontrollgruppe ohne Training. Die Leistung der Kinder in der Aufgabe zum Sehen und Wissen und in der Metakognitionsaufgabe verbesserte sich mehr in der mentale Zustände-Trainingsgruppe als in den anderen beiden Gruppen. Jedoch resultierte das Komplementsyntaxtraining überraschenderweise nicht in einer deutlicheren Verbesserung in der Komplementsyntaxaufgabe verglichen mit den anderen beiden Gruppen. Daher kann das Fehlen signifikanter Effekte des Komplementsyntaxtrainings auf die Fähigkeiten zum Schlussfolgern über mentale Zustände nur mit Vorsicht interpretiert werden. Die Effekte des mentalen Zustände-Trainings deuten darauf hin, dass, wie von Harris (1996, 1999) vorgeschlagen, die wiederholte Involvierung in Unterhaltungen, in denen Unterschiede zwischen den Wissenszuständen verschiedener Personen diskutiert werden, von Vorteil für den Erwerb einer Theory of Mind bei Kindern ist. Artikel 3 stellte fest, dass das mentale Zustände-Training das Verstehen epistemischer Zustände bei den Kindern begünstigte. Wenn in den Analysen der Trainingseffekte jedoch für die Komplementsyntaxfähigkeiten der Kinder vor dem Training kontrolliert wurde, verschwand der Effekt des mentale Zustände-Trainings auf die Metakognitionsaufgabe und die Komplementsyntaxfähigkeiten mit 33 Monaten erwiesen sich als Prädiktor von Metakognition über Gruppen und Messzeitpunkte hinweg. Dieses Resultat ergänzt das Ergebnis aus Artikel 2, dass Komplementsyntax für Kinder eine Rolle in der Entwicklung

der Theory of Mind im dritten Lebensjahr spielt. Insgesamt präsentierte Artikel 3 eine der ersten Theory of Mind-Trainingsstudien, die mit Kindern jünger als drei Jahren durchgeführt wurden. Zudem förderte, anders als in anderen Trainingsstudien, eine linguistische Intervention erfolgreich das Verstehen anderer mentaler Zustände als falsche Überzeugungen. Artikel 3 fand ebenso heraus, dass der produktive Wortschatz kognitiver Partikel vor dem Training bei den mentale Zustände-Trainingskindern die durch das Training bedingte Verbesserung in der Aufgabe zum Sehen und Wissen vorhersagte. Der produktive Wortschatz kognitiver Partikel, als besonders komplexer Bestandteil kognitiver Begriffe, könnte also ein Indikator für das Potential der Kinder, von einer solchen linguistischen Intervention zu profitieren, sein.

Insgesamt zeigt diese Dissertation die Fähigkeit von Kindern im dritten Lebensjahr über mentale Zustände nachzudenken auf und erläutert, wie Fortschritte in dieser Fähigkeit Hand in Hand mit zunehmend kompetenteren Sprachfertigkeiten gehen. Sowohl das Wissen der Kinder über Komplementsyntax als auch ihre Teilnahme an perspektivenwechselndem Diskurs unterstützten ihr Schlussfolgern über visuelle Wahrnehmung und Wissen schon im dritten Lebensjahr. Diese Ergebnisse erweitern, was zahlreiche Studien über den Einfluss von Komplementsyntax, mentaler Sprache und Teilnahme an Konversationen auf das Verstehen falscher Überzeugungen im Alter zwischen drei und fünf Jahren herausgefunden haben. Das dritte Lebensjahr ist eine kritische Zeitspanne in der sozio-kognitiven Entwicklung der Kinder insofern als, dass sie von einem eher impliziten Verständnis zu einem expliziten Nachdenken über mentale Zustände übergehen. Etwa zur gleichen Zeit machen Kinder beträchtliche Fortschritte in ihrer Sprachentwicklung. Das Ergebnis dieser Dissertation, dass beide Entwicklungen im dritten Lebensjahr miteinander zusammenhängen, muss jedoch weiter durch Studien, die sich auf diesen Zeitraum fokussieren, erforscht werden. Darüber hinaus muss das Ausmaß des impliziten Schlussfolgerns über falsche Überzeugungen bei jungen Kindern weiter mit neuartigen Paradigmen und Stimulusmaterialien erfasst werden. Diese Dissertation stellte weitere Evidenz dar, dass antizipatorische Blickzeitparadigmen möglicherweise nicht zuverlässig das implizite Verständnis falscher Überzeugungen erfassen und betonte die Relevanz groß angelegter Studien, um die Replizierbarkeit theoretisch bedeutsamer Effekte zu bewerten. Verlässliche und reproduzierbare Evidenz einer impliziten Theory of Mind im Säulingsalter zu finden, würde theoretische Vorstellungen darüber, wie und in welchem Alter Kinder die Gedanken anderer verstehen, vorantreiben.

Table of Contents

1 General Introduction

Chapter 1 provides the theoretical background to the research presented in this dissertation. Section 1.1 defines the term 'Theory of Mind', gives a historical overview of how Theory of Mind research emerged, delineates in which order children acquire the most relevant aspects of a Theory of Mind, and lastly presents three Theory of Mind theories whose explanations have dominated the field until the early 21st century. Section 1.2 describes evidence of implicit Theory of Mind abilities before age four, which has challenged some of these theories, and introduces new theoretical ideas formulated to account for these novel findings. Section 1.3 briefly highlights the relevance of language for acquiring a Theory of Mind before providing empirical findings of the advances young children make in their development of mental state vocabulary, complement syntax, and conversational abilities. Then, Sections 1.4 and 1.5 go into more detail about the role certain aspects of language (i.e., complement syntax, mental state vocabulary, and perspective-shifting discourse) play in children's Theory of Mind development. The second parts of these sections present empirical evidence supporting these theoretical ideas. Lastly, Section 1.6 highlights the overall research questions underlying this dissertation, describes the study design, and gives a short overview of the methodology and research questions examined in the three articles presented here. Translations of the used German/English example sentences to English/German can be found in Appendix A.

1.1 What is a Theory of Mind?

1.1.1 Definition of 'Theory of Mind' and Historical Outline

The term "Theory of Mind" entered research on children's socio-cognitive abilities in 1978, when Premack and Woodruff published a seminal study on the chimpanzee's ability to infer a human's intentions. Adopting this term from philosophy, they wondered whether the chimpanzee possesses a Theory of Mind. Before that date, research on children's conception and understanding of the mind had either been conducted in the Piagetian tradition (e.g., Masangkay et al., 1974 on children's visual perspective-taking skills) or had been focused on children's metacognitive skills, mainly their metamemory (e.g., Wellman, 1978). Premack and Woodruff (1978) defined the term "Theory of Mind" in the following way: "In saying that an individual has a theory of mind, we mean that the individual imputes mental states to himself and to others" (p. 515). Based on this definition, some researchers emphasize that having a Theory of Mind also entails understanding how humans' actions are shaped by their mental

states (Wellman, 2011) as well as explaining and predicting people's behavior based on their mental states (Apperly, 2012).

Commenting on Premack and Woodruff's (1978) study, three philosophers pointed out that a purely behavioral interpretation of the chimpanzee's actions in the study could not be ruled out by the utilized study design. Based on this criticism, they independently suggested a potential paradigm suitable to assess whether an ape (or a human) truly possesses a Theory of Mind (Bennett, 1978; Dennett, 1978; Harman, 1978). They all ended up proposing a procedure similar to what is now known as the standard change-of-location false belief task (see Section 1.1.2). Premack and Woodruff (1978), as well as Dennett (1978), Bennett (1978), and Harman (1978), were largely unconcerned with whether children possess a Theory of Mind. Due to Piaget's influence on developmental psychology, many childhood researchers assumed that children were egocentric in their reasoning about others. After 1978, however, the term "Theory of Mind" was quickly taken up by researchers interested in children's understanding of the mind and inspired researchers to examine young infants' gestural and verbal communication for references to mental states (Bretherton et al., 1981). In 1983, Wimmer and Perner conducted the first systematic study of belief understanding in children, adapting the theoretical ideas by Bennett (1978), Dennett (1978), and Harman (1978) to create the unexpected-transfer paradigm. Since then, numerous studies have investigated the onset, breadth, and depth of young children's understanding of mental states. In the past 40 years, Theory of Mind research has investigated antecedents and consequents of Theory of Mind, earlier and later developments than false belief understanding, reasoning about other mental states, predictors and correlates of Theory of Mind, intra- and intercultural differences in Theory of Mind, implicit Theory of Mind, and more (Flavell, 2000, 2004). According to Apperly (2012), the various routes Theory of Mind research has taken support three different ways of thinking about and conceptualizing Theory of Mind – as a conceptual problem to be solved, as a set of cognitive processes, or as a social competence that might vary among individuals. However, Apperly (2012) contends that all three ways constitute essential aspects of Theory of Mind.

1.1.2 False Belief Understanding and False Belief Tasks

As mentioned above, Bennett (1978), Dennett (1978), and Harman (1978) independently suggested a potential paradigm to assess an animal's or human's comprehension of another person's false belief as a litmus test to determine whether the animal or human possesses a Theory of Mind. They suggested the following procedure: The participant watches

how an individual puts an object into container A and leaves. The participant then observes how someone else moves the object to container B in the individual's absence. The participant should be credited with some understanding of belief if the participant expects that the first individual searches for the object in location A and not B. The critical problem faced was how to design a paradigm which allows the chimpanzee to demonstrate with a specific behavior what they expect (Dennett, 1978). This problem is circumvented when applying such a paradigm to children, who can verbally report their expectations about the individual's search behavior. Wimmer and Perner (1983) applied such an unexpected-transfer (or change-of-location) paradigm to children and found that the ability to attribute such false beliefs to others starts to develop between ages four and five, whereas three- to four-year-olds show no sign of ascribing a false belief. Children correctly attributing a false belief to an agent were mostly also able to construct a deceitful utterance aimed at leading an antagonist to the incorrect hiding location based upon the agent's point of view. Wimmer and Perner (1983) argue that "representing wrong beliefs requires the construction of two different models of the world […] and the explicit representation of the falseness relation between propositions in one model and the corresponding propositions in the other model" (p. 123). In contrast, true belief tasks can also be solved by egocentrically assuming that others know the same things as oneself and are thus not as good a test of belief understanding (Flavell, 2000).

Wimmer and his research group also constructed a second type of false belief paradigm, now known as the unexpected-content false belief task. In this task, children are presented with a familiar box (e.g., Smarties tube or Band-Aid box) and are asked to guess its (familiar) content. Subsequently, the true content of the container is revealed. Then, children are either queried about their own previous (false) belief about the content or are asked to report what another person, unaware of the true content, would believe is in the box. Only a few three-year-olds, but half of the four-year-olds, and the majority of five-year-olds succeeded in such a false belief task (Hogrefe et al., 1986; Perner et al., 1987). Two related concepts to false belief understanding are comprehending representational change and making the appearance-reality distinction. Gopnik and Astington (1988) found that both of these concepts prove difficult for three-year-olds, emerge between ages four and five, and are related to false belief understanding.

1.1.3 Age Trajectory of Theory of Mind Abilities

A meta-analysis by Wellman and colleagues (2001) aggregated the results of numerous primary studies to examine at which age children come to understand false beliefs. Across all suitable studies, half of the children succeeded on false belief tasks by the age of 44 months; and after that age, performance dramatically improved with age. By 48 months, children performed significantly above the chance level of 50% correct (Wellman et al., 2001). Additionally, the meta-analysis investigated several potential moderator variables. A deceptive motive underlying the transformation, participants' involvement in the transformation, the target object not being present, and the protagonist's belief being salient improved performance in all age groups. Temporal marking in the test question ("Where will Sally look **first** for her marble") only increased performance in the older age group. Other aspects of the task such as type of false belief task, type of test question (e.g., "Where does Sally **think** her marble is?" in contrast to "Where will Sally **look for** her marble?"), nature of the protagonist, or nature of the target object did not influence performance (Wellman et al., 2001). Thus, in sum, understanding of false beliefs emerges between ages three and four, with children performing above chance level throughout the fifth year of life.

Despite the multitude of studies on children's false belief understanding, children's comprehension of other mental states has also been emphasized as an essential part of Theory of Mind research (e.g., Beaudoin et al., 2020; Bloom & German, 2000). A meta-analysis by Wellman and Liu (2004) revealed that, in general, children understand desires before beliefs, diverse beliefs before false beliefs, and ignorance/knowledge before false beliefs as well. One exemplary finding of the latter developmental sequence was found in a study by Hogrefe et al. (1986), which applied the unexpected-content paradigm. In this study, three-year-olds correctly indicated that another child waiting outside the room does not **know** the unexpected content of the familiar container but failed to ascribe a **false belief** about the container's content to the other child. Based on the above-mentioned findings of a developmental sequence in understanding different mental states, Wellman and Liu (2004) constructed a Theory of Mind scale. The tasks on this scale form a general progression such that most children consistently pass easier tasks before more complex tasks. Thus, as children get older, they tend to pass more tasks on the scale. A German adaptation of the scale has been validated with three- to five-year-old children and yielded comparable results (Kristen et al., 2006).

Already in the first year of life, infants develop some crucial socio-cognitive abilities such as engaging in joint attention, following adults' gazes or pointing, and directing an adult's attention to an interesting object (Carpenter, Nagell et al., 1998). Around their first birthday, infants can use an adult's intentional behavior (e.g., pointing) to locate a hidden object (Behne et al., 2005) and differentially imitate only intentional but not accidental actions (Carpenter, Akhtar et al., 1998). Around the same age, infants understand the goal-directedness of others' behavior by predicting others' goal-based actions (Woodward, 1998) and by differentially interpreting an adult's action in terms of its underlying goal (Carpenter et al., 2005). By age two, children show some appreciation of others' desires (e.g., Wellman & Woolley, 1990). Children between ages two and three start to understand that another person's visual perspective might differ from their own (Level 1 perspective-taking). Already 24-month-olds have been found to understand this difference in visual perspectives within an action-based perspective-taking task (Moll & Tomasello, 2006), whereas earlier research had only credited children from 2.5 years on with such Level 1 perspective-taking abilities (Flavell et al., 1981; Masangkay et al., 1974). A recent large study observed that children learn to introspect their own visual perspective about seven months earlier than they understand another person's visual perspective (Gonzales et al., 2018). After children have acquired Level 1 visual perspective-taking abilities, they start to understand the causal relation between seeing and knowing between ages three and four (e.g., Pillow, 1989; Pratt & Bryant, 1990; Ruffman & Olson, 1989). This means, they learn that visual access to, for instance, the content of a box leads to knowledge about this content. Gonzales et al. (2018) report that children introspect their own epistemic perspective about seven months earlier than they attribute the correct epistemic perspective to someone else. While three-year-olds can infer knowledge from someone's informational access, they have more problems identifying and remembering the sources of their beliefs compared to four- and five-year-olds (Gopnik & Graf, 1988; O'Neill & Gopnik, 1991).

Recent research by Harris and colleagues emphasizes, however, that even two-year-olds have a better understanding of epistemic states than assumed: Children were found to produce a flip gesture to signal ignorance already in the second year of life, and verbal statements of ignorance appeared some months later (Harris et al., 2017a). Confronted with pictures of hard-to-name objects, 16- to 37-month-olds communicated their ignorance about the object's name in various ways (Harris et al., 2017a). From 18 months on, children also asked a

growing proportion of information-seeking questions and persisted with such questions if the adult interlocutor did not provide a satisfactory response (Harris et al., 2017a). From 30 months on, the number of information-seeking questions dominated the number of interaction-seeking questions (Harris et al., 2017b). However, even three-year-olds were not yet able to refute an unjustified or counterintuitive knowledge claim expressed by an adult experimenter, even if they had been given time to falsify the counterintuitive claim themselves (Fedra & Schmidt, 2018; Harris et al., 2017b).

Around the same time at which children's false belief understanding develops, they also start to engage in Level 2 visual perspective-taking, that is, they appreciate that the same object can be seen and represented differently depending on one's visual perspective (Flavell, 1988; Flavell et al., 1981; Masangkay et al. 1974). As mentioned in Section 1.1.2, also between ages four and five, children understand representational change (Gopnik & Astington, 1988) and learn to make the appearance-reality distinction (Flavell et al., 1983). The ability to metarepresent one's own or others' beliefs or perspectives and to simultaneously consider two alternative conflicting representations of the same object or fact underlies all four of these abilities (Gopnik & Astington, 1988) and thus seems to develop between ages three and five. This ability is preceded by children's understanding of other mental states, including intentions, desires, perception, and knowledge. After children have acquired false belief understanding, they begin to correctly attribute second-order false beliefs between ages five and six (Perner & Wimmer, 1985) and more complex mental states such as double bluff, deception, sarcasm, faux-pas, or white lies in middle childhood and adolescence (Devine & Hughes, 2013; Happé, 1994; Osterhaus & Koerber, 2021). Even higher-order judgments of false beliefs have been investigated in adult participants (Kinderman et al., 1998; Valle et al., 2015), but it can be argued that such higher-order judgments are mainly more complex because they place higher demands on working memory and executive functions, not because they require increasingly elaborate Theory of Mind reasoning (Apperly, 2012). More naturalistic assessments have also been used to assess differences in adults' Theory of Mind abilities and have been found to create substantial variance (e.g., Golan et al., 2006; Lawson et al., 2004). A recent review provides an overview of how advanced Theory of Mind abilities are and can be assessed in adolescents and adults (Osterhaus & Bosacki, 2022).

1.1.4 Theory of Mind Theories

In light of findings that children only understand false beliefs from around age four on, different groups of researchers have formulated theories to explain how a Theory of Mind emerges in the course of children's cognitive development. The three most influential theories formulated in the early 1990s are briefly summarized below:

Theory-Theory

The theory-theory by Wellman and colleagues (e.g., Gopnik & Wellman, 1992; Wellman & Gelman, 1998) compared children's acquisition of a Theory of Mind with the formulation and refinement of a scientific theory. They argued that between ages 2.5 and four, children change from a premature psychological theory to another more elaborate theory and that this change involves several gradual transitions, similar to the process of shifting from one scientific theory to a new one in the light of accumulating contradictory findings (Gopnik & Wellman, 1992). Two-year-olds' theory does not yet postulate the existence of representational states but is primarily concerned with the comprehension of desires and perceptions in a non-representational way, that is, as drives towards objects and awareness of objects (*desire psychology*). At age three, children show some signs of understanding beliefs but in a non-representational way, that is, modeled on their earlier understanding of desires and perceptions. Other people's beliefs are still barely considered in children's explanations of others' behavior, and at first, these belief contents directly reflect the world. However, as they grow older, three-year-olds become capable of explaining a finished action in terms of the agent's incorrect belief and show some conceptual understanding of states such as pretense or dream (*desire-belief psychology*). By age four to five, children have acquired a representational understanding of the mind, and all mental states are considered propositional attitudes, that is, as involving representations of reality rather than reality itself (*belief-desire psychology*). Now, new predictions and explanations of others' behavior are possible, and previous explanations need to be revised in light of the new insights (Gopnik & Wellman, 1992). Gopnik and Wellman (1992) argued that shifts in children's explanations of others' actions provide evidence for the theory-theory and that children's task performance and behavior between ages 2.5 and five demonstrate that a theory change in children's reasoning about others is taking place. Accordingly, accumulating evidence about people's behavior which cannot be explained by the initial *desire psychology* first leads to an extension of the theory into the *desire-belief psychology*, in which others' beliefs are at least sometimes considered in post-hoc explanations of behavior.

As further counter-evidence accumulates, the initial theory and its extensions can no longer explain the newly observed phenomena. Thus, the initial theory becomes falsified, and a new theory that can account for both the previously explained phenomena and the not-yet-explained phenomena is formulated (*belief-desire psychology*).

Modularity Theory

While the theory-theory involves a theory change within children's reasoning, modularity theorists such as Leslie (1994) claim that children are equipped with a domain-specific Theory of Mind mechanism (ToMM) for computing meta-representations from birth on. The ToMM is "a module that spontaneously and post-perceptually processes behaviors that are attended, and computes the mental states that contribute to them" (Scholl & Leslie, 2001, p. 697). According to Leslie (1994), one of the earliest observed products of the ToMM is children's capacity to pretend, which emerges between 18 and 24 months. However, the ToMM needs to be assisted by a selection processor (SP) for handling general executive processes. Within a typical false belief task, the ToMM computes the meta-representation of the belief, and the SP executes the domain-general demands such as inhibiting a prepotent response. Due to younger children's limitations in terms of SP capacity, three-year-olds fail false belief tasks despite correctly calculating the false belief (Leslie, 1994). Leslie and colleagues argue that the problems children and adults diagnosed with autism spectrum condition (ASC) have with Theory of Mind stem from an impairment of the ToMM since they master tasks with comparable executive demands as well as typically developing controls do (Leslie, 1994; Scholl & Leslie, 2001).

Simulation Theory

With striking difference to theory theorists, Harris (1992) assumes that children succeed in reasoning about others' mental states through a simulation process. As they grow older, children's capacity to perform increasingly accurate simulations develops further through practice within social environments. Initially, towards the end of children's first year of life, they come to understand others' intentions by adopting others' attentional or emotional stance towards the target object, which influences their own subsequent behavior towards the object (Harris, 1992). In a second step, the simulated emotional or attentional stance no longer directly affects children's behavior, but children attribute the simulated stance to the interaction partner, which allows them to engage in intentional interaction with the other person (e.g., remove the target to change the partner's emotional stance). Next,

children learn to anticipate another person's reaction even if this person's mental state differs from their own. To do so, children need to put aside their own stance and imagine another person's stance. They anticipate that the interaction partner sees or wants an object that the child themselves does not see or want. In the last step, children need to anticipate such an intentional stance towards counterfactual targets by using increased imaginative power. In the end, children can also deal with hypothetical situations contrary to reality and imagine that someone sees something or believes something that the child assumes is not the case (Harris, 1992). This entire course of development is based upon children's innate capacity to engage in joint attention with others. Younger children's failures on false belief tasks are assumed to be due to children making an egocentric simulation based on their own current mental state (Gopnik & Wellman, 1992; Harris, 1992).

Evidence for and Relevance of these Theories

According to Gopnik and Wellman (1992), children's justifications for their responses in false belief tasks support the idea of a gradual theory change taking place in children's reasoning about others. Initially, children explain others' behavior solely based on others' perceptions and desires. Around age three, children consider others' (false) beliefs at least post-hoc when explaining others' unexpected behavior. An example of this would be when others exhibited behavior that did not satisfy their desire, since they had an incorrect belief. Later, at age four, children prioritize others' beliefs over others' desires in explaining actions (Gopnik & Wellman, 1992). According to Leslie (1992), both the emergence of children's ability to pretend in the second year of life and the impairment of children and adults with ASC in reasoning about representational mental states support the idea of an innate ToMM, which is selectively impaired in ASC. Harris' (1992) simulation theory is in line with the empirically well-documented role of children's social and family environment for developing a Theory of Mind.

To date, these theories are still relevant for Theory of Mind research. Theory-theory and modularity theory lead to different predictions about infants' and young toddlers' mental state reasoning abilities and thus give different explanations for observed successful performances in implicit Theory of Mind paradigms (see Section 1.2.1). Simulation theory directly points to the role of children's engagement in conversations within their families and social environments for their development of mental state reasoning abilities (see Section 1.5). Both of these topics remain a focus of current Theory of Mind research and serve as a foundation

for this dissertation's investigation of children's Theory of Mind abilities in the third year of life and the role different aspects of language play in the acquisition of a Theory of Mind.

1.2 Early False Belief Understanding

1.2.1 Empirical Evidence of Early False Belief Understanding

Spontaneous-Response Paradigms

In 1994, Clements and Perner provided first evidence that explicit false belief understanding assessed within the traditional verbal false belief tasks might be developmentally preceded by some sort of implicit knowledge of false beliefs. They enacted a standard change-of-location false belief task (see Wimmer & Perner, 1983) accompanied by verbal narration. However, before asking the typical false belief test question, they prompted children to anticipate the protagonist's action by saying, "I wonder where he is going to look" and recorded whether children's anticipatory eye gaze went to the belief-based or the true location of the target object. They found that from age 2;11 years on, children reliably looked at the belief-based location in false belief and true belief trials. This spontaneous looking behavior indicates some implicit knowledge about false beliefs (Clements & Perner, 1994). Performance on the explicit false belief question increased with age and was substantially worse than implicit responding. Moreover, a large gap between successful performance on the implicit and the explicit measures was observed (Clements & Perner, 1994). A training study found that children who succeeded on such an implicit measure were more likely to profit from a false belief training (Clements et al., 2000). Thus, false belief understanding might develop gradually rather than all at once. Follow-up investigations using adaptations of the paradigm with three rather than two locations revealed that children's accurate looking behavior was likely not caused by attributing ignorance to the agent (Garnham & Ruffman, 2001). Moreover, the children's implicit knowledge of false beliefs appeared to be highly unconscious to them. Children bet tokens on at which of the two locations they expect the agent to show up. Betting was only affected by children's explicit, not implicit responses (Ruffman et al., 2001). Furthermore, children in a transitional period between failing and passing the explicit false belief question were rather uncertain about their explicit responses, as evident in their betting behavior.

Following these early findings on two-year-olds' implicit understanding of false belief, researchers also became interested in infants' early understanding of false belief. Leslie (1992) had assumed that the postulated ToMM is present from infancy on and should already start operating at around 18 months when children's abilities to engage in pretense become

apparent. Thus, using spontaneous-response paradigms, it should be possible to assess the depth and breadth of infants' and young toddlers' implicit knowledge of false beliefs. Violation-of-expectation, anticipatory-looking, and interactive helping paradigms have been utilized for this purpose and will be briefly characterized below.

Violation-of-expectation paradigms are derived from habituation paradigms and were first used by Baillargeon and colleagues (1985) in an infant study on object permanence. In such paradigms, infants are first familiarized with a situation. Subsequently, either a possible or an impossible event or, in the case of false belief studies, a belief-congruent or a belief-incongruent test event is presented. Higher looking times at the incongruent/impossible event compared to the congruent/possible event indicate infants' surprise or increased interest in the incongruent/impossible event (Aslin, 2007; Baillargeon et al., 1985). Following Woodward's (1998) adaptation of such a paradigm for assessing infants' expectations about goal-based actions, Onishi and Baillargeon (2005) used this type of paradigm to measure infants' expectations about belief-based actions. Infants were confronted with either a belief-congruent or a belief-incongruent event in a change-of-location situation where an agent came to hold either a true or a false belief about an object's location. Looking times revealed that infants expected the agent to search in the belief-based location, independently of whether the agent's belief was true or false (Onishi & Baillargeon, 2005). This finding was interpreted as evidence of infants' representational Theory of Mind skills.

Southgate et al. (2007) employed an anticipatory-looking paradigm to determine whether infants really attributed a false belief or merely ignorance to the agent in the Onishi and Baillargeon (2005) study. Anticipatory-looking paradigms rely on humans' tendency to anticipate actions and their outcomes when observing them. This tendency is already present in human infants by the end of the first year of life (e.g., Falck-Ytter et al., 2006; Flanagan & Johansson, 2003). Within anticipatory-looking paradigms, actions are presented which are performed based upon certain mental states of the agent (e.g., intentions or true/false beliefs). Then, participants' spontaneous predictive looking behavior is measured to infer whether participants considered the agent's mental state when anticipating their action. Several studies report that infants and adults correctly predict an agent's goal-based action in such paradigms (e.g., Cannon & Woodward, 2012; Paulus et al., 2017). Southgate and colleagues (2007) presented a video in which an agent watched how an object was originally hidden in location A. Next, the object was transferred to location B and then removed from

the scene. Dependent on the false belief condition, the agent either watched the transfer from A to B (FB1 condition) or not (FB2 condition), while the agent never watched how the object was removed from the scene. Thus, the agent held a false belief about the object's location. In combination, the conditions FB1 and FB2 serve as mutual controls that children really engage in belief-tracking to solve the task (Baillargeon et al., 2018; Kulke et al., 2019; Southgate et al., 2007). Southgate et al. (2007) argued that in the study by Clements and Perner (1994), children younger than 2;11 years did not succeed because their knowledge about the object's current location interfered with their false belief reasoning and created a reality bias. Therefore, Southgate et al. (2007) removed the object from the scene and found that most 25-month-olds directed their first look to the correct belief-based location and spent significantly more time looking at the belief-based than at the other location.

Interactive helping paradigms exploit infants' understanding of others' goals and their tendency to help others achieve these goals, which can be observed in infants as young as 18 months (Warneken & Tomasello, 2006). Buttelmann et al. (2009) confronted 2.5-year-olds, 1.5-year-olds, and 16-month-olds with a situation in which an adult needed the child's help unlocking a box and retrieving a hidden toy. For half of the children, the experimenter held a true belief about the toy's location and was deliberately searching in another box. For the other half of the children, the adult's belief about the toy's location was false, leading to erroneous searching in the wrong box. The majority of 2.5- and 1.5-year-olds helped the adult open the box they were trying to open in the true belief condition and the box now containing the toy in the false belief condition. This indicates their comprehension of the adult's goal-directed behavior and false belief about the toy's location (Buttelmann et al., 2009).

A series of studies and paradigms followed up on and extended these initial studies, probing the depth and flexibility of infants' and toddlers' early false belief reasoning abilities. Surian et al. (2007) found that 13-month-olds also expected animated agents, and not only human actors, to behave according to their true and false beliefs. Moreover, 17-month-olds attributed a false belief to a simple geometric shape if the video suggested that two shapes were chasing each other, and children anticipated the chaser to follow the chasee to the belief-based location (Surian & Geraci, 2012). Apart from the type of protagonists, the types of belief-inducing situations were also varied, primarily using violation-of-expectation paradigms: 15-month-olds realized that an actor deliberately causing a location change while looking away held a true and not a false belief about the object's new location (Träuble et al.,

2010). Furthermore, 14.5-month-old infants expected agents to behave based on their false perceptions of a situation (Song & Baillargeon, 2008), 18-month-olds successfully attributed a false belief about object identity to an agent (Scott & Baillargeon, 2009), and 2.5-year-olds correctly attributed false beliefs about the content of a container (He et al., 2011). These findings demonstrate some flexibility in infants' early false belief reasoning. Additionally, above-chance anticipatory-looking in three different types of belief-inducing situations – change-of-location, misinformation, and unexpected-content – was observed in a large sample of Chinese three- and four-year-olds (Wang et al., 2012). In contrast, a different anticipatory-looking study with three- and four-year-olds, as well as adults, found that all age groups anticipated correctly in the usual change-of-location trial but failed to attribute a false belief about object identity (Low & Watts, 2013).

In a critical test of infants' belief reasoning, 18-month-olds attributed different mental states to the same agent in the same situation depending on their own prior experience with either regular, opaque blindfolds or transparent, "trick" blindfolds (Senju et al., 2011). Only if the infants had experienced the opaque blindfolds did they expect an agent who had worn the same blindfolds during the location change of the target object to hold a false belief about the object's current location. Moreover, 2.5-year-olds held different expectations about the searching behavior of an agent who was either absolutely ignorant about the current location of the target object or held a false belief about the current location (He et al., 2011). This again suggests that children do not merely attribute ignorance to agents in false belief scenarios (see also Garnham & Ruffman, 2001 using anticipatory-looking).

Anticipatory-looking studies were also employed to investigate older children's and adults' implicit processing of beliefs. Using a paradigm with multiple trials, adults correctly looked longer at the no-object location in false than in true belief trials despite additionally engaging in a distractor task (Schneider, Bayliss et al., 2012). Thus, implicit mental state reasoning skills are not only present in infants and young toddlers, but also adults possess implicit false belief skills, which appear to work unconsciously. In line with this, Schuwerk et al. (2015, 2016) found that while neurotypical eight-year-olds and adults succeeded on an anticipatory-looking false belief task, children and adults diagnosed with ASC were impaired in their belief-based action prediction compared to the matched control groups even though the groups did not differ in their explicit Theory of Mind abilities. These findings corroborate and extend results by Senju et al. (2010), who found that six- to eight-year-old neurotypical children

performed better than age-matched children with ASC on both an implicit and an explicit false belief task. Adopting the paradigm by Surian and Geraci (2012), Meristo and colleagues (2012, 2016) investigated deaf two- and four- to eight-year-old children born into hearing families, who are significantly delayed in their explicit Theory of Mind. The deaf children performed substantially worse than groups of age-matched hearing children in the anticipatory-looking false belief task. While the hearing controls succeeded on the task, all two-year-old deaf children anticipated the incorrect location, and the four- to eight-year-old deaf children performed at chance level. Moreover, four- to eight-year-olds' performance in the anticipatory-looking task was related to their language skills and explicit false belief understanding (Meristo et al., 2016).

The findings on language-delayed deaf children raise the question of potential correlates of implicit Theory of Mind. It is well-established that explicit Theory of Mind is related to language development (Milligan et al., 2007) and executive functions (Devine & Hughes, 2014). The correlates of implicit Theory of Mind have so far rarely been assessed, and the existing studies offer a mixed picture: A crucial difference between the anticipatory-looking paradigms by Surian and Geraci (2012) and Southgate et al. (2007) is the presence or absence of the target object during the anticipatory phase. For explicit false belief tasks, the target object's presence during the response phase is assumed to evoke a "pull of the real", increasing the inhibitory demands of the task and aggravating successful task solutions. Thus, the paradigm by Southgate et al. (2007) can be considered easier since it avoids the "pull of the real". To investigate whether implicit false belief reasoning is affected by increased inhibitory demands, Wang et al. (2012) compared three- and four-year-olds' performance in a transfer version (target object present) and a removal version (target object absent) of the Clements and Perner (1994) paradigm and also assessed explicit answers. The presence of the object had a modest effect on explicit performance but barely an effect on implicit performance since implicit performance levels were close-to-ceiling in both versions. Thus, manipulation of the object's presence appeared to not affect implicit performance – at least not in older children. In contrast, Wang and Leslie (2016) applied the paradigm by Southgate et al. (2007) and compared two- to three-year-olds' and adults' performance in a low-demand false belief (identical to the original FB2 condition), high-demand false belief (object not removed), and true belief trial (agent watched transfer). Participants were more likely to look at the correct belief-based location in the low-demand false belief and in the true belief trials. However, they showed no

preference for a location in the high-demand false belief trial. Thus, children's and even adults' knowledge of the object's real location hindered their implicit false belief reasoning (Wang & Leslie, 2016). In line with this finding, Schneider, Lam et al. (2012) also found that manipulation of external cognitive load (engagement in no, a low-demand, or a high-demand distractor task) affected adults' false belief reasoning in the paradigm by Schneider, Bayliss et al. (2012).

Regarding relations with language skills, Meristo et al. (2016) argued that the impairment of deaf children born to hearing families in implicit false belief tasks stems from their lack of exposure to relevant communication, and that children's linguistic skills affect performance on implicit false belief tasks. In contrast, Low (2010) and Grosse Wiesmann et al. (2017) found that performance in an anticipatory-looking false belief task was not related to complement syntax or general language abilities in typically developing three- and four-year-olds.

To date, few researchers have investigated relations between explicit and implicit false belief reasoning. Clements and Perner (1994) and Ruffman et al. (2001) observed dissociations between performance on the implicit and the explicit part of their paradigm. Also, Grosse Wiesmann et al. (2017, 2018) observed that three- and four-year-olds' performance in an anticipatory-looking paradigm was not associated with performance in two explicit false belief tasks. However, Low (2010) found that implicit false belief understanding predicted performance in explicit false belief tasks. Only two studies have investigated longitudinal predictive relations between infants' implicit false belief understanding and their later explicit Theory of Mind: One longitudinal project found significant relations between implicit false belief understanding at 18 months and explicit false belief understanding at 48 (Thoermer et al., 2012), 50, 60, and 70 months (Kloo et al., 2020), as well as belief-based intention understanding at 60 months (Sodian et al., 2016). Another longitudinal study which, however, utilized different types of implicit false belief tasks, found no evidence of continuity between 14- and 18-month-olds' performance in violation-of-expectation and interactive helping false belief tasks and their explicit false belief understanding at four to five years (Poulin-Dubois et al., 2020).

Elicited-Response Paradigms

The previous section focused on so-called spontaneous-response paradigms, which demonstrate children's implicit false belief reasoning skills. In light of these new findings and tasks, paradigms relying on children's explicit responses, such as the original task by Wimmer and Perner (1983), have been termed elicited-response paradigms. From the first study

showing that children acquire false belief understanding around age four, researchers have sought to identify factors and task modifications allowing children to succeed at a younger age.

A study by Mitchell and Lacohée (1991) reported that making the false belief about the content of a container more physically real by posting a picture of the initial belief into a postbox increased three-year-olds' false belief performance on an unexpected-content task. This effect was not observed if an irrelevant picture was posted. Also, control conditions ascertained that children's correct replies did not refer to the content of the postbox but to the content of their initial belief (Mitchell & Lacohée, 1991).

A number of studies have adapted the formulation of the standard test question in the change-of-location false belief task. It remains unclear whether asking the think-question ("Where does Sally **think** her marble is?") or the look for-question ("Where will Sally **look for** her marble?") result in different false belief performance levels (Hala & Chandler, 1996; Wellman et al., 2001). However, several studies found that asking three-year-olds, "Where will Sally look **first** for her marble?" increased performance compared to the standard questions to either above-chance or chance performance (Siegal & Beattie, 1991; Surian & Leslie, 1999; Yazdi et al., 2006). Performance on true belief stories was unaffected by this modification, assuring that the "look first" modification did not simply trigger children to indicate the first hiding location of the target (Siegal & Beattie, 1991; Surian & Leslie, 1999; Yazdi et al., 2006). Siegal and Beattie (1991) argued that young children might misinterpret the pragmatic intent of the standard question and thus find the "look first" version easier. Surian and Leslie (1999) instead assumed that the "look first" question increases the salience of the correct location and therefore reduces the need to inhibit the prepotent incorrect response.

A further popular modification of the standard false belief task is the removal of the target object from the scene (after or instead of the displacement of the target) to reduce the so-called "reality bias" or "curse of knowledge", which has been assumed to hinder younger children's successful false belief representation. Already, Bartsch (1996) observed that three-year-olds performed better on a false-location task if the desired object was in neither of two possible locations than if it was in the container contrary to the protagonist's belief. More recently, Ghrear et al. (2021) found that in a four-container version of a change-of-location false belief task, three-year-olds performed as good as four- to six-year-olds if they were merely told that the object was moved to *another* container. However, they performed

significantly worse than the older children if they were told *which* container the object was moved to. Also, if the target object switched identity during its transfer, and three-year-olds were queried about the initial nature of the object, false belief performance tended to increase (Kikuno et al., 2007). Further research yielded that three-year-olds' involvement in planning a deception on someone facilitated their false belief reasoning while merely observing how somebody else deceived the agent did not (Hala & Chandler, 1996). Despite the above-described facilitative effects for three-year-olds obtained by different task modifications, the meta-analysis by Wellman et al. (2001) still revealed a clear age trend in false belief performance across all false belief tasks and adaptations published at that point. Thus, even though single studies found increased performance levels in three-year-olds, examining all studies together showed that three-year-olds do not consistently perform above chance level in false belief tasks.

Recently, however, Rubio-Fernández and Geurts (2013, 2016) developed a novel Duplo false belief task based on a typical change-of-location story, but they included some modifications to facilitate children's tracking of the agent's perspective: First, the agent merely turned her back to the scene but remained visible during the transfer, and second, children were required to interactively act out the agent's behavior after the transfer. Only in combination did the two modifications increase three-year-olds' false belief performance to significantly above-chance performance (Rubio-Fernández & Geurts, 2013). In a follow-up study, the authors found that keeping these two modifications but using the question "where will the agent go now?" instead of the more open-ended question "what will happen next?" still resulted in above-chance performance. In contrast, using the original Duplo task but increasing the saliency of the target object decreased performance to below chance level (Rubio-Fernández & Geurts, 2016). Recently, three replication attempts of this paradigm did not replicate the originally reported high performance levels (Dörrenberg et al., 2019; Kammermeier & Paulus, 2018; Priewasser et al., 2020) and two of these found no significant performance differences between the Duplo task and a standard false belief task (Dörrenberg et al., 2019; Priewasser et al., 2020).

While the previously described modifications increased the performance of three-year-olds to chance or even above-chance level, younger children have rarely been investigated using elicited-response false belief tasks. A recent study by Setoh et al. (2016) not only reduced inhibitory demands by removing the target object from the scene but also included

two practice questions following the same format as the test question to reduce response generation demands. They found that 30- and 33-month-olds performed significantly above chance in this low-demands change-of-location false belief task. This finding has recently been replicated in an independent lab (Grosso et al., 2019) and has been extended by the original authors to a false belief task on object identity (Scott et al., 2020). The authors further found that both simplifications were indeed necessary since a high-inhibition version with practice questions resulted in below-chance performance (Setoh et al., 2016). Moreover, 2.5-year-olds only succeeded if the type of wh-question in the practice questions matched the type of wh-question in the test question (Scott et al., 2020). These impressive findings of false belief reasoning already among two-year-olds have not gone unchallenged: Some researchers argue that the participants might have solved the task using lower-level strategies evoked by the introduction of the practice questions (Fenici & Garofoli, 2020; Rubio-Fernández et al., 2017; but see Scott et al., 2017 and Setoh et al., 2020 for counterarguments).

1.2.2 Mentalist versus Minimalist Accounts of Early Theory of Mind

The novel findings regarding infants' and young toddlers' implicit false belief reasoning raise the question of how this new evidence can be reconciled with the well-documented finding that only children between three or four to five years old start to attribute false beliefs to themselves and others (Wellman et al., 2001; Wimmer & Perner, 1983). Results from traditional false belief tasks have established the view that a fundamental change in children's thinking must take place around their fourth birthday. This change in thinking is supposed to drive their acquisition of a representational Theory of Mind. The accumulating evidence from studies using implicit measures or simplified versions of the explicit false belief task has challenged this view. In an attempt to reconcile these findings of early false belief understanding with the findings from traditional tasks, several influential theories have been formulated. They can be divided into mentalist theories and minimalist theories. With regard to the former, the following section focuses on the early-competence view mainly put forward by René Baillargeon and Rose Scott and the idea of conceptual-continuity presented in Setoh et al. (2016). Then, some important minimalist accounts, the dual-systems account by Apperly and Butterfill (2009), and the idea of developmental enrichment will be discussed.

Early-Competence/Conceptual-Continuity View and the Processing-Load Account

The early-competence view was derived from findings that infants as young as 15 months showed sensitivity to an agent's false belief in a violation-of-expectation paradigm

(Onishi & Baillargeon, 2005) and other accumulating evidence of infants' and young toddlers' early Theory of Mind abilities (Baillargeon et al., 2010; Baillargeon et al., 2016; Scott, 2017; Scott & Baillargeon, 2017). Scott (2017) argues that infants' and children's performance observed in such spontaneous-response tasks demonstrates their ability to attribute false beliefs to agents. Infants' and young toddlers' abilities to reason about a variety of different mental states (Baillargeon et al., 2016) and to flexibly attribute false beliefs (Scott, 2017) are considered evidence that "infants possess a psychological-reasoning system that provides a skeletal causal framework for interpreting and predicting the actions of agents in terms of their mental states (including false beliefs)" (Scott, 2017, p. 70). The idea that infants can attribute not only motivational or epistemic states but also counterfactual states to others (Scott & Baillargeon, 2017) clearly contradicts the traditional view that a representational Theory of Mind only develops around age four and that this development involves a fundamental change in children's thinking about others (e.g., Wellman et al., 2001).

However, if infants and young toddlers already possess a fully-fledged Theory of Mind, how can two- and three-year-olds' poor performance in traditional elicited-response false belief tasks be explained? Scott (2017) proposes a processing-load account to explain this discrepancy (see also Baillargeon et al., 2010; Scott, 2014). She argues that traditional elicited-response tasks involve at least three executive function demands: the response selection process, the response inhibition process, and working memory demands. These combined demands are assumed to overwhelm young children, leading to their failure in traditional tasks despite their capacity to attribute false beliefs. In contrast, spontaneous-response tasks impose lower processing demands, allowing even young children to succeed. As their processing skills improve, children can start to master traditional false belief tasks. Other researchers assume that inhibitory demands (Surian & Leslie, 1999; Wang & Leslie, 2016) or children's insufficient pragmatic skills in interpreting the test questions (e.g., Rubio-Fernández & Geurts, 2013, 2016; Siegal & Beattie, 1991) alone cause younger children to fail in traditional false belief tasks and explain their success in simplified tasks. However, Setoh et al. (2016) found that only reducing all three executive function demands resulted in even 2.5-year-olds' above-chance performance in an explicit change-of-location false belief task.

A common idea underlying these explanations is the assumption that substantial continuity between infants' and young children's performance in non-traditional tasks and their later performance in traditional false belief tasks should be observed (Setoh et al., 2016). A

couple of longitudinal studies observed relations between infants' comprehension of the goal-directedness of actions and their false belief understanding some years later (e.g., Aschersleben et al., 2008; Sodian et al., 2016; Wellman et al., 2004, 2008; Yamaguchi et al., 2009). Moreover, longitudinal relations were discovered between 18-month-olds' performance in an anticipatory-looking false belief task and their later explicit false belief understanding at ages four, five, and six (Kloo et al., 2020; Sodian et al., 2016; Thoermer et al., 2012). Poulin-Dubois et al. (2020), however, found no relations between infants' performance in a violation-of-expectation false belief task and their explicit false belief understanding at age five. Thus, the pattern of findings regarding continuity between earlier and later Theory of Mind abilities is mixed but has yielded some evidence that infants' psychological reasoning abilities are conceptually continuous with their later explicit Theory of Mind (for a recent review, see Sodian et al., 2020). However, correlations between earlier and later socio-cognitive abilities are also compatible with dual-systems views or the idea of developmental enrichment, which will both be described in the following sections.

Submentalizing, Behavior Rules, and Dual-Systems

The most radical non-mentalist explanation of infants' and young toddlers' surprising false belief reasoning skills was put forward by Heyes (2014). She argues that infants do not possess any capacities to attribute mental states at all but merely react to the low-level perceptual novelty of the test stimuli compared to the stimuli observed before. Similarly, Perner and Ruffman (2005) initially argued that neurons in the brain might code for recency of exposure and remember certain associations, allowing the child to respond faster to more recent and less novel test events. This would lead to reduced looking time at these events compared to novel, less recent test events.

A more exhaustive account by Ted Ruffman and Josef Perner claims that infants succeed in nonverbal false belief paradigms due to their excellent statistical learning skills and their reliance on specific behavior rules (e.g., 'an agent will search for his object where he last saw it', Perner, 2010; Perner & Ruffman, 2005; Ruffman, 2014; Ruffman & Perner, 2005). Perner (2010) argues that findings from nonverbal false belief paradigms with infants (and great apes) can easily be explained by assuming that infants rely upon specific learned behavior rules. Ruffman (2014) further elaborates on this idea: To solve nonverbal false belief tasks, infants build upon their innate capacity for statistical learning to understand patterns of behavior. Moreover, they can derive an agent's future behavior from the agent's current

perception and previous behavior. Both, Ruffman (2014) and Perner (2010) contend that infants can arrive at the correct action predictions in nonverbal false belief tasks if they directly link the agent's perception to the agent's subsequent behavior, without considering the agent's underlying mental state.

Other researchers have tried to conciliate the diverging findings from implicit and explicit false belief tasks by proposing two systems for mental state reasoning. Here, the focus will be on the most prominent dual-systems account by Apperly and Butterfill (2009). Their argumentation is based on the different demands humans' mental state reasoning system has to face: On the one hand, mental state reasoning needs to be fast and highly efficient to enable participation in quickly changing social interactions. On the other hand, mental state reasoning needs to be flexible enough to deal with various mental states and situations. Drawing analogies with number cognition, Apperly and Butterfill (2009) propose two distinct mental state reasoning systems – one efficient but inflexible system and one flexible but cognitively demanding system (see also Low et al., 2016). The efficient system works rather automatically (Frith & Frith, 2012) and allows infants to succeed in implicit nonverbal false belief tasks. The acquisition of the flexible system depends upon linguistic abilities and executive functions and enables solving explicit verbal false belief tasks (Apperly & Butterfill, 2009). In support of this idea, Low et al. (2016) presented evidence demonstrating that the cognitive processes involved in solving implicit nonverbal tasks are limited to certain types of mental state reasoning tasks (e.g., Level 1 but not Level 2 perspective-taking), highlighting the inflexibility and signature limits of these processes (see also Low & Watts, 2013).

Developmental Enrichment

Dual-systems accounts claim that the system involved in solving implicit false belief tasks and the system involved in solving explicit false belief tasks operate mainly independently. Other researchers assume that some sort of developmental enrichment happens between infants' first success on implicit tasks and the age at which children pass traditional explicit false belief tasks. Carruthers (2013) – a proponent of the mentalist view – assumes that Theory of Mind tasks differ in their complexity due to their higher or lower executive control demands. Thus, the mind-reading system already present in infants develops into the adult system through gradual conceptual enrichment and improved interaction with executive control systems (Carruthers, 2013, 2016). Ruffman (2014) argues that under the influence of growing language skills and exposure to parental talk (e.g., about mental states), infants'

implicit knowledge about behavior transforms into explicit knowledge about mental states. Consequently, individual differences in Theory of Mind might be explained by differences in parental input. A plethora of research relating parental mental state language input to children's Theory of Mind competencies (see Section 1.5.5) support this reasoning.

Tomasello (2018) explains that infants can succeed in implicit false belief tasks by predicting the agent's action solely based upon what the agent sees and knows, that is, by imagining and tracking the agent's current knowledge state. He argues that a fully-fledged explicit Theory of Mind develops based on the emergence of joint attention behavior in infancy and continued participation in conversations. Thereby, children understand that people can have different subjective perspectives on the same fact and that an objective perspective exists as well. It is not until age four to five that children understand how these different perspectives can be coordinated, which leads to their success in traditional false belief tasks at this age (Tomasello, 2018).

1.2.3 Replication Crisis of Implicit Theory of Mind

From 2005 on, Theory of Mind research was flooded with studies aimed at the new research field of infants' and young toddlers' early implicit Theory of Mind abilities. These studies initially provided more and more evidence of situations in which infants demonstrate impressive mental state reasoning abilities. However, "scientific progress requires replication of theoretically important findings" (Crivello & Poulin-Dubois, 2018, p. 55), and in the past few years, unsuccessful and partial replication studies have started to dominate the field. A recent meta-analysis by Barone et al. (2019) found that positive findings of infants' false belief understanding mainly stem from the earlier published studies assessing only small samples. Furthermore, higher success rates are obtained using violation-of-expectation paradigms than anticipatory-looking or interactive paradigms (Barone et al., 2019). In line with this, performance often does not consistently correlate across the different paradigms and gaze measures used (Dörrenberg et al., 2018; Poulin-Dubois & Yott, 2018). A systematic data article and overview by Kulke and Rakoczy (2018), as well as a special issue in *Cognitive Development* (Paulus & Sabbagh, 2018) dedicated to replication attempts of studies on infants' social cognition, have uncovered many studies which did not or only partially replicate the findings of the original implicit false belief paradigms.

While numerous studies from Baillargeon's research group have found implicit false belief understanding among infants in violation-of-expectation paradigms, replication

attempts by other researchers have often failed to find clear evidence of false belief understanding in these paradigms (Dörrenberg et al., 2018; Poulin-Dubois & Yott, 2018; Powell et al., 2018; Yott & Poulin-Dubois, 2016; but see e.g., Yott & Poulin-Dubois, 2012 for a successful replication).

Results from replication attempts of the interactive helping paradigm by Buttelmann et al. (2009) are mixed: Powell et al. (2018) replicated performance in the false belief condition but did not examine the true belief condition as a control. Poulin-Dubois and Yott (2018) found that 18-month-olds opened the box with the toy in the false belief condition, but the selected box did not differ between the false belief and the true belief condition. This indicates that infants might have chosen the box with the toy independently of the agent's belief. Crivello and Poulin-Dubois (2018) even found that 18-month-olds performed at chance level in both the false belief and the true belief condition. In contrast, Priewasser et al. (2018) observed above-chance performance only in the false belief condition but chance performance in the true belief condition. Moreover, in an additional condition with a third box, where the toy was never placed, the infants always opened the box containing the toy, independently of condition. This means, infants ignored that the agent deliberately searched in the third box that had never contained the toy. Priewasser et al. (2018) thus argued that the infants did not pay attention to the agent's mental state and simply always helped retrieve the toy.

Replication attempts of anticipatory-looking paradigms have focused on the originally assessed age groups but also on older children and adults. While the paradigm by Clements and Perner (1994) has been replicated several times (e.g., Clements et al., 2000; Garnham & Ruffman, 2001; Low, 2010; Ruffman et al., 2001; Wang et al., 2012), the paradigms by Southgate et al. (2007) and Surian and Geraci (2012) have been found particularly difficult to replicate: The paradigm by Surian and Geraci (2012) consists of a false belief and a true belief condition, and crucially longer looking at the no-object location in false belief than in true belief trials is required to attribute false belief reasoning. A replication attempt by Kulke, Reiß et al. (2018) observed that children tended to look longer at the object location independently of condition (see also Kulke & Rakoczy, 2019), whereas adults looked longer at the object location in false than in true belief trials. However, crucially, no looking time difference for the no-object location was found (Kulke, Reiß et al., 2018; Kulke, von Duhn et al., 2018). Even including verbal narration into the task to make it more similar to the paradigm by Clements and Perner (1994) did not improve children's performance (Kulke & Rakoczy, 2019).

In the Southgate et al. (2007) paradigm, above-chance performance in the FB1 and the FB2 condition is required to attribute false belief reasoning. Replication attempts usually only found above-chance performance in the FB1 condition, with the FB2 condition evoking either chance performance or even below-chance performance in children (Burnside et al., 2018; Dörrenberg et al., 2018; Kulke & Rakoczy, 2019; Kulke, Reiß et al., 2018; Kulke, von Duhn et al., 2018; Kulke et al., 2019; Schuwerk et al., 2018) and only rarely above-chance performance in adults (Burnside et al., 2018). Even a recent direct replication attempt by the original authors, using the original age group, stimuli, and measures, only found chance performance in the FB1 and FB2 conditions (Kampis et al., 2021). This led the authors to conclude that the paradigm and stimuli do not reliably elicit (belief-based) action anticipation. While Kampis et al. (2021) argue that a failure to replicate the original findings does not mean that infants do not understand false beliefs, they still emphasize that, given their unsuccessful replication of the Southgate et al. (2007) paradigm, "an important piece of evidence that has to date supported the existence of early competence of mental-state attribution should no longer serve this function" (p. 8).

Against this background, the original studies' authors were invited to comment on the puzzling picture of failed and partial replications: Regarding violation-of-expectation paradigms, Baillargeon et al. (2018) argued that methodological differences largely account for the negative findings. Similarly, unsuccessful replication attempts of the interactive helping paradigm were attributed to differences in the set-up, materials, procedure, sample size, sample population as well as interference by other tasks administered in the same session (Baillargeon et al., 2018). Regarding anticipatory-looking studies, Baillargeon et al. (2018) admit that high exclusion rates resulting from participants' failure to anticipate correctly in the second familiarization trial (Grosse Wiesmann et al., 2018; Kulke, Reiß et al., 2018; Kulke, von Duhn et al., 2018; Schuwerk et al., 2018) indicate that the task does not consistently elicit action prediction. Moreover, Baillargeon et al. (2018) question whether test trial performance can even be interpreted if only half or even less than half of the sample meets the inclusion criterion. The finding that participants often only performed above-chance in the FB1 condition could mean that they followed a last-location strategy. Baillargeon et al. (2018), however, argue that such a strategy should also lead to better performance in the familiarization trials and consistently below-chance performance in the FB2 trials. Instead, the FB2 condition might be more challenging in terms of working memory and attention demands (Baillargeon et al.,

2018; Grosse Wiesmann et al., 2018; Senju et al., 2010), turning it into an unreliable control condition.

Since evidence of infants' and young toddlers' implicit false belief understanding has been falling away, novel implicit paradigms are needed to address some of the methodological issues existing paradigms face and to reliably assess the nature of infants' early mental state reasoning abilities. The anticipatory-looking paradigm created by Grosse Wiesmann et al. (2017) is a promising such paradigm. It uses a removal version with FB1 and FB2 trials (see e.g., Southgate et al., 2007) and analogous TB1 and TB2 trials, in which the agent holds a true belief. Moreover, it employs a multi-trial design to avoid high exclusion rates due to participants' failure to anticipate and includes several familiarization trials to facilitate inference of the agent's goal. The content of the trials is based upon the paradigm by Surian and Geraci (2012) in that two animated agents chase each other and one of the two hides from the other in a y-shaped tunnel. Using this paradigm, Grosse Wiesmann et al. (2017) found above-chance performance in a combined measure of FB1 and FB2 trials in three- and four-year-olds. There is a great need to replicate this paradigm to corroborate these findings and find out whether the paradigm is also sensitive to younger children's false belief understanding. Moreover, investigating interrelations among different aspects of children's early Theory of Mind abilities helps to empirically evaluate the different theories formulated on implicit Theory of Mind abilities. Relations between implicit and explicit Theory of Mind abilities, for instance, have rarely been examined, and observing such relations would provide some support for conceptual-continuity and developmental enrichment theories.

1.3 Language Acquisition

The following sections first briefly point out the importance of language for acquiring a Theory of Mind and then describe the emergence of mental state terms in toddlers' vocabulary, their acquisition of complement syntax, and some advances in their conversational and discourse skills. Then, Sections 1.4 and 1.5 go into more detail on why each of these aspects of language should be relevant for developing mental state reasoning abilities.

1.3.1 The Role of Language for Theory of Mind

Theory of Mind research on typically developing children (Astington & Jenkins, 1999; Ruffman et al., 2003; Slade & Ruffman, 2005), children with ASC (Happé, 1995; Tager-Flusberg, 2000), and deaf children born to hearing parents (Meristo et al., 2007; Peterson & Siegal, 2000) suggests that language plays a vital role in children's acquisition of mental state

reasoning abilities. A meta-analysis by Milligan et al. (2007) corroborated that three- to six-year-old typically developing children's language skills are related to their false belief understanding, even when controlling for the effects of age. A longitudinal study that assessed language and Theory of Mind at three time points found that language was not significantly predicted by earlier Theory of Mind abilities when controlling for earlier language. However, early language competencies explained unique variance in later Theory of Mind when controlling for earlier Theory of Mind (Astington & Jenkins, 1999). In contrast, a longitudinal study by Slade and Ruffman (2005) detected a bi-directional relation between language and Theory of Mind. The above-mentioned meta-analysis on the relation between false belief understanding and language discovered that while both predictive relations were significant, the effect of earlier language abilities on later false belief understanding was stronger than the reverse relation (Milligan et al., 2007). In line with this finding, significant predictive relations between earlier language skills and later false belief understanding have been found across a time span of two years (Farrar & Maag, 2002; Watson et al., 2001).

There continues to be a hot debate about which aspect of language is most relevant for children's acquisition of a Theory of Mind. Sections 1.4 and 1.5 will present three different theories proposing causal roles of complement syntax, mental state vocabulary, or perspective-shifting discourse for children's Theory of Mind. Mental state terms are assumed to be essential for Theory of Mind development because mental state terms are necessary to express the content of mental states (see Section 1.5.1). However, mental state verbs often require complement syntax constructions to express the content of a mental state. Moreover, complement sentences can express an incorrect statement without making the entire sentence incorrect if they are embedded under certain verbs (see Section 1.4.1). In perspective-shifting discourse – which might use both mental terms and complement sentences – children are repeatedly confronted with differences in people's knowledge and belief states, which might foster their comprehension of such states (see Section 1.5.3). Following the description of each theory, empirical evidence for each theory will be discussed. Before that, the upcoming sections (Sections 1.3.2, 1.3.3, and 1.3.4) provide some background on how far young children's language abilities are developed in each of these areas by age three to four.

1.3.2 Acquisition of Mental State Vocabulary

Children produce their first words around their first birthday (Bockmann et al., 2020; Kauschke & Hofmeister, 2002) and tend to achieve the 50-words benchmark by 18 months on

average (Kauschke & Hofmeister, 2002). Children's first words mainly refer to what children experience in their daily lives and to everything that is hearable, seeable, touchable, and manipulable (Kauschke, 2000, 2012). They initially produce basic-level terms, terms for actions, places, events, and routines. Then, object properties, times, places, and later abstract terms such as *work* or *life* follow (Kauschke, 2000). German children start to talk about mental states at around 18 months, and at 30 months, 90% of all children use mental state terms (Kauschke, 2012).

The majority of research on children's acquisition of mental state terms has been conducted on English-speaking children. Mental state terms, or internal state terms, are words used to explicitly refer to mental states (Beeghly et al., 1986). Historically, Bretherton and colleagues (1981) were the first to systematically assess infants' usage of internal state terms and reported that a majority of 20-month-olds already labeled internal states. The same authors conducted a literature search for studies documenting young children's language production. This search yielded that some internal state terms (e.g., *want*, *hungry*) can be frequently observed in young children's speech but are mainly applied to refer to themselves rather than to others. Moreover, children below age three already questioned others about their internal states, revealing their awareness that internal states cannot always easily be inferred (Bretherton et al., 1981). Systematic analyses of natural speech data from two- to five-year-old children demonstrated that mental state terms emerge in children's spontaneous speech late in the second year of life, with an increasing frequency of occurrence throughout the third year of life (e.g., Bartsch & Wellman, 1995; Bretherton & Beeghly, 1982; Shatz et al., 1983). Already at 28 months, "children interpret their own and other people's mental states, comment on their own or someone else's expected and past experiences and discuss how their own or someone else's [mental] state might be changed" (Bretherton & Beeghly, 1982, p. 919).

Certain aspects of children's acquisition of mental state language have been observed across numerous empirical studies: Children's frequency and variety of mental state talk increase with age (Bretherton & Beeghly, 1982; Beeghly et al., 1986; Furrow et al., 1992; Moore et al., 1994), and children initially talk about desires and emotions and later increasingly refer to cognition (Bartsch & Wellman, 1995; Pascual et al., 2008; Ruffman et al., 2002; Shatz et al., 1983; Taumoepeau & Ruffman, 2006, 2008). Before 2.5 years, children's talk about desires dominates; later, their use of belief terms increases until around the fourth birthday, when

talk about beliefs dominates (Bartsch & Wellman, 1995). Most of children's desire utterances include the term *want*, and children mainly use *think* or *know* to refer to cognitive states (Bartsch & Wellman, 1995; Bretherton & Beeghly, 1982; Furrow et al., 1992; Moore et al., 1994; Pascual et al., 2008 for Spanish children). Individual differences in children's total usage of mental states terms and usage of mental state terms from specific categories (e.g., emotion, moral, cognition) are highly consistent over time (Bretherton & Beeghly, 1982; Ruffman et al., 2002; Taumoepeau & Ruffman, 2006, 2008). Moreover, children's acquisition and usage of mental state language are strongly influenced by parental mental state language input. Thus, numerous studies revealed that mothers' mental state talk is not only related to but also predictive of children's later mental state talk (Beeghly et al., 1986; Furrow et al., 1992; Moore et al., 1994; Ruffman et al., 2002; Taumoepeau & Ruffman, 2006, 2008).

Regarding the order in which children acquire different mental state categories, children at 13 months mostly use perception and volition terms (Beeghly et al., 1986). At 20 and 28 months, children mainly use volitional, physiological, and perception terms, followed by affect and obligation terms, while cognition terms lag behind (Bretherton & Beeghly, 1982). A longitudinal study by Taumoepeau and Ruffman (2006, 2008) found that between 15 and 24 months, children mainly increased their usage of emotion and desire terms, while between 24 and 33 months, also the amount of *think*/*know* utterances and modulations of assertions grew. By the age of 40 to 47 months, terms for *think*, *know*, *pretend*, and *memory* were the most common ones used (Brown et al., 1996). Second-born children at this age also talk more often about mental states with their friends and siblings than with their mothers, even though their mothers talk more to them about mental states than their friends or siblings do (Brown et al., 1996).

Cognitive terms (e.g., *think* and *know*) can already be found in two-year-olds' spontaneous speech. However, early research on children's mental state vocabulary has emphasized that children initially use cognitive terms mainly for conversational uses and that genuine references to thoughts and beliefs only appear shortly before the third birthday (Bartsch & Wellman, 1995; Furrow et al., 1992; Shatz et al., 1983). Montgomery (2002), however, argues that it is difficult to distinguish between children's truly referential uses of mental state terms and more conversational uses because it is unclear whether children also distinguish between these different uses. Therefore, it seems fruitful to also analyze pragmatic or conversational uses of mental state terms. Harris and colleagues (2017c) re-analyzed mental state utterances

previously discarded as conversational uses in the context of the surrounding conversation. In doing so, they found that two-year-olds already referred to their own and, to a lesser extent, also to others' knowledge states spontaneously and in a manner connected to the ongoing conversation. Children affirmed and denied their own knowledge and affirmed and queried their interlocutor's knowledge (Harris et al., 2017c). However, only between ages three and four did children explicitly talk about contrasts between mental states and reality or between different mental states held by different people (Bartsch & Wellman, 1995; Shatz et al., 1983).

Research on German children's mental state language acquisition has corroborated many of the results obtained for English-speaking children. At 18 months, German children mainly refer to desires and emotions. Around 23 months, children talk about emotions and physiology, and at 30 months, they also start talking about volition, ability, obligation, and to a lesser extent about cognitions (Kauschke, 2012; Kauschke & Klann-Delius, 1997). Moreover, longitudinal studies have shown that the usage of internal state terms strongly increases between 13 and 36 months (Kauschke, 2012; Kauschke & Klann-Delius, 1997). While children's internal state language mainly consisted of expressive interjections at 13 months, they started using internal adjectives such as *yummy* around 15 months. At 21 months, children used evaluative and modifying terms, verbs with volitional and intentional content, and the noun *fear* (Kauschke, 2012; note, that the author referred to the German expression "Ich habe Angst." which is, however, best translated to English with the phrase "I am afraid/scared."). At 36 months, children had acquired a variety of terms, including nouns for stating feelings and talking about cognitive processes, mental and modal verbs referring to physiological, emotional, ability, volitional, intentional, and cognitive processes, and adjectives for describing emotional states (Kauschke, 2000). A cross-sectional study using a parent questionnaire to assess German children's mental state language at 24, 30, and 36 months found similar patterns of production across the earlier two measurement points with children producing more terms for physiology, perception, and volition than for emotion and morals, and cognition lagging behind all other categories (Kristen et al., 2012). Moreover, mental state language scores improved across the measurement points, and inter-individual differences remained stable over time (Kristen et al., 2012). Cross-linguistic synchrony in German and English children's acquisition of mental state language was corroborated by a study assessing German, Italian, French-Canadian, and English-Canadian 30- to 32-month-olds' mental state language with a parent questionnaire. In all languages, desire talk preceded talk about beliefs and was more varied

than talk about beliefs. Furthermore, the consistency of scores and developmental sequences of the categories was high across the languages (Kristen et al., 2014).

While children's production of mental state terms begins in the second year of life, their comprehension of cognitive terms as expressing different degrees of certainty develops later. In a longitudinal study, children's comprehension of mental verbs at age four was related to their use of desire terms at age two and their use of belief terms at age three (Moore et al., 1994). These findings indicate consistency in the acquisition of mental state language: Children who were talking more about desires at age two, talked more about beliefs at age three, and better understood degrees of speaker certainty at age four (Moore et al., 1994).

In a hidden-object task, three- to six-year-old and eight-year-old children heard two conflicting statements, involving the mental verbs *know*, *think*, and *guess*, about the location of a desired object (Moore et al., 1989). While three-year-olds showed no differentiation among the terms, four-year-olds started comprehending the different degrees of speaker certainty in the *know-think* and *know-guess* contrasts, and this understanding was complete by age five. The distinction between *think-guess* was not mastered at any age (Moore et al., 1989). Using a longitudinal design and the hidden-object task, Kristen-Antonow et al. (2019) found that German children reach above-chance performance in the *know-think* and the *know-guess* contrast by 60 months but above-chance performance in the *think-guess* contrast only by 70 months. Moreover, performance on the *know-think* contrast improved longitudinally from 60 to 70 and from 70 to 94 months, whereas performance on the *know-guess* and *think-guess* contrasts only improved significantly between 70 and 94 months (Kristen-Antonow et al., 2019). Another hidden-object study that assessed comprehension of modal verbs and modal adjuncts found that, again, four-year-olds had some ability to differentiate among the terms (Moore et al., 1990). Overall, high-certainty and low-certainty terms could be distinguished by age four, whereas the evidence for the distinction between low-certainty terms was less clear. In a slightly different paradigm, four- and five-year-olds and first- and third-graders first made assumptions about the location of an object (Johnson & Wellman, 1980). They were then queried about whether they guessed, knew, or remembered the object's location. In different conditions, children either had a present knowledge base about the location, a past knowledge base about the location, or received feedback on their present performance in indicating the correct or incorrect location. The youngest children were largely unaware of the conceptual differences among the verbs *guess*, *know*, and *remember* (Johnson &

Wellman, 1980). Between ages four and five, children's interpretation of the verb *guess* changed such that they no longer denied guessing when they guessed incorrectly. By age five, they also differentiated *guess* from *remember* and *know*. A detailed comprehension of the verb *remember* was only accomplished in school age. Before, *remember* was also used to refer to correct knowledge about present events (Johnson & Wellman, 1980). All in all, children's appropriate use of mental terms precedes their detailed comprehension of differences among the terms' meanings.

1.3.3 Acquisition of Complement Syntax

Combinations of words only emerge in children's spontaneous speech after they have acquired a critical productive vocabulary of more than 50 words at around 1.5 years old (Kauschke, 2012). At 2.5 years old, children have mastered the generalized verb-second position, which is typical in the German language, and around their third birthday, children begin to produce subordinate clauses (Kauschke, 2012). The first complex sentences appear when children's simple sentences become around four words long, and the major types of complex sentences emerge between ages two and four (Bowerman, 1979). Bloom et al. (1989) define complex sentences as "sentences with two verbs that express two propositions" (p. 101). According to Bowerman (1979), "[c]omplex sentences are structures that are built up of simpler sentences through the recursive operations of *co-ordination* and *embedding"* (p. 285, italicizing in original). In a co-ordinated complex sentence, neither sentence syntactically depends on the other sentence. In an embedded complex sentence, one sentence is subordinated to the other sentence and functions, for instance, as a subject or object of the main clause (Bowerman, 1979). The embedding of the subordinate clause into the main clause is also referred to as complementation. Thus, in complementation, "one proposition serves as an argument WITHIN another proposition" (Bloom et al., 1989, p. 102, capitalization in original). Bloom et al. (1989) argue that the acquisition of complementation thus depends upon the child being able to simultaneously hold two propositions in mind, where one proposition can be expressed as a simple sentence and the other expresses the mental attitude towards that proposition. Due to this feature, many mental verbs require complement syntax constructions to express the content of a mental state (see also Section 1.3.1).

At first, children acquire object complementation using complement-taking verbs such as *want*, *see*, *look*, *like*, *need*, *ask*, and *say*. In the third year of life, the complement-taking verbs *think*, *know*, *hope*, and *show* also become common (Bowerman, 1979). Thus, by their

third birthday, children can already produce complement clauses with a variety of different complement-taking verbs (Brandt et al., 2010). Three types of complement clauses can be distinguished: S-complements such as “I think (that) the sun will shine tomorrow”, wh-complements such as “You don’t know where the chocolate is”, and if-complements such as “We asked if they want soup for dinner”. Each type occurs with only a limited number of complement-taking verbs (Diessel & Tomasello, 2001).

Complement clauses serve two primary functions, namely the assertive and performative use. In the assertive use, the main clause expresses the main proposition (e.g., “I think” or “He said”), and the complement clause adds a piece of information to the main clause (e.g., the content of the thought or communicative act). In the performative use, however, the main clause only refers to a specific aspect of the interaction between the two communication partners (Diessel & Tomasello, 2001). This use is based on the fact that utterances can not only describe something but can also perform an action themselves. This latter case can always be expressed in a performative complement construction such as “I suggest that we leave before it begins to rain.” (Diessel & Tomasello, 2001, p. 103) or “I promise that I will help you with this work.” (Diessel & Tomasello, 2001, p. 103). Thus, the complement construction is used to perform an action such as making a promise or providing a suggestion. However, such cases can always be replaced with a simple construction by leaving out the content of the main clause (e.g., “We better leave before it begins to rain.”, Diessel & Tomasello, 2001, p. 103). In contrast, in the assertive case, the main clause contains information that cannot be left out without changing the proposition. A third use of complement constructions is the formulaic use, where the main clause is primarily used as an attention getter (e.g., “Look what I can do.”) or as an epistemic marker (Diessel & Tomasello, 2001).

Natural language data has been used to study children’s acquisition of complementation. Complement clauses emerge already shortly after children’s second birthday, with S-complements being the earliest complements used (Diessel & Tomasello, 2001). Bloom et al. (1989) observed that complement clauses emerge between ages two and three, and that epistemic verbs were used with complements in almost two-thirds of their appearances, whereas complements accompanied perception verbs in only around one-eighth of all cases. Moreover, the subjects of the main clauses were very limited and did not include talk about third persons (Bloom et al., 1989). Based upon the above-described functions of complement clauses, Diessel and Tomasello (2001) analyzed natural language data from one- to five-year-

old children and observed that most of the early uses of complement clauses were formulaic: Epistemic verbs (e.g., *think*, *guess*) were initially only utilized for formulaic uses, which were then supplemented by the other two functions as children grew older, but assertive uses remained rare even at age five. Deontic markers (*wish*, *hope*) and discourse directives (*see*, *look*, *remember*) only appeared in formulaic uses, with the exception of the verb *see*. Communication verbs (*say*, *tell*, *pretend*) are less abstract. The verbs *say* and *tell* were only used to refer to a verbal activity thus mainly in an assertive function. Interestingly, *pretend* was used in a rather concrete sense to refer to a type of game and mainly served a performative function (e.g., "I am pretending I am a princess!"). Complements with communication verbs appeared a few months after complements with *think*, *know*, and *see*. Futher, communication complements were more likely to occur with a complementizer and were used in more complex and diverse main clauses than complements with any of the other complement-taking verbs (Diessel & Tomasello, 2001).

In contrast to English, in German, the verb can either take a verb-second or a verb-final position in S-complement clauses, depending on whether or not the complement clause begins with a complementizer (e.g., "Ich glaube, der Osterhase **versteckt** morgen Schokolade." or "Ich glaube, *dass* der Osterhase morgen Schokolade **versteckt**."). Brandt et al. (2010) found that German two- to five-year-old children use both verb-second and verb-final complement clauses from an early age with a high number of different complement-taking verbs and seem to learn the construction using a variety of different forms in parallel. A sentence repetition paradigm showed that German three- and four-year-olds are sensitive to the correct word order in complement clauses and often correct complement clauses with an incorrect word order upon repetition (Grosse Wiesmann et al., 2017). Sentence repetition studies can be used to examine children's implicit knowledge of grammar and sentence structures, since three-year-olds have been found to modify sentences with an incorrect word order into a grammatically correct subject-verb structure upon repetition (Abbot-Smith et al., 2001). Sentence repetition paradigms revealed that two- to five-year-old children were more likely to correct an ungrammatical complement sentence commenced by a complement-taking verb appearing with high frequency in naturalistic speech (Kidd et al., 2006). Moreover, four- and six-year-old children were better at recalling a complement structure containing a high-frequency rather than a low-frequency complement-taking verb (Kidd et al., 2010). Thus, children's knowledge about complementation might initially be learned from examples appearing with high

frequency and is only later extended to low-frequency verbs (Kidd et al., 2006). This supports Bloom et al.'s (1989) finding that children seem to acquire the different complement-taking verbs step by step. In contrast, when four- and five-year-olds were asked to repeat complement clauses and thereby change the first-person subject in the main clause into a third-person subject, performance was not influenced by the frequency of the utilized complement-taking verb (Brandt et al., 2011).

Regarding children's comprehension of complement constructions, Brandt et al. (2016) found that four-year-olds better understand third- than first-person complements when they need to extract information from the complements to identify the location of a hidden object. Moreover, de Villiers and de Villiers (2000) argue that children below 3.5 years old still misrepresent and therefore misunderstand the structure of embedded complements such as "what did she say she bought?". Three-year-olds answer such questions with what the person actually bought, and only at age four do children instead report the content of the communicative act (de Villiers, 2005; de Villiers & de Villiers, 2000). Thus, despite the early onset of producing complement constructions in the third year of life, a complete comprehension of complement clauses and usage of complement clauses for assertive purposes only develop somewhat later (de Villiers, 2005; Diessel & Tomasello, 2001).

1.3.4 Acquisition of Conversational Abilities

Children's acquisition of conversational and discourse abilities is part of the development of pragmatic aspects of language. From seven months on, infants vocalize during the speech pauses of their caretakers' talk to them, indicating first signs of turn-taking (Kauschke, 2012; Pan & Snow, 1999; Rohlfing, 2019). However, young children's turn-taking skills only appear sophisticated in the context of conversations with their parents: Children's attempts to enter into conversation with their peers are often unsuccessful (Pan & Snow, 1999), and turn-taking in conversations with peers only starts to work well between ages three and four (Rohlfing, 2019). By nine to ten months, infants begin to produce intentional communication and direct an adult's attention to an object of interest (Pan & Snow, 1999). By their first birthday, infants recognize others' intentions and selectively imitate only others' intentional actions (Carpenter, Akhtar et al., 1998; Kauschke, 2012). Between nine and 12 months, infants' interaction with their interlocutor becomes triadic through the emergence of joint attention skills. Between 12 and 15 months, infants begin to appreciate their interaction partner as an intentional agent with whom they can cooperate to reach goals (Kauschke, 2012). At 18

months, children start to initiate conversations, but only at around 36 months do they adapt their utterances to the current topic of the conversation (Rohlfing, 2019). At 20 months, children only respond to about 33% of the questions and conversational prompts directed at them, but by 29 to 36 months, this proportion increases to more than 50% (Kauschke, 2012; Pan & Snow, 1999). Children continue to show a much lower response rate to their mother's declarative statements, indicating that they understand something about the conversational obligation to answer a question (Pan & Snow, 1999). Between 21 and 36 months, the share of utterances relating to the communication partner's previous utterance also strongly increases (Pan & Snow, 1999).

At about age two, children have acquired a variety of different speech acts (e.g., requesting, stating a proposition, answering questions), but mainly use them in a limited set of communicative interchanges, such as negotiating the immediate activity, directing the hearer's attention, and discussing joint focus (Pan & Snow, 1999; Snow et al., 1996). At 32 months, their repertoire of communicative interchanges also includes discussing the non-present, talking about recent events, and discussing non-observable thoughts and feelings (Snow et al., 1996). The emergence of so-called displaced references – which are utterances directing the hearer's attention to some information that is not perceptible in the interaction partner's environment – was studied by Morford and Goldin-Meadow (1997). By 16 months, children were already referring to currently non-present objects, actions, attributes, or locations. By 21 months, they started talking about proximal (i.e., recent) events (e.g., "Look Mum, I flipped over"). However, only by 31 months did they begin to refer to more distal (i.e., past or future) events and non-actual or hypothetical events. Occurrences of all three types of displaced references increased with age, and the majority of all displaced references were initiated by the children (Morford & Goldin-Meadow, 1997).

Children initially do not understand everything they are told, but two-year-olds already ask for clarification in such cases (Pan & Snow, 1999; Rohlfing, 2019). 20- to 42-month-olds also respond appropriately to parental verbal clarification queries 85% of the time and sometimes even use strategies other than repetition (e.g., revision, elaboration) in response. Thus, children might appreciate that message transmission can fail for more than one reason (Pan & Snow, 1999). In an experimental study, 2.5-year-old children were prompted to request one of two objects (an interesting action object or a boring object) from an experimenter. Fulfillment of the child's request and whether the experimenter expressed comprehension or

misunderstanding in their verbal response to the child were varied. Children provided more repetitions of their request when the experimenter had expressed misunderstanding even if they had received the correct requested object (Shwe & Markman, 1997). This finding shows that young children demonstrate high persistence in clarifying their communicative signals.

Young children also consider their interlocutor's knowledge state when making requests: Two-year-olds conveyed more information (i.e., naming location or gesturing towards location) to their parent when requesting help in retrieving a toy if the parent had been absent while the toy had been hidden (O'Neill, 1996). Moreover, O'Neill (2005) reports on a study in which 22-month-olds were repeatedly presented with the same object, with one feature of the object changing during the last presentation only. Children's vocalizations and behaviors during the presentations demonstrated their understanding "that the first appearance of a toy, or a change in its property, is something worth talking about" (O'Neill, 2005, p. 92).

In total, young children's early production of extended discourse relies on and is supported by "adult willingness to interpret unclear utterances, to pose clarification questions, to supply needed background information, and to resonate to children's evaluations" (Pan & Snow, 1999, p. 246). Basic functions and features of conversations are primarily acquired in structured interactions with parents, such as book-reading situations (Kauschke, 2012).

1.4 The Role of Complement Syntax for Theory of Mind

1.4.1 Linguistic Determinism Theory by Jill de Villiers

Linguistic determinism theory is based upon the temporal co-emergence of mental state talk in children's spontaneous speech and their success on standard verbal false belief tasks (de Villiers & de Villiers, 2003). It attributes a causal role for children's false belief understanding to their acquisition of a specific grammatical structure, the syntax of complementation (de Villiers & de Villiers, 2000, 2003, 2009). Complement clauses appear when a verb does not take merely a noun phrase as an object but an entire embedded proposition called a complement (de Villiers, 2005). The verb under which the complement is embedded is called the matrix verb. Linguistic determinism theory postulates that children who become capable of using the syntax of complementation have acquired a new capacity to represent propositional attitudes (i.e., labels for possible states of mind) (de Villiers & de Villiers, 2000, 2003). Thus, complement syntax used to explicitly express a mental state as in "Anna thinks that zebras are black and white" enables the representational changes necessary to understand something like (false) beliefs. Only the "language of propositional attitudes opens up a classification into

worlds, or different points of view on reality" (de Villiers, 2005, p. 188) by introducing a new point of view – the subject's perspective – into the sentence (de Villiers, 2005).

De Villiers and de Villiers (2009) emphasize some crucial features of complements: Complements appear under verbs of desire (e.g., *want*), communication (e.g., *say*), and mental state (e.g., *think*). They can be false independently of the truth of the matrix sentence; this means, the complement can be false while the whole sentence remains true (e.g., "Anna thought that zebras were purple."). In contrast, "[o]ther types of clauses […] must contain true propositions or derail the truth of the entire sentence" (de Villiers & de Villiers, 2003, p. 350) (e.g., "Anna is crying because zebras are purple."). Moreover, complements allow recursion and are as recursive as propositional attitudes can get (e.g., "Anna believed that Peter thought that Lisa was home."). Taken together, these features render complement clauses specifically helpful for representing others' mental states, particularly their mistakes, lies, and false beliefs, because complement clauses can at the same time "represent the truth in someone else's mind and attribute it only to that person" (de Villiers & de Villiers, 2009, p. 170). This means, complement clauses can express the state of the world in another person's mind and make clear that this is only the thought, belief, or statement of the specific person and not of the speaker (de Villiers & de Villiers, 2009). According to linguistic determinism theory, only once children have acquired the syntax of complementation, can they represent something like a false belief (de Villiers & de Villiers, 2009). Therefore, mastery of complement syntax is required to succeed on false belief tasks.

De Villiers and de Villiers (2000) explicitly describe how children reach false belief understanding via increasingly elaborate complement comprehension: At first, children acquire basic sentences in which the sentence's content always matches reality. This is followed by instances in which children notice discrepancies between what is said and reality. Such discrepancies might arise when people pretend something is the case or make a mistake. The first embedded sentences are then mastered under verbs of communication, desire, or mental state. Around this time, children become familiar with the fundamental syntax necessary for complement clauses but do not yet realize that the truth values of the matrix clause and the complement clause are independent: For instance, if Jack is searching for his favorite food in the fridge and not in the cupboard, where it really is, children will deny the entire sentence "Jack thinks his food is in the fridge" and not just the part "his food is in the fridge". However, through repeated exposure in conversations, children recognize occasions which suggest that

complement clauses can be false when embedded under certain matrix verbs. These first instances are reports of lying or mistakes (e.g., "Jack said the dog is in the kitchen" whereas the dog really is in the bathroom) because "statements are overt and can be compared to reality" (de Villiers & de Villiers, 2000, p. 197). Due to the syntactic analogy between complements under communication verbs, such as *say*, and complements under mental state verbs, such as *think*, children can generalize their knowledge about expressing falsehood in communication complements to mental state complements. Thus, children can then understand a statement as implying a false belief even though thoughts are not overt and must be inferred from actions (e.g., if Jack repeatedly searches for something in the fridge, Jack must **think** his desired object is in the fridge). Having acquired these insights, children can represent possible worlds in other minds even when they deviate from their own world or reality (de Villiers & de Villiers, 2000). Evidence that the ability to fully understand such complement clauses develops roughly around the same age at which children begin to master false belief tasks comes from experiments asking children for their memory of false complements: Children are told stories in which a character is doing one thing (e.g., drinking a cup of coffee) but claiming to have done another thing (e.g., drinking a cup of tea). If three-year-olds are then asked, "What did he say he drank?" they will confidently reply with "coffee" as if they had been asked "What did he drink?" (see de Villiers & de Villiers, 2000, 2009), whereas four-year-olds no longer commit this mistake (de Villiers, 2005). Only after this accomplishment, the complement structure can be considered fully understood (de Villiers, 2005).

1.4.2 Evidence for Linguistic Determinism Theory

A longitudinal study by de Villiers and Pyers (2002) provided supporting evidence for linguistic determinism theory. Three-year-old children were repeatedly assessed with false belief tasks as well as a memory-for-complements task and their spontaneous production of complement sentences was recorded. The memory-for-complements task used a novel task format: In several stories, a character made a mistake, told a lie, or had a false belief which was reported using a complement clause and was contrasted with reality (e.g., "He thought he found his ring, but it was really a bottle cap.", de Villiers & Pyers, 2002, p. 1043). Children were then required to report the content of the mistake, lie, or false belief in response to the question "What did he/she say/think?". Cross-sectionally, stronger relations between measures of complement syntax competence and false belief understanding were observed than between more general language measures and false belief understanding (de Villiers &

Pyers, 2002). Longitudinally, memory-for-complements performance explained unique variance in false belief understanding three months later when accounting for general language abilities. Earlier false belief understanding, in contrast, did not predict memory-for-complements performance three months later (de Villiers & Pyers, 2002). Training studies by Hale and Tager-Flusberg (2003), as well as Lohmann and Tomasello (2003), found that training on complement syntax improved three- and four-year-olds' false belief reasoning, whereas training on relative clauses did not have such an effect (Hale & Tager-Flusberg, 2003). However, training based on false belief tasks (Hale & Tager-Flusberg, 2003) and training consisting of perspective-shifting discourse about deceptive objects, which avoided the use of complement syntax (Lohmann & Tomasello, 2003), also resulted in improved false belief performance. Thus, training on complement syntax seems to not be a necessary condition for fostering false belief understanding.

Cross-linguistic research appears to provide counterevidence to linguistic determinism theory: In German, verbs of desire also require a full tensed complement clause when they refer to a desire for something to happen or for someone to do something ("Ich will, **dass du die Küche aufräumst**"), whereas in English an infinitival complement is used ("I want you **to clean up the kitchen**"). Perner et al. (2003) assessed three- and four-year-olds' false belief understanding and their memory for desire, communication, and belief complements which were discrepant with reality. Children's memory for *want*-complements was significantly better than their memory for *say*- and *think*-complements and their false belief understanding. Perner et al. (2003) thus argue that since German children acquire desire complements much earlier, while still failing false belief tasks, knowledge of the specific grammatical structure of complements cannot be the determining factor for successful false belief reasoning. Studies on Chinese-speaking children have extended these findings: In Cantonese and Mandarin, the same grammatical structure is used to express desires and lies or beliefs without making use of tensed complements. Still, both Cantonese- and Mandarin-speaking children talk about desires well before they talk about beliefs (Tardif & Wellman, 2000), indicating that they master the correct syntactic form for *want* before they master it for *think*. Moreover, Cheung and colleagues (2004) found that Cantonese three- to five-year-olds performed significantly better on *want*- than on *say*-complements in a memory-for-complements task, and that neither performance on *want*- nor on *say*-complements explained unique variance in false belief understanding after controlling for the effect of general language ability. These findings cast doubt

on the universality of linguistic determinism theory. However, de Villiers (2005) argues that complements under *want* are *irrealis* since they relate to hypothetical or future states. Only mastery of *realis* complements, which talk about present or past states, whose truth or falsehood can be judged, should be the determining factor for acquiring false belief understanding. Furthermore, not all languages might use complements, and there might be other aspects in the respective languages which serve the same purpose (de Villiers & de Villiers, 2009). Still, Mo et al. (2014) reported beneficial effects of a sentential complements training on Mandarin-speaking three- and four-year-olds' false belief skills, supporting the causal role of complement syntax for false belief reasoning in a language other than English.

Over the following years, additional controversial evidence accumulated: In a series of three studies, Low (2010) found that three- and four-year-old children's memory for complements explained unique variance in their explicit false belief understanding even when effects of general language, age, nonverbal intelligence, and performance in an implicit false belief task had already been accounted for. Ng et al. (2010) hypothesized that the relation between standard false belief tasks and memory for complements might be driven by a "reality/desire pull" built into both the false belief and complements tasks. In line with this idea, children's false belief performance was related to infinitival rather than tensed complements if infinitival complements tasks contained a strong "reality/desire pull". Furthermore, the relation between false belief and complementation vanished when complements avoided a "reality/desire pull," and the "reality/desire pull" in the false belief tasks was diminished (Ng et al., 2010).

Convergent evidence on the relation between complement syntax and false belief understanding comes from studies assessing deaf children. Deaf children born to hearing parents have no access to communications with their parents in the first few years of life due to having no shared means of communication. As a result, these children are often significantly delayed in their language acquisition, while at the same time having normal nonverbal abilities and a high interest in social activities (de Villiers, 2005; Schick et al., 2007). In contrast, deaf children born to deaf parents are exposed to and acquire a natural language from birth on, and their language acquisition proceeds similarly to hearing children (de Villiers, 2005). Interestingly, deaf children born to deaf parents perform similarly to hearing children on different types of Theory of Mind tasks, whereas deaf children with language delays are significantly delayed in their comprehension of cognitive states but not in their comprehension of desires or intentions (de Villiers, 2005). A large study comparing false belief understanding between three- to

six-year-old hearing children, deaf children with language delays, and deaf children who were native American Sign Language signers revealed that deaf native signers performed equally well as hearing children in verbal and low-verbal false belief tasks. Deaf children born to hearing parents, in contrast, performed significantly worse than both other groups on verbal and low-verbal false belief tasks (de Villiers, 2005; Schick et al., 2007). Moreover, vocabulary and memory for false complements independently predicted false belief performance in both groups of deaf children. Thus, de Villiers (2005) highlights that a mature comprehension of false belief and similar cognitive states depends upon the mastery of complex linguistic structures, since deaf children's problems with complement syntax go hand in hand with their difficulties in reasoning about cognitive states.

Another interesting group to study the relation between language and Theory of Mind is children with ASC. It is well-known that children with ASC fail explicit false belief tasks at a significantly higher rate than matched control groups (e.g., Baron-Cohen et al., 1985). Yet, in studies assessing ASC children's false belief understanding, a minority of participants pass the tasks, contradicting the assumption that ASC children have specific deficits in Theory of Mind (e.g., Baron-Cohen et al., 1985; Happé, 1995). Happé (1995), however, found that success on false belief tasks was closely related to ASC children's verbal ability. Tager-Flusberg and Joseph (2005) argue that it is not general language per se but specifically the acquisition of complement syntax which enables children with ASC to succeed on false belief tasks. This assumption is empirically corroborated by significant links between performance on complements tasks and false belief understanding in children and adolescents with ASC (Durrleman, 2017; Lind & Bowler, 2009; Tager-Flusberg & Joseph, 2005), which even extended to nonverbal assessments of Theory of Mind (Durrleman, 2017). Moreover, a longitudinal study found that competence in extracting the content of false complements predicted Theory of Mind skills both cross-sectionally and longitudinally, after the effects of age and general language had been accounted for (Tager-Flusberg & Joseph, 2005). Thus, individuals with ASC who succeed on explicit Theory of Mind tasks might use a verbally-mediated way to arrive at the correct solution (Lind & Bowler, 2009). In line with this reasoning, a training study with typically developing children, children with ASC, and children with developmental language disorder compared the effectiveness of a complement sentences training to a training focused on lexical enrichment (Durrleman et al., 2019). Participants in the complement sentences training significantly improved their performance on verbal and low-verbal false belief tasks and false

complements tasks independently of group. The lexical training did not result in such improvements (Durrleman et al., 2019). A recent meta-analysis and qualitative review compared the effect sizes of complementation and general language abilities on children's false belief understanding (Farrar et al., 2017). While significant heterogeneity in studies with typically developing children was observed, complement syntax competence was consistently related to false belief understanding in children with ASC. Farrar et al. (2017) thus assume different linguistic pathways to acquiring false belief understanding, with general language skills being more relevant for typically developing children and complement syntax being necessary for children with ASC (Farrar et al., 2017; Tager-Flusberg & Joseph, 2005).

Most of the studies presented so far have used either the memory-for-false-complements task by de Villiers and Pyers (2002) or a similar false complements task by Hale and Tager-Flusberg (2003). However, some researchers have criticized these task formats: They argue that to understand the report of a lie, a mistake, or a false thought, participants already need to have some degree of false belief understanding (Ruffman et al., 2003). According to these researchers, it is not surprising that competence with false complements relates to performance on false belief tasks since both tasks share the demand to represent falsity. So far, the few studies that have tried to disentangle the grammatical properties of complement syntax tasks and the requirement to represent falsity have yielded mixed results. A study by Brandt et al. (2016) applied the hidden-object task (see Moore et al., 1989) and analyzed relations between three- and four-year-olds' comprehension of first- or third-person complements with *think* and *know* and their false belief understanding. Brandt et al. (2016) found that comprehension of third-person complements was related to children's understanding of their own and others' false beliefs. A longitudinal study by De Mulder and colleagues (2019) used a complement syntax task in which two characters talked about different objects using complement sentences. Their statements never involved mistakes, lies, deceptions, or false beliefs. Complement syntax was cross-sectionally associated with Theory of Mind and predicted four-year-old children's Theory of Mind eight months later. However, complement syntax did not explain unique variance in later Theory of Mind when the effects of age and earlier Theory of Mind had been accounted for. De Mulder and colleagues (2019) argue that the lack of a substantial relation might be due to the nature of the task, which avoided expressions of falsity in the complements. A study by Grosse Wiesmann et al. (2017) applied a sentence repetition task to assess complement syntax skills. Three- and four-year-olds heard complement

sentences with or without a complementizer with either a correct or an incorrect word order in the complement clause. As mentioned in Section 1.3.3, in German, a different verb position in the S-complement is required depending on whether the complement starts with a complementizer or not. Children's modifications and corrections of incorrect complement clauses were recorded as a purely syntactic measure of their complement syntax skills. Children's explicit false belief understanding was associated with their complement syntax performance (Grosse Wiesmann et al., 2017). However, the relations were not stronger than those with performance in the syntactic subtest of a standardized language assessment, in which comprehension of causal and temporal clauses, negation, and prepositions was assessed. The authors thus argued that the relation between false belief and complement syntax might rather be driven by common demands of processing complex hierarchical structures than by the specific properties of complement structures (Grosse Wiesmann et al., 2017). Another recent study assessed complement syntax comprehension by having three- to eleven-year-old children point at one of four pictures that best matched a specific complement sentence (Burnel et al., 2021). When controlling for age, comprehension of complement sentences was related to both performance on the Theory of Mind scale (Wellman & Liu, 2004) and nonverbal false belief understanding. In sum, the usage of alternative complement syntax tasks has yielded some evidence of an association with false belief understanding and Theory of Mind more generally. These findings suggest that the relation between complement syntax and false belief might not solely be driven by semantic aspects of complement clauses.

Increasing evidence of false belief understanding before age three to four (see Section 1.2.1) and the breadth of children's mental state reasoning skills in the third year of life (see Section 1.1.3) raise the question of whether complement syntax already supports children's acquisition of Theory of Mind earlier in life. If children understand false beliefs before age three, and if complement syntax skills are required to succeed on false belief tasks, relations between the two concepts should already be observed in the third year of life. A recent longitudinal study assessed two- to three-year-olds' complement clause proficiency using repetition and comprehension tasks (Boeg Thomsen et al., 2021). While children's complement clause proficiency predicted their false belief understanding six months later, the opposite relation was also significant, contrary to the assumptions of linguistic determinism theory. Two-year-olds already possess some knowledge about complement structures (see Section 1.3.3). Thus, a systematic investigation of cross-sectional and longitudinal links between early

complement syntax and early Theory of Mind is required to find out whether complement syntax also affects the emergence of very early Theory of Mind abilities, such as the comprehension of perception or knowledge. Finding such links would show two things: First, children's complement syntax skills are relevant before the standard explicit false belief task is mastered, and second, complement syntax might enable children to understand false beliefs at an earlier age than assumed.

1.5 The Role of Mental State Vocabulary and Perspective-Shifting Discourse for Theory of Mind

1.5.1 The Role of Mental State Vocabulary for Theory of Mind

As outlined in Section 1.3.2, children start to acquire terms for mental states in the third year of life, with increasing occurrences after their third birthday. Olson (1988) argues that this corresponds roughly to the age at which children start to gain insights into the representational nature of mental states, that is, the age at which children acquire a Theory of Mind. He states that children first use language to formulate propositions or requests and to represent them. Later on, children use their acquired metalanguage (i.e., mental state terms) to talk about these propositions and meta-represent them. While children may have acquired such metalanguage earlier, according to Olson (1988), children only start to apply it around age four to refer to their own and others' talk and actions. In sum, Olson (1988) proposes that "it is the acquisition of this metalanguage that […] is central to the development of a theory of mind" (p. 424).

Other researchers, however, remain skeptical as to whether solely the acquisition of mental state terms drives children's development of a representational Theory of Mind since mental state terms and expressions of mental state content are usually embedded in complement syntax constructions. Therefore, it can be argued that children do not yet have a complete comprehension of mental state terms before they are able to use them flexibly within complement constructions (Astington, 2000). Still, it seems plausible that children's exposure to (and usage of) mentalistic comments containing mental state terms at least focuses their attention on mental processes (Harris et al., 2005).

1.5.2 Evidence for Mental State Vocabulary Fostering Theory of Mind

A large strand of research has focused on relating children's production of or productive vocabulary for mental state language to their emerging Theory of Mind. Three- and four-year-olds' mental state talk measured in conversations with friends, siblings, or parents has

been found to be related to children's false belief understanding (Brown et al., 1996; Hughes & Dunn, 1997, 1998), their ability to engage in pretend play (Nielsen & Dissanayake, 2000), their understanding of emotion and deception (Hughes & Dunn, 1997, 1998), their affective perspective-taking (Dunn et al., 1991), and their ability to provide false belief explanations (Dunn et al., 1991). Brown et al. (1996) discovered that three- to four-year-old children's use of contrastives and modulations of assertions was particularly related to their false belief understanding. Moreover, children's mental state talk at age three to four was positively associated with performance in the strange stories task, mental state talk, and the amount of prosocial behavior at age ten (Carr et al., 2018).

Other researchers have assessed children's productive vocabulary for mental state terms with a parent questionnaire: Chiarella et al. (2013) found that two- and three-year-olds' mental state vocabulary was related to their desire understanding as well as their emotional and visual perspective-taking. Olineck and Poulin-Dubois (2007) observed that children's mental state vocabulary at 32 months was associated with performance on a standard false belief task and on the Theory of Mind scale 1.5 years later. Similarly, Brooks and Meltzoff (2015) discovered that children's mental state vocabulary at 2.5 years predicted their Theory of Mind and, in particular, their false belief understanding two years later. The most substantial relation was found for children's productive vocabulary for cognition terms.

A different strand of research has focused on relating children's comprehension of mental state terms to their Theory of Mind. Using a version of the hidden-object task, four-year-olds' false belief understanding was associated with their ability to comprehend and distinguish *know* and *think* as well as *must* and *might* as expressing different levels of speaker certainty (Moore et al., 1990). De Mulder (2015) also discovered that four-year-olds' comprehension of mental state terms was predictive of their false belief understanding eight months later, with the inverse relation being significant as well. This bi-directional relation might indicate that a common conceptual development underlies the acquisition of false belief understanding and the comprehension of mental state terms (Moore et al., 1990). According to a study by Cheung et al. (2009), children's comprehension of non-factive mental verbs, in particular, is related to their false belief understanding. San Juan and Astington (2017) conducted an epistemic verb training in which two- to four-year-olds watched videos of an agent sorting objects by object type. In half of the cases, the agent witnessed the deceptive nature of a target object (true belief situation), whereas, in the other half, the agent did not (false belief

situation). The agent's sorting action was either described neutrally, described using a familiar epistemic verb, or described using a novel epistemic verb in the false belief and a familiar epistemic verb in the true belief cases. Training did not affect children's performance on the explicit false belief tasks. However, children in both epistemic verb conditions improved their implicit false belief understanding (measured by their accuracy of first fixations; San Juan & Astington, 2017). Thus, hearing descriptions of a belief-based action that contained epistemic verbs influenced children's processing of beliefs. In sum, research shows that children's production and comprehension of mental state vocabulary foster their Theory of Mind development. However, evidence from training studies that demonstrate how increased exposure to mental state language improves children's mental state reasoning abilities is largely missing, particularly in children below age three.

1.5.3 The Role of Perspective-Shifting Discourse for Theory of Mind

Harris (1996, 1999) put forward a strong claim about the role of conversational experiences for children's Theory of Mind development. He argues that children are in general involved in two different types of activity with their parents – planful actions and communication through language (Harris, 1996). These activities require different types of understanding of others. To engage in planful actions with their parents, children only need to consider their parents as planful agents who want to and try to achieve certain goals. Thus, children need to understand their parents' intentions and desires. However, to engage in meaningful conversations with their parents, children need to consider their parents as epistemic subjects with whom information can be exchanged (Harris, 1996). Between 18 months – when children have learned to flexibly engage in joint attention and have started to participate in communication about external objects – and 36 months – when children start to understand the mental states of knowledge and belief – children's "conception of, and capacity for, communication undergoes a gradual revolution" (Harris, 1999, p. 101): During this time, conversation becomes increasingly disconnected from current activity, and children learn that conversation can constitute an activity of its own. Due to children's increasing competence and repeated engagement in conversation, they understand others as epistemic agents who can receive and provide information (Harris, 1996). Particularly, the emergence of conversations about displaced topics (see Section 1.3.4) – in other words, about past, future, or hypothetical events or currently not observable objects – which has no instrumental purpose exposes children to situations in which they and their parent differ regarding their knowledge and belief status

(Harris, 1996, 1999): The parent, for instance, knows that the child's friend's birthday party last weekend lasted another two hours after they left it. The child, in contrast, only knows that the party lasted until 6 PM when they personally left the party. Conversations about displaced topics can be considered a prerequisite for explicit talk about knowing and thinking. To engage in such conversations, children must learn to put their point of view aside at times and to shift between their own perspective and their conversation partner's perspective (Harris, 1996).

After their third birthday, children start to understand that others might also acquire misleading or partial information. According to Harris (1999), this development builds upon the earlier emerging skill to engage in information-exchanging conversations, which repeatedly confront children with situations in which knowledge and beliefs differ. Thus, engagement in discourse in which different points of view are presented and discussed facilitates children's Theory of Mind development. Consequently, variations in mothers' pragmatic intention to introduce different perspectives into conversations with their children resulted in performance differences in children's Theory of Mind (Harris, 2005). Furthermore, parents' sensitivity to different individual perspectives fosters such sensitivity in their children as well (Harris et al., 2005). Moreover, Dunn and Brophy (2005) assume that also children's participation in dyadic relationships with peers and engagement in connected, smooth, successful conversations within these relationships positively affects their Theory of Mind development.

1.5.4 Evidence for Perspective-Shifting Discourse Fostering Theory of Mind

A review by Symons (2004), which combines evidence from cross-sectional, longitudinal, and training studies using a variety of different methodologies, contends that mental state discourse is relevant for children's social understanding. A recent study, for instance, not only found links between parents' cognitive talk and their three- and four-year-old children's cognitive talk during a joint puzzle-solving activity, but also that the occurrence of a conflict of perspectives during the activity mediated the relation between parents' self-referent cognitive talk and their child's cognitive talk (McLoughlin et al., 2020). In general, parental mental state talk is one of the most relevant aspects of children's engagement in conversations and will be discussed separately in Section 1.5.5.

Apart from the beneficial effects of parental mental state talk, family size has also been found to be associated with children's Theory of Mind: Having a higher number of siblings is related to better performance on Theory of Mind tasks in three-, four-, and five-year-olds (Jenkins & Astington, 1996; Perner et al., 1994). A meta-analysis yielded a modest effect of

the number of siblings on three- to seven-year-old children's false belief understanding which remained when controlling for children's verbal ability and was independent of whether all or only older siblings were considered (Devine & Hughes, 2018). Another aspect of children's family environment that shapes their conversational experiences is parents' mind-mindedness. Meins and Fernyhough (1999) found that maternal mind-mindedness when children were three years old predicted children's false belief and emotion understanding two years later. The aforementioned meta-analysis also reported a modest, significant association between parents' mind-mindedness and three- to five-year-olds' false belief understanding (Devine & Hughes, 2018). The way mothers react to certain behaviors of their children, such as their misconduct, also affects children's Theory of Mind. Ruffman et al. (1999) administered a questionnaire to mothers asking them to respond to five everyday situations in which they need to discipline their child somehow. Mothers' tendency to refer to the victim's emotional perspective when disciplining their child was related to children's belief understanding. In a similar vein, Peterson and Slaughter (2003) looked at mothers' speech and actions towards their children in a broader range of situations (e.g., forgetting, preparing a surprise, pretending). They offered four prototypical reactions that either referred to an underlying mental state or did not and were either elaborated or not. Mothers' preference to elaborate upon mentalistic themes was associated with four- and five-year-olds' false belief understanding but not their emotion understanding. However, this finding may be due to the questionnaire tapping mainly epistemic states (Peterson & Slaughter, 2003).

The previously reported studies assessed the effects of certain aspects of children's conversational environments on their Theory of Mind. However, the effects of perspective-shifting discussions have also been investigated within training studies, in which one subgroup of children was exposed to a heightened amount of mental state discussions to analyze their causal influence on children's Theory of Mind. In the training study by Lohmann and Tomasello (2003), for instance, the discourse-only condition, in which the deceptive and real nature of objects were merely discussed without mental state terms or complement syntax, also improved three- and four-year-old children's Theory of Mind. A more recent training study by Sellabona et al. (2013) was based upon the study by Lohmann and Tomasello (2003): In all training conditions, except the sentential complements training, deceptive objects and their real and apparent identity were presented. A discourse-only training, a sentential complements training, a no-language training, and a labeling training, in which children only observed

how the apparent and real identity of the object were labeled without being involved in a discussion, were compared. While three-year-old children in the discourse condition improved their performance on a representational-change task from before to after the training, the labeling group outperformed all other groups on various false belief tasks after the training (Sellabona et al., 2013). The authors argued that focusing on the double perspective of deceptive objects was helpful for children's comprehension of false beliefs. Thus, the labeling condition might have been even more helpful than the discourse condition, since it better focused children's attention to this most vital aspect of the training (Sellabona et al., 2013).

Several training studies explicitly modified children's exposure to conversations and discussions about mental states. Three- to five-year-old children repeatedly listened to storybooks enriched with mental state talk and mental state scenarios and were engaged in subsequent discussions of the story characters' mental states or in games aimed at eliciting mental state discussions. Training groups improved their subsequent emotion (Grazzani & Ornaghi, 2011) as well as their false belief and deception understanding (Guajardo & Watson, 2002; Tompkins, 2015) compared to control groups that either did not undergo training or merely listened to the stories without subsequent discussions. Two recent training studies focused on two-year-olds who were exposed to storybooks covering a range of different mental states or emotions (Grazzani, Ornaghi, & Brockmeier, 2016; Grazzani, Ornaghi, Agliati et al., 2016). After the stories, the children in the training groups participated in discussions about the presented mental states or emotions, whereas the children in the control groups engaged in free play or conversations about the characters' physical properties. The training conditions resulted in greater improvements in emotion comprehension, mental state language, desire, and true belief understanding than in the control groups. Moreover, in a training study by Gola (2012), several groups of three- and four-year-olds' watched videos in which characters were surprised after the actual identity of a misleading object had been revealed and in which multiple viewpoints were discussed. Mental verb utterances either addressed the character's mental state in the first person or the second/third person, as a statement or as a question, and in an interactive form directed towards the participants or not. Only children who overheard statements or questions about another person's mental state (i.e., second/third person statement or question conditions) improved their false belief understanding (Gola, 2012). Thus, hearing others explicitly discuss someone else's perspective bolstered children's mental state reasoning, whereas solely being presented with different viewpoints was less important.

1.5.5 Evidence for Parental Mental State Talk Fostering Theory of Mind

Supporting evidence for facilitative effects of both mental state language and conversational input from the environment stems from studies assessing the effects of parental mental state talk on children's developing Theory of Mind abilities. Two recent meta-analyses found a small to modest, significant relation between parental mental state talk and three- to seven-year-old children's false belief and emotion understanding (Devine & Hughes, 2018; Tompkins et al., 2018). The relation was independent of the context used to elicit parental mental state talk and of children's verbal ability.

Cross-sectional as well as longitudinal studies have uncovered positive relations between both the frequency and the variety of maternal mental state talk and two- to six-year-old children's belief (Adrián et al., 2005, 2007; Ensor et al., 2014; Ruffman et al., 2002; Slaughter et al., 2007), desire (Ruffman et al., 2002), and emotion understanding (Beeghly et al., 1986; Taumoepeau & Ruffman, 2006, 2008). Two recent longitudinal studies even found that maternal mental state talk (particularly mothers' cognition talk) when children were two or three years old was related to children's mental state language and their performance in the strange stories task at age ten (Carr et al., 2018; Ensor et al., 2014). Moreover, maternal mental state talk when children were three to four years old was associated with a higher amount of prosocial behavior and a lower amount of externalizing problem behavior when children were ten years old. This suggests that mothers' mental state talk serves as a protective factor even years later (Carr et al., 2018). Some studies found that particularly maternal cognition or "think-know" talk fosters children's Theory of Mind abilities (Ensor et al., 2014; Roby & Scott, 2018; Ruffman et al., 2002; Slaughter et al., 2007). In line with these findings, one of the meta-analyses also found that mothers' talk about cognition had a stronger effect on children's false belief and emotion understanding than talk about desires and emotions (Tompkins et al., 2018). Some studies also observed stronger effects of elaborating, explanatory, or causal mental state talk compared to merely mentioning mental states (Peterson & Slaughter, 2003; Slaughter et al., 2007). However, this finding was not corroborated in the meta-analysis by Tompkins et al. (2018). Longitudinal studies tracking mothers' mental state talk over a longer time indicate that mothers adapt the content of their mental state talk to the child's age and initially refer more to desires and later talk more about cognitions (Adrián et al., 2007; Beeghly et al., 1986; Ensor et al., 2014; Jenkins et al., 2003; Longobardi et al., 2016). Also, maternal mental state talk is predictive of children's mental state talk, whereas usually neither

children's mental state talk nor their Theory of Mind are associated with mothers' mental state talk (Beeghly et al., 1986; Jenkins et al., 2003; Ruffman et al., 2002; Taumoepeau & Ruffman, 2006, 2008; but in Ensor et al., 2014 both relations were significant). This pattern of findings suggests that maternal mental state talk plays a causal role in children's Theory of Mind development (Adrián et al., 2007; Ruffman et al., 2002; Taumoepeau & Ruffman, 2008).

A study by Howard et al. (2008) looked in more detail at specific ways in which mothers' use of certain mental verbs in conversation with their three- and four-year-old children is related to children's mental state reasoning and found a positive effect of mothers' use of questions and mental verbs referring to others. Welch-Ross (1997) focused on mother-child conversations about past events. While mothers' elaborative statements about past events were positively associated with children's understanding of conflicting mental representations (in false belief and representational-change tasks), mothers' repetitions were negatively related to children's understanding of knowledge. Thus, the amount of detail that mothers provide when talking about past events with their children serves as a source for children to understand different representations of the same event (Welch-Ross, 1997).

To summarize, the effects of conversation-based interventions and parental mental state talk point to a vital role of children's participation in perspective-shifting discourse in acquiring mental state reasoning abilities. However, conversation-based training studies rarely controlled for other aspects of language, such as complement syntax, such that beneficial effects of the training could also be due to exposure to complement syntax. Focusing on the effects of a perspective-shifting training about epistemic states combined with mental state vocabulary in young children could show whether the training effects present in the literature are independent of complement syntax and can also be found for younger children and other types of mental states.

1.6 The Present Study

1.6.1 Motivation and Overarching Research Questions

As outlined in the previous sections, the third year of life is a critical time period in children's linguistic and socio-cognitive development. In recent years, numerous studies have revealed that infants and young toddlers already possess impressive Theory of Mind abilities, as evident in their performance in various versions of violation-of-expectation, anticipatory-looking, and interactive helping paradigms (see e.g., Baillargeon et al., 2010; Scott & Baillargeon, 2017). There is an ongoing debate among early childhood researchers on how to

interpret these novel findings and how to align them with the well-known finding that children only succeed in traditional explicit false belief tasks from age three or four on. This debate has become even more heated as studies not or only partially replicating the original findings of implicit Theory of Mind have accumulated (e.g., Kulke & Rakoczy, 2018; Paulus & Sabbagh, 2018). Paradigms that can reliably and robustly measure the extent of infants' and young toddlers' early Theory of Mind abilities are much needed. Moreover, additional research capturing young toddlers' explicit Theory of Mind abilities and particularly interrelations among aspects of early Theory of Mind in the third year of life could help elucidate how children proceed from an implicit understanding of mental states in infancy to an explicit, representational Theory of Mind at age three to five.

A potential mechanism facilitating this progression from early implicit to later explicit mental state reasoning abilities is language (e.g., San Juan & Astington, 2012). As seen in Section 1.3, around the same time at which children start to develop explicit Theory of Mind abilities, they also make impressive progress in their language abilities in several domains: Their mental state vocabulary becomes richer and starts to include cognition terms, they become more proficient at using the syntax of complementation, allowing to directly contrast two beliefs or a belief with reality, and they develop their conversational abilities, enabling them to engage in more elaborate discourse with their interlocutors. While the contribution of each of these aspects of language to Theory of Mind development in three- to six-year-old children has been intensely researched and corroborated, it remains unclear what role language plays in younger children's acquisition of early Theory of Mind abilities.

Recent research has revealed that already two- and young three-year-olds show sensitivity to complement syntax and that these early complement syntax abilities are related to their Theory of Mind (Boeg Thomsen et al., 2021; Grosse Wiesmann et al., 2017). Thus, it is possible that young children's beginning competencies with complement syntax foster their mental state reasoning abilities earlier than assumed, particularly since complement sentences are also needed and used to express the content of other mental states, such as perception or knowledge.

Moreover, Harris and colleagues' re-analyses of children's naturalistic speech have demonstrated that two-year-olds already readily affirm, deny, and query their own and their interlocutor's knowledge and ask information-seeking questions (Harris et al., 2017c). Two-year-olds also use gestures to express their ignorance and start to explicitly, verbally express

their ignorance some time later (Harris et al., 2017a; Harris et al., 2017b). Thus, children begin to communicate their knowledge and ignorance earlier than thought, which – together with children's engagement in discourse with their caretakers – might foster their mental state reasoning abilities already in their third year of life.

Against this backdrop, this dissertation investigated the following three overarching research questions:

1. What early Theory of Mind abilities do children possess in the third year of life, and how are they interrelated?
2. What role does children's sensitivity to the syntax of complementation play in their acquisition of early Theory of Mind abilities in the third year of life?
3. What role do children's emerging mental state vocabulary and their engagement in perspective-shifting discourse play in their acquisition of early Theory of Mind abilities in the third year of life?

The following section will describe the overarching longitudinal study within which the three studies making up this dissertation were conducted. Afterward, a brief overview of each article and study will be given.

1.6.2 Study Design

Three studies were conducted within an overarching longitudinal study spanning children's third year of life. Children were assessed at 24, 27, 33, and 36 months, and approximately 70% of the families were willing to return for a follow-up assessment when their child was 52 months old. Figure 1 provides an overview of the overall longitudinal study. In addition to the tasks and measures described here and in the articles, further measures were collected at each measurement point.

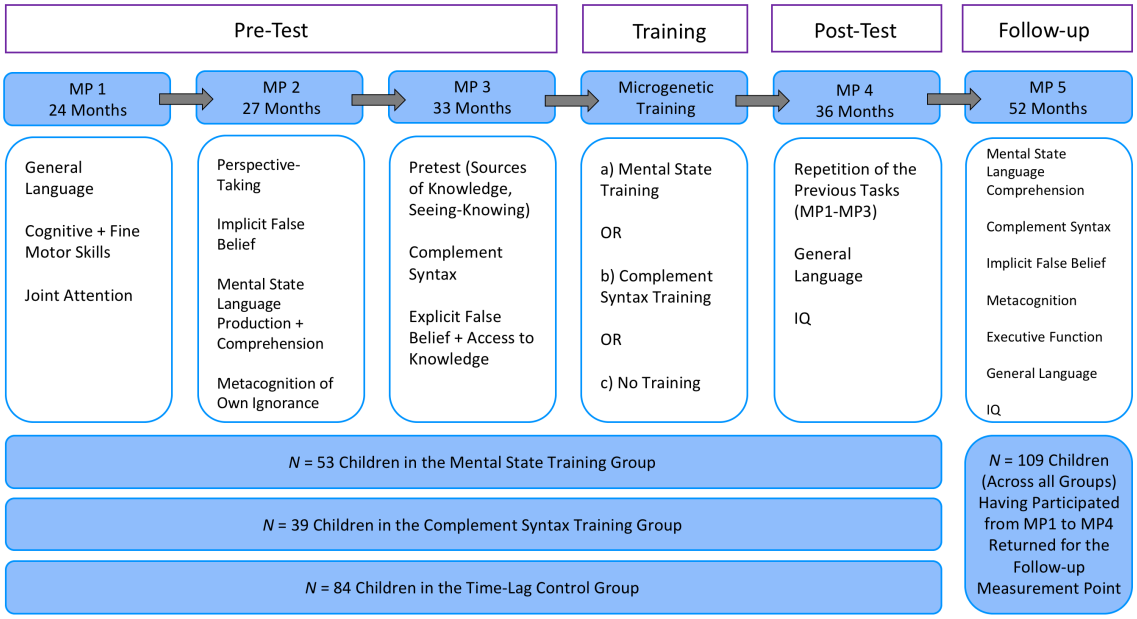
One group of *N* = 84 children only participated at these measurement points and served as a time-lag control group to the two training groups. Approximately one month after the assessment at 33 months, two groups of children began participating in one of two microgenetic interventions. Children participated in ten to 14 training sessions, which were spread over a period of six to seven weeks. One group of *N* = 53 children took part in a mental state training in which children participated in perspective-shifting discourse with two experimenters. These children were repeatedly questioned about their own and other people's knowledge about the content of several containers, and about their sources of knowledge. Moreover, these children had to refute unjustified knowledge claims, and explicitly

differentiate between their own previous and current beliefs about the content of a container. Explicit feedback on children's replies was repeatedly provided and highlighted the relation between informational access and knowledge and utilized numerous instances of the mental verbs *know*, *think*, *guess*, *see*, and *tell*. A second group of *N* = 39 children participated in a comparable training on complement syntax. These children were repeatedly presented with correct and incorrect complement clause constructions and were prompted to repeat and thereby possibly correct them. The same materials as in the mental state training were used.

Children from all three groups were re-invited for the measurement point at 52 months and were assessed with a variety of Theory of Mind and language tasks as well as tasks measuring their general cognitive abilities (e.g., intelligence and executive functions).

Figure 1

Design of the Underlying Longitudinal Study



Note. *N* = 13 of the mental state training children, and *N* = 9 of the complement syntax training children were only invited to participate at MP3, MP4, and in the microgenetic training to compensate for participant loss due to laboratory shutdowns during the first wave of the Covid-19 pandemic.

1.6.3 Overview of Article 1

The first article was conceptualized as a close replication attempt of the anticipatory-looking false belief task by Grosse Wiesmann et al. (2017). As outlined in Section 1.2.3, implicit Theory of Mind research is currently facing a replication crisis, such that reliable and replicable evidence of implicit false belief understanding is needed to evaluate the extent of young children's early Theory of Mind. In Article 1, the anticipatory-looking false belief task by Grosse

Wiesmann et al. (2017) was conducted at 27, 36, and 52 months. Furthermore, general language abilities were assessed at 24, 36, and 52 months and children's explicit false belief understanding was measured at 36 and 52 months. In addition to the direct replication attempt, further analyses of children's progression through the task, their performance in the first false belief trial only, and the consistency between the main effects obtained in the replication study and in the original study were analyzed. The research presented in Article 1 mainly addresses overarching research question 1 by assessing children's early Theory of Mind using an implicit task format. The following research questions were investigated:

1. Can the original finding of above-chance false belief performance in three- and four-year-old children be replicated, and is the paradigm also sensitive towards the implicit tracking of beliefs in children below age three?
2. Does performance in the anticipatory-looking false belief task longitudinally change across the three measurement points?
3. Is performance in the anticipatory-looking false belief task related to children's general language abilities and their explicit false belief understanding?

1.6.4 Overview of Article 2

The second article examined several aspects of children's early mental state reasoning abilities – their visual and epistemic perspective-taking, their metacognition of their own ignorance, and their early explicit false belief understanding – once at 27 or 33 months and once at 36 months. Moreover, children's early sensitivity to complement syntax was assessed at 33 and 36 months. Cross-sectional and longitudinal relations between complement syntax and mental state reasoning abilities were analysed, mainly pursuant to overarching research question 2. Additionally, relations among early mental state reasoning abilities at 36 months were exploratorily examined, contributing to overarching research question 1. Article 2 investigated the following two research questions:

1. Is children's sensitivity to complement syntax at 33 months related to various mental state reasoning skills (visual and epistemic perspective-taking, metacognition, and false belief understanding) three months later?
2. Are children's mental state reasoning skills at 27 and 33 months related to their later sensitivity to complement syntax?

1.6.5 Overview of Article 3

The third article utilized the longitudinal study's underlying training design. In the control group and both training groups, visual and epistemic perspective-taking, metacognition of one's own ignorance, sensitivity to complement syntax, and performance in a near-transfer task assessing comprehension of the seeing-knowing relation were measured before and after the training phase. Moreover, children's joint attention skills at 24 months, productive mental state vocabulary at 27 months, and general language and intelligence at 36 months were assessed as predictors of mental state training-related improvement. Lastly, mental state training children's performance in the individual training sessions was scored similarly to the near-transfer task. The research presented in Article 3 pursued the overarching research questions 2 and 3 and investigated the following research question:

1. Do children in the mental state group improve their performance on the near-transfer task, perspective-taking tasks, and their metacognition of their own ignorance more than children in the complement syntax group and in the time-lag control group?

2 Article 1: Evidence for Goal- and Mixed Evidence for False Belief-Based Action Prediction in 2- to 4-Year-Old Children: A Large-scale Longitudinal Anticipatory Looking Replication Study

Reference

Kaltefleiter, L. J., Schuwerk, T., Grosse Wiesmann, C., Kristen-Antonow, S., Jarvers, I., & Sodian, B. (2022). Evidence for goal- and mixed evidence for false belief-based action prediction in 2- to 4-year-old children: A large-scale longitudinal anticipatory looking replication study. *Developmental Science*, *25*(4), e13224. https://doi.org/10.1111/desc.13224

Abstract

Unsuccessful replication attempts of paradigms assessing children's implicit tracking of false beliefs have instigated the debate on whether or not children have an implicit understanding of false beliefs before the age of four. A novel multi-trial anticipatory looking false belief paradigm yielded evidence of implicit false belief reasoning in 3- to 4-year-old children using a combined score of two false belief conditions (Grosse Wiesmann, C., Friederici, A. D., Singer, T., & Steinbeis, N. [2017]. *Developmental Science*, *20*(5), e12445). The present study is a large-scale replication attempt of this paradigm. The task was administered three times to the same sample of *N* = 185 children at 2, 3, and 4 years of age. Using the original stimuli, we did not replicate the original finding of above-chance belief-congruent looking in a combined score of two false belief conditions in either of the three age groups. Interestingly, the overall pattern of results was comparable to the original study. Post-hoc analyses revealed, however, that children performed above chance in one false belief condition (FB1) and below chance in the other false belief condition (FB2), thus yielding mixed evidence of children's false belief-based action predictions. Similar to the original study, participants' performance did not change with age and was not related to children's general language skills. This study demonstrates the importance of large-scaled replications and adds to the growing number of research questioning the validity and reliability of anticipatory looking false belief paradigms as a robust measure of children's implicit tracking of beliefs.

Larissa Kaltefleiter's contributions to the article: Conducted the study, was responsible for data acquisition, data coding, data pre-processing, and data analysis, and took the lead in writing and revising the manuscript.

3 Article 2: Does Syntax Play a Role in Theory of Mind Development Before the Age of 3 Years?

Reference

Kaltefleiter, L. J., Sodian, B., Kristen-Antonow, S., Grosse Wiesmann, C., & Schuwerk, T. (2021). Does syntax play a role in Theory of Mind development before the age of 3 years? *Infant Behavior and Development*, *64*, 101575. https://doi.org/10.1016/j.infbeh.2021.101575

Abstract

Language plays an important role in Theory of Mind development. Specifically, longitudinal and training studies indicate that the acquisition of complement syntax has an effect on three- to five-year-old children's mastery of the concept of false belief. There is evidence for both a beginning explicit understanding of the mind and mastery of complement syntax in children before their third birthday. In the present study, we investigated longitudinally whether an early sensitivity to complement syntax is related to early development of Theory of Mind abilities in a sample of *N* = 159 German-speaking 27- to 36-month-old children. Children's sensitivity to formal properties of complement syntax at 33 months was associated with their perspective-taking skills and their metacognition of own ignorance three months later. This relation remained significant when controlling for the effects of general language abilities. Furthermore, children's sensitivity to complement syntax was concurrently related to their early false belief understanding. These findings support the view that complement syntax shares representational demands with an understanding of epistemic states and that language begins to support the acquisition of epistemic concepts earlier than was previously thought.

Larissa Kaltefleiter's contributions to the article: Planned and conducted the study, was responsible for data acquisition, data coding and scoring, data analysis, and took the lead in writing and revising the manuscript.

4 Article 3: "I Know What's Inside Because You Showed Me": Training 33-Month-Old Children to Attribute Knowledge and Ignorance to Themselves and Others

Reference

Kaltefleiter, L. J., Sodian, B., Schuwerk, T., & Kristen-Antonow, S. (2022). "I know what's inside because you showed me": Training 33-month-old children to attribute knowledge and ignorance to themselves and others. *Infant Behavior and Development*, *68*, 101744. https://doi.org/10.1016/j.infbeh.2022.101744

Abstract

The present study investigated the role of language in two-year-old children's early understanding of knowledge and ignorance. An intense microgenetic training consisting of 12 to 14 training sessions within six to seven weeks was conducted between 33 and 36 months. One training group experienced and participated in discourse about epistemic states in theoretically relevant situations which highlighted, for instance, the relation between seeing and knowing or contrasts between different people's knowledge states. The other training group was trained on complement syntax using sentence repetition tasks. An age-matched control group received no training. The complement syntax training was not effective in improving complement syntax competence more than in the other two groups. In contrast, the mental state training led to higher improvements in the mental state training group than in the other two groups on tasks assessing comprehension of the targeted concepts (e.g., comprehension of the seeing-knowing relation). The mental state training also had an effect on children's metacognitive awareness of their own ignorance, which was, however, not independent of complement syntax competence assessed at 33 months. No effect was obtained on epistemic perspective-taking skills. Our findings indicate that the use of mental state language in discourse promotes children's acquisition of epistemic concepts even before their third birthday.

Larissa Kaltefleiter's contributions to the article: Planned and conducted the study, was responsible for data acquisition, developed the coding scheme for the near-transfer task and the mental state training sessions, coded, scored, and analyzed the data, took the lead in writing and revising the manuscript.

5 General Discussion

Chapter 5 provides a general discussion of the articles' findings and aims to answer the overarching research questions introduced in Section 1.6.1. Section 5.1 shortly summarizes the articles' main results. Sections 5.2 and 5.3 discuss the findings in relation to the three overarching research questions. Further, these two sections sketch some theoretical implications of the obtained findings and provide practical suggestions for future research on the respective topics. Section 5.4 describes overall limitations of the research presented and gives some global ideas for future research on Theory of Mind in the third year of life. Lastly, the general discussion closes with a conclusion in Section 5.5.

5.1 Summary of the Articles' Main Findings

5.1.1 Summary of Article 1

Article 1 (Kaltefleiter, Schuwerk et al., 2022) aimed to replicate three- and four-year-olds' above-chance belief-congruent anticipatory-looking in a novel multi-trial anticipatory-looking false belief task. The anticipatory-looking false belief task was conducted three times – at 27, 36, and 52 months. *N* = 185 children took part in the study, with *N* = 62 children participating in all three measurement points. Additionally, children's explicit false belief understanding was measured at 36 and 52 months, and a standardized language assessment was administered at 24, 36, and 52 months. The original finding of above-chance belief-congruent looking in a combined score of two false belief conditions, that mutually controlled each other for the application of lower-level strategies, could not be replicated. As in the original study, performance did not change with age and was unrelated to children's general language abilities. A direct comparison of performance in the two false belief conditions (FB1 and FB2) yielded that children performed substantially above-chance in the FB1 condition and below-chance in the FB2 condition at all three measurement points. Moreover, children performed well above chance level in the familiarization trials, requiring only goal-based action predictions. An analysis of performance in only the first false belief trial of each participant resulted in similar findings. Comparing children's looking durations at the target's initial and final hiding location revealed that children looked longer at the final hiding location in the test phase at all three measurement points, independently of the false belief condition. A statistical comparison of the original study's effect sizes with the replication study's effect sizes demonstrated that the original study's findings were not inconsistent with the null effects observed in the replication attempt. Lastly, performance in the FB1 condition was correlated with

children's performance in the explicit false belief tasks, which was not the case for the FB2 condition or the familiarization trials. In sum, Article 1 yielded only mixed evidence of implicit false belief reasoning in two- to four-year-olds.

5.1.2 Summary of Article 2

Article 2 (Kaltefleiter et al., 2021) investigated the role that complement syntax plays for the acquisition of mental state reasoning abilities in the third year of life within a sample of *N* = 159 children. Children were longitudinally assessed at 27, 33, and 36 months. Children's sensitivity to the correct syntax of complementation and their early false belief understanding were measured at 33 and 36 months. Their visual and epistemic perspective-taking skills, as well as their metacognition of their own ignorance, were assessed at 27 and 36 months. One of the false belief tasks used was a novel low-demands false belief task with reduced response-generation and inhibitory demands. Children's sensitivity to complement syntax at 33 months was related to their visual and epistemic perspective-taking skills, their metacognition, but not their false belief understanding three months later. However, at 33 months, children's complement syntax competence was cross-sectionally associated with their false belief understanding. Children's performance in the mental state reasoning tasks at 27 and 33 months did not predict their later complement syntax performance at 33 and 36 months. Together, these results provide first evidence that children's early complement syntax abilities play a causal role in children's acquisition of a Theory of Mind already before age four. Further, children's performance in the visual perspective-taking tasks at 36 months was cross-sectionally related to performance in two false belief tasks, which indicates that children did not solve the low-demands false belief task using lower-level non-mentalistic strategies.

5.1.3 Summary of Article 3

Article 3 (Kaltefleiter, Sodian et al., 2022) reported on a combined longitudinal and training study taking place in children's third year of life. Children either participated in one of two microgenetic interventions between 33 and 36 months old (*N* = 53 children in the mental state training and *N* = 39 children in the complement syntax training) or in a time-lag control group (*N* = 83 children). In the mental state training, children were exposed to perspective-shifting discourse and were asked numerous questions about their own and others' epistemic states and sources of knowledge. In the complement syntax training, children were required to repeat correct and incorrect complement sentences. Children's participation in the complement syntax training neither affected their complement syntax skills nor their reasoning

about epistemic states compared to the other two groups. However, children in the mental state training improved more than the other two groups in a near-transfer seeing-knowing task, focused on the same concepts as taught in the training, and in a metacognition of own ignorance task. Yet, the second effect vanished when controlling for children's pre-training complement syntax skills. Moreover, the only significant positive predictor of mental state training-related improvement was children's productive vocabulary for cognitive particles assessed at 27 months. Analyses of mental state training children's progression through the training revealed that children substantially improved their performance throughout the training and that some of the conveyed concepts were acquired more quickly and more easily than others. In sum, the results indicate that young children's mental state reasoning abilities can be fostered with a linguistic intervention already below age three.

5.2 Discussion of the Findings on Theory of Mind Abilities in the Third Year of Life

This dissertation aimed to enhance our understanding of children's Theory of Mind abilities in the third year of life. Previous research has already shown that children become proficient in reasoning about intentions in the first and second year of life (Carpenter et al., 2005; Phillips & Wellman, 2005) and start to attribute desires and visual perception to themselves and others shortly after the second birthday (Gonzales et al., 2018; Moll & Tomasello, 2006; Wellman & Woolley, 1990). Explicit reasoning about false beliefs and other epistemic states, however, develops between ages three and four (Gonzales et al., 2018; Gopnik & Graf, 1988; Pratt & Bryant, 1990; Wellman et al., 2001).

Over the past two decades, novel task formats which measure children's implicit understanding of false beliefs, using spontaneous-response tasks, have questioned this view (e.g., Baillargeon et al., 2010; Scott & Baillargeon, 2017). Moreover, a new low-demands false belief task produced evidence of explicit false belief reasoning already in two-year-olds (Setoh et al., 2016). Recently, the replication crisis of implicit Theory of Mind (Poulin-Dubois et al., 2018) and criticism towards the low-demands false belief task (Fenici & Garofoli, 2020; Rubio-Fernández et al., 2017) have raised concerns about whether these novel paradigms are reliable and valid measures of early false belief understanding. Article 1 attempted to shed more light on the replicability and validity of implicit false belief tasks. Thus, a promising new multi-trial anticipatory-looking false belief task (Grosse Wiesmann et al., 2017) was conducted trying to replicate the original finding of above-chance belief-congruent looking in three- and four-year-olds. Since three- and four-year-olds already start succeeding in explicit false belief tasks,

the task was also administered to younger children at age two. Moreover, in Articles 1 and 2, relations between performance in the low-demands false belief task and other Theory of Mind tasks were investigated.

Article 1 only partially replicated the original findings of the multi-trial anticipatory-looking false belief task in 27-, 36-, and 52-month-old children. Children performed above chance level in the FB1 condition but below chance level in the FB2 condition at all three ages. Regarding the low-demands false belief task, Article 2 found significant above-chance performance as in the original study (Setoh et al., 2016). Moreover, Article 1 discovered that performance in the FB1 condition tended to be positively related to performance in the low-demands false belief task at 36 months. This was not the case for performance in the FB2 condition or the familiarization trials. In combination with the finding that performance in the FB1 condition was associated with performance in two well-established false belief tasks (Wellman & Liu, 2004) at 52 months, those correlations indicate that both the low-demands and the anticipatory-looking false belief task share representational demands with false belief reasoning. Critics might argue that both the FB1 condition and the low-demands false belief task could be solved by tracking the last location of the target object. However, object-tracking does not allow solving the two explicit false belief tasks from the Theory of Mind scale correctly. Moreover, in Article 2, a sum score of the low-demands false belief task and a false belief picture book task (Hughes & Ensor, 2007), and both false belief tasks separately, were correlated with performance in Level 1 visual perspective-taking at 36 months. Performance in the two false belief tasks and the Level 1 visual perspective-taking tasks also tended to be related to performance in a task assessing children's metacognition of their own ignorance. The present findings suggest cross-sectional interrelations among three-year-olds' abilities to reason about visual and epistemic states. Much of previous research on interrelations among Theory of Mind measures had only focused on relations among different false belief tasks. Other aspects of Theory of Mind had mainly been investigated as precursor abilities to explicit false belief understanding or Theory of Mind around age four (e.g., Brooks & Meltzoff, 2015; Colonnesi et al., 2008; Kloo et al., 2020; Sodian et al., 2016; Yeung et al., 2019). However, the findings from Articles 1 and 2 contribute some evidence that early Theory of Mind skills are interrelated around the third birthday. These findings extend both cross-sectional (Low, 2010) and longitudinal relations (Kloo et al., 2020; Sodian et al., 2016; Thoermer et al., 2012) between implicit and explicit measures of false belief understanding. However, neither Article 1

nor Article 2 investigated longitudinal relations among different Theory of Mind measures, which would provide even more profound evidence of conceptual continuity in children's mental state reasoning. Moreover, Grosse Wiesmann et al. (2017, 2018) did not observe cross-sectional relations between three- and four-year-olds' performance on one implicit and two explicit false belief tasks. Furthermore, Poulin-Dubois (2020) did not find longitudinal relations between performance in a violation-of-expectation and an interactive helping task and later explicit false belief understanding. Also, Warnell and Redcay (2019) recently reported minimal coherence among different explicit Theory of Mind measures in early childhood, and Chiarella et al. (2013) observed no links among visual and emotional perspective-taking as well as desire understanding in 2.5-year-olds. Thus, the evidence of interrelations of Theory of Mind measures in young children is mixed and requires further systematic investigation.

The pattern of findings from the anticipatory-looking false belief task in Article 1 provided no clear evidence of young children's implicit false belief reasoning and could be explained by assuming that children do not consider the agent's belief and just look at the last location they had observed the target object (Baillargeon et al., 2018; Southgate et al., 2007). This is supported by Article 1's finding that children looked longer at the last hiding location of the target object independently of the false belief condition. However, across other replication attempts of anticipatory-looking paradigms, some studies found that children performed at chance level (Grosse Wiesmann et al., 2018; Kulke, Reiß et al., 2018; Schuwerk et al., 2018), whereas other studies found that children performed below chance level in the FB2 condition (Grosse Wiesmann et al., 2018; Kulke, Reiß et al., 2018; Kulke, von Duhn et al., 2018). Moreover, some studies did not even find consistent above-chance performance in the FB1 condition, which would be expected if children just track the last location of the object (e.g., Dörrenberg et al., 2018; Kampis et al., 2021; Kulke et al., 2019). A study including an additional TB1 condition, in which the agent also watched transfer and removal of the target object, found that four- and five-year-olds' first saccades did not indicate that children generally followed a last location strategy, since they showed no preference for either of the two locations in the TB1 condition (Kulke et al., 2019). Yet, children's looking durations revealed a tendency to track the object's location: Children looked longer at the last location of the object not only in the FB1 but also in the TB1 condition, even though the agent had witnessed the object's removal from the last location in the TB1 condition (Kulke et al., 2019).

The overall pattern of findings from replication studies of anticipatory-looking paradigms indicates that children might not follow a last location strategy, but that the FB2 condition is not a reliable control condition because it might invoke higher working memory and inhibitory demands. In FB2 trials, children need to keep in mind where the agent last witnessed the target object while at the same time monitoring the object's final movements. Article 1 also found that children's anticipatory looking times were in general longer in the FB1 condition than in the FB2 condition. Children were possibly more confused by the FB2 condition and may not have formed a clear expectation about where the agent will search in FB2 trials. Thus, they did not anticipatorily gaze at either of the two areas of interest but instead looked at other parts of the screen, which led to shorter anticipatory-looking durations. Consequently, new implicit false belief conditions, which are more comparable in working memory and inhibitory control demands, are needed to investigate the extent of young children's belief-based action anticipations and relations with other aspects of Theory of Mind.

Article 1's findings, combined with previous research, suggest that children might be sensitive to agents' goal- and belief-based actions already before age three, but that their reasoning about belief-based actions is easily hindered by task modifications or heightened task demands (e.g., load on working memory). Wang and Leslie (2016), for instance, found that not removing the target object from the scene prior to the anticipatory phase, and thus increasing the inhibitory demands of the task, aggravated two- and three-year-olds' and even adults' false belief reasoning. Also, Setoh et al. (2016) only found below-chance performance of 2.5-year-olds when using a high-inhibition version of their false belief task. Further, evidence from replication studies of implicit Theory of Mind shows that infants and young children demonstrate false belief reasoning only in a few specific conditions. Slight methodological changes of the original stimuli or procedures often resulted in chance performance in contrast to the original findings. Moreover, children's performance in implicit tasks often does not consistently correlate across different false belief measures and paradigms (Dörrenberg et al., 2018; Kulke, Reiß et al., 2018), while studies on explicit false belief understanding often observe medium correlations (e.g., Gopnik & Astington, 1988; Grosse Wiesmann et al., 2017; Thoermer et al., 2012) or high internal consistency (Astington & Jenkins, 1999; Slade & Ruffman, 2005) among different false belief tasks. Thus, if infants and young children are indeed capable of attributing and understanding false belief, this understanding might be rather

fragile since it only influences behavior under certain circumstances (Crivello & Poulin-Dubois, 2018).

Theoretical Implications

Discovering interrelations among different Theory of Mind measures informs theoretical ideas about how and at which age children acquire mental state reasoning abilities. Sodian et al. (2020) emphasize that cross-sectional and longitudinal relations between early Theory of Mind abilities, or Theory of Mind precursor abilities, and later explicit Theory of Mind, or false belief understanding, support both conceptual-continuity and developmental enrichment theories. Proponents of conceptual-continuity theories argue that infants already possess impressive psychological reasoning abilities (Baillargeon et al., 2016; Scott & Baillargeon, 2017). They assume that children understand false beliefs already before age four, but that external task demands hinder their successful performance (Scott & Baillargeon, 2017). Article 1's finding of above-chance performance in the FB1 condition of the anticipatory-looking task and the assumption that children failed the FB2 condition due to higher inhibitory demands support this account. The finding that children's performance in the anticipatory-looking task did not improve with age also supports conceptual-continuity claims. This finding indicates that implicit false belief reasoning might be present from early on and might not undergo significant changes. Moreover, the observed cross-sectional relations between implicit and explicit false belief reasoning substantiate the idea of conceptual continuity. However, additionally, systematic longitudinal relations between earlier implicit and later explicit false belief understanding would have been expected. Article 2's cross-sectional relations between reasoning about visual perception and false belief align with Ruffman's (2014) idea of developmental enrichment and the idea that implicit false belief tasks might be solved by merely tracking the visual percepts of the protagonist (Perner & Roessler, 2012). According to Ruffman (2014), young children's successful performance on implicit false belief tasks can also be explained by tracking others' visual perception and observing statistical regularities in others' behavior. Thus, these accounts assume that success in the anticipatory-looking and the low-demands false belief task is not due to a representational understanding of mental states. Rather, an early understanding of others' behavior, based upon monitoring visual perception, later develops into a representational Theory of Mind (Ruffman, 2014). The correlations between early false belief understanding and visual perspective-taking in Article 2 could be considered support for this argumentation.

Practical Implications and Suggestions for Future Research

To fully discriminate between the ideas of conceptual-continuity and developmental enrichment, obtaining reliable and replicable evidence of implicit false belief understanding in young children is necessary. Different, newly created implicit false belief conditions should vary contents and external demands of implicit mental state reasoning paradigms to better understand if and how infants arrive at the correct belief-based anticipations. Paradigms that require children's verbal responses might at first seem more suitable to understand their underlying reasoning by having them justify their responses. However, young two-year-olds' language skills are still limited, and several studies showed that children give correct belief-based judgments before they can provide adequate, belief-based justifications of their answers (e.g., Clements & Perner, 1994; Clements et al., 2000; Wimmer & Mayringer, 1998).

In addition to creating novel paradigms and conditions, Wellman (2018) recommends obtaining a more progressive developmental picture of infants' and young children's implicit Theory of Mind abilities. Instead of assessing children at different ages with different types of spontaneous-response paradigms, he suggests to systematically measure children's implicit mental state reasoning abilities at essential time points in development with the same paradigms. This would allow to see how different age groups perform within the same situation under the same external task demands. Thus, developmental differences between infants' and three- or four-year-olds' performance in spontaneous-response paradigms could more easily be inferred and would provide insightful information on whether implicit Theory of Mind abilities undergo substantial developmental changes.

Not finding correct, belief-based action anticipations in some spontaneous-response studies cannot be equated with young children having no false belief understanding, since children who solve explicit false belief tasks correctly, as well as adults, also often perform at or below chance level in these studies (e.g., Grosse Wiesmann et al., 2018; Kulke, von Duhn et al., 2018; Kulke et al., 2019; Schuwerk et al., 2018). Thus, participants' actual predictions might not be reflected in the pattern of their saccades, resulting in poor performance (Kampis et al., 2021). Furthermore, the time window available for anticipations might be too short for participants to arrive at a final anticipation, especially in conditions with higher task demands (Wang & Leslie, 2016). Moreover, the traditionally utilized videos and stimuli might not be engaging enough or motivating enough for participants to form anticipations (Kampis et al., 2021; Kulke et al., 2019). Thus, using more modern stimuli that attract children's interest

seems commendable (Kampis et al., 2021). A first initiative in this direction, which also systematically investigates the reliability and robustness of findings from spontaneous-response Theory of Mind tasks, is the *ManyBabies 2* project (Schuwerk et al., 2021).

A general problem of many spontaneous-response Theory of Mind tasks, and also of the Duplo false belief task (Rubio-Fernández & Geurts, 2013), is that participants need to find out by themselves what the agent wants to or aims to do based on only little familiarization with the situation (Priewasser et al., 2020). This contrasts with most traditional false belief tasks in which children are explicitly told what the agent wants to do. In the most frequently successfully replicated spontaneous-response paradigm by Clements and Perner (1994), children were also told what the agent would do next. This procedure even allowed to measure implicit and explicit false belief understanding within the same paradigm. While one study revealed that adding such verbal explanation of the story events to other anticipatory-looking paradigms did not significantly affect performance (Kulke & Rakoczy, 2019), it still seems advisable to familiarize children better with the situation presented in spontaneous-response tasks. For instance, Grosse Wiesmann et al. (2017) included several familiarization trials in their paradigm to facilitate inference of the agent's goal. However, the familiarization trials were intermixed with the false belief trials. Still, unlike studies finding low correct anticipation rates in the familiarization trials, the paradigm by Grosse Wiesmann et al. (2017) led to substantially above-chance performance in the familiarization trials, both in the original study and in the replication study (see Article 1). Repeated familiarization with the agent's goal before presenting the first false belief-based action seems even more advisable. Thus, the plans of the *ManyBabies 2* project have also included conducting several familiarization trials to better familiarize participants with the agent's goal of approaching the target. First results of the *ManyBabies 2* project showed that both toddlers and adults engaged in spontaneous goal-based action anticipation within the presented familiarization trials, and that exclusion rates due to failure to anticipate were lower than in previous studies (Schuwerk et al., 2021).

In sum, introducing novel spontaneous-response paradigms that allow controlling external demands and assessing early mental state reasoning abilities within different conditions are promising new avenues for research on infant Theory of Mind. To date, infant Theory of Mind research suffers from a lack of standardization of paradigms, task procedures, gaze measures, and criteria. These issues are particularly pronounced for violation-of-expectation paradigms (see Rubio-Fernández, 2019 for criticism and proposals for future violation-of-

expectation studies). Thus, establishing standard criteria (e.g., for the length of the anticipatory phase or end-of-trial criteria in violation-of-expectation studies), improving familiarization procedures, and ensuring that participants are motivated to anticipate actions are vital steps for further advancing research on infant Theory of Mind.

Article 1's results additionally emphasized the practical relevance of well-powered, large-scale studies for replicating theoretically relevant findings, since the substantial performance differences between FB1 and FB2 trials only became evident in the large-scale replication. Article 1 also provided a further analysis which examined the consistency between the effect sizes obtained from the original and the replication study, following a procedure by Anderson and Maxwell (2016). This analysis revealed that, despite being significantly above chance, the original study's effect size was not inconsistent with the null effect obtained in the replication study. Statistically comparing the original and the replication study's effect sizes (based on Anderson & Maxwell, 2016) appears to be a promising additional goal for future replication studies. This method considers the different sample sizes and shows whether the original effect size is, similarly to Article 1, so small relative to the original sample size that it is compatible with a non-significant effect observed in the replication study. Instead of merely talking about successful or failed replications, this approach can also demonstrate whether the effect sizes strongly differ from each other or are comparably large.

5.3 Discussion of the Effects of Language on Early Theory of Mind

This dissertation's second and third aims were to investigate the effects of different aspects of language on children's Theory of Mind abilities in the third year of life. As shown in Sections 1.2 and 5.2, children and infants demonstrate implicit and explicit understanding of false beliefs earlier than assumed. The accumulating evidence of implicit false belief understanding raised the question of how children transition from an early implicit to a later explicit understanding of false beliefs. Some researchers point out that children's developing language abilities play a vital role in this transition (Ruffman, 2014; San Juan & Astington, 2012; Tomasello, 2018).

In 1988, Olson proposed that children's acquisition of mental state language enables their subsequent Theory of Mind development. Several primary studies have traced the emergence of mental state vocabulary back to age 1.5 to two (Bartsch & Wellman, 1995; Bretherton & Beeghly, 1982; Bretherton et al., 1981; Kauschke, 2012; Shatz et al., 1983). However, initially, children mainly refer to physiological states, desires, or perceptions and rarely talk

about cognitive states (Bartsch & Wellman, 1995; Pascual et al., 2008; Ruffman et al., 2002; Shatz et al., 1983; Taumoepeau & Ruffman, 2006, 2008). Moreover, some researchers emphasized that early uses of cognitive terms mainly serve conversational purposes and do not truly refer to underlying cognitive states (Bartsch & Wellman, 1995; Shatz et al., 1983). However, this view has been questioned by Harris and colleagues (2017c). They re-analyzed young children's naturalistic speech and discovered that already two-year-olds frequently refer to their own and others' knowledge. If children refer to cognitive states at age two, their early knowledge about cognitive terms could affect their developing mental state reasoning abilities already in the third year of life. However, various mental state terms require complement sentences to refer to a mental state, such as a perception or belief. Research has shown that already two-year-olds produce complement sentences with a variety of different complement-taking verbs (Bloom et al., 1989; Bowerman, 1979; Brandt et al., 2010; Diessel & Tomasello, 2001), but they mostly use them for conversational purposes. Exceptions to this usage are complements under communication verbs, which are mainly used for assertive purposes from early on (Diessel & Tomasello, 2001). According to linguistic determinism theory (see Section 1.4.1), children's comprehension and production of complement sentences should be related to their understanding of false beliefs (de Villiers & de Villiers, 2009). Linguistic determinism theory is only formulated for children's explicit comprehension of false beliefs. However, the emergence of complement syntax already around the second birthday makes it seem likely that complement syntax skills also play a role in children's reasoning about other, earlier developing mental states, such as perception or knowledge, which begin to be understood in the third year of life. Moreover, the mental verbs *see* and *know*, which are used to talk about one's perception or knowledge, often take complements. Since they are factive verbs, they require the complement sentence to be true. Thus, complement sentences under *see* or *know* cannot express a false proposition without rendering the entire sentence incorrect. If, however, the syntactic aspects of complement syntax drive the link with false belief understanding, children's complement syntax proficiency may also be linked to their comprehension of other mental states, such as perception and knowledge. In previous training studies, the effects of mental state vocabulary and complement syntax on children's Theory of Mind could sometimes hardly be separated, since the training material contained both complement sentences and mental state terms. Thus, some previous training studies did not clearly disentangle the effects of complement syntax from the effects of mental state

terms. Dependent on the study's research question, an effect of the intervention was interpreted as evidence for either the role of complement syntax (Lohmann & Tomasello, 2003) or the role of mental state language (Gola, 2012; San Juan & Astington, 2012). In contrast, in Articles 2 and 3, the effects of complement syntax and mental state vocabulary on Theory of Mind were studied separately.

Article 2 found that children's sensitivity to complement syntax at 33 months was longitudinally related to their visual and epistemic perspective-taking as well as their metacognition of their own ignorance three months later. While the complement syntax training in Article 3 did not have the expected effects on children's subsequent complement syntax or mental state reasoning abilities, Article 3's findings corroborated the significant relation between complement syntax and metacognition. Moreover, Article 3 found that an intense linguistic training, exposing children to perspective-shifting discourse and mental state terms, had a beneficial effect on children's subsequent metacognition and knowledge about the seeing-knowing relation and sources of knowledge. The findings from Articles 2 and 3 point out that, as expected, children's language skills already influence their comprehension of mental states before age three. The results extend the well-documented relation between three-year-old and older children's Theory of Mind and their proficiency with complement syntax and mental state language to children in the third year of life. Thus, Articles 2 and 3 showed that language abilities are already relevant at an age at which, according to traditional findings, children do not have a representational understanding of mental states yet.

A combined training of mental state terms and perspective-shifting discourse ameliorated children's understanding of the seeing-knowing relation and their metacognition, but the effects of both aspects could not be studied separately. However, unlike in other training studies, the usage of complement syntax was avoided as much as possible. Thus, the training effects can be considered caused by exposure to mental state terms and discourse and not by exposure to complement syntax. This result confirms Harris' (2005) argumentation that children's involvement in discourse about different people's mental states facilitates their understanding of mental states. The evidence for the effects of complement syntax on children's mental state reasoning abilities in the third year of life was less strong, since the complement syntax training itself was not successful. However, complement syntax showed longitudinal relations with later mental state reasoning, whereas earlier mental state reasoning did not predict later complement syntax. This pattern of findings indicates a causal role of

complement syntax. Moreover, including complement syntax as a covariate also diminished the mental state training's effect on children's metacognition.

To engage children in perspective-shifting, information-exchanging discourse in daily life, it is almost impossible to avoid the usage of complement syntax or mental state terms, particularly if mental states are discussed. Thus, all three investigated aspects of language might rather be working together to create such conversations. Particularly in the third year of life, developments in one area of language often also result in improvements in another area of language. For instance, children probably acquire many new mental state terms from their occurrences in complement syntax constructions. Thus, they learn not only the new mental state term but also a possible grammatical structure in which the term can be used. Moreover, it appears difficult to imagine, how parents can engage children in perspective-shifting discourse about others' mental states without using mental state terms to name and highlight the discussed mental state and without using complement sentences to contrast and compare different mental states. When discussing their own or others' epistemic states, children need to utilize their mental state vocabulary to refer to an epistemic state, but they can hardly express the content of someone's epistemic state without using complement syntax constructions. Also, children can compare two different epistemic states by using two complement sentences, but it only becomes clear that these sentences refer to epistemic states when introducing them with the appropriate mental state terms. While independent training effects of mental state language, complement syntax, or perspective-shifting conversations have been documented, Lohmann and Tomasello (2003), for instance, reported the most substantial effect on three- and four-year-olds' false belief understanding for an intervention that combined all of these aspects. According to Harris et al. (2005), however, the driving factor for Theory of Mind acquisition are children's experiences in discourse, in which parents or other knowledgeable adults convey and elucidate various mental states, and not specific terms or syntactic constructions. Thus, either the observed effects of mental state vocabulary and complement syntax might be coincidentally due to them frequently occurring in mental state conversations or mental state vocabulary and complement syntax might additionally support children's mental state reasoning abilities.

The following two sections will go into some more detail on the observed separate effects of complement syntax and mental state language combined with perspective-shifting discourse on children's Theory of Mind in the third year of life.

5.3.1 Effects of Complement Syntax on Early Theory of Mind

Article 2 also investigated the relationship between children's early false belief understanding and their complement syntax skills. While complement syntax was not predictive of false belief understanding, complement syntax and early false belief understanding were cross-sectionally related at 33 months. The complement syntax task used in Article 2 was a sentence repetition paradigm based on Grosse Wiesmann et al. (2017), which did not include false statements, lies, or false beliefs in the test sentences, unlike other frequently used complement syntax measures (e.g., de Villiers & Pyers, 2002; Hale & Tager-Flusberg, 2003). This might be one reason why the evidence of a longitudinal link between complement syntax and false belief understanding in Article 2 was weaker than in other studies. De Mulder et al. (2019), for instance, also used a complement syntax measure that did not involve false statements or false thoughts. In their study, four-year-olds' comprehension of complement sentences did not explain unique variance in their false belief understanding eight months later when controlling for children's general language skills.

Also, the complement syntax training conducted in Article 3 required children only to repeat correct and incorrect complement sentences. This rather subtle complement syntax training might not have affected subsequent complement syntax performance because children's complement syntax skills still undergo significant improvements in the third year of life. Thus, the maturational effect in the control group and both training groups may have been stronger than any effect of an intervention.

In Articles 2 and 3, the complement syntax measure only involved complements under communication verbs, and performance on communication complements was related to understanding perception and knowledge. This finding is in line with de Villiers' (2005) suggestion that children initially learn the features of complement syntax from examples with communication verbs and later generalize these insights to complements under different verbs. This argumentation seems plausible since, as Diessel and Tomasello (2001) found, complements with communication verbs are initially the only complements used for assertive purposes. Thus, particularly for young children, competence with communication complements seems most relevant for their comprehension of mental states. However, in Articles 2 and 3, a complement syntax task containing epistemic or perception verbs was not conducted such that the effects of different types of complements on early mental state reasoning abilities could not be investigated. A recent study by Davis and Landau (2021) found that even though

perception verbs are used more frequently than epistemic verbs, children use complements under epistemic verbs as a model for complements under perception verbs. Davis and Landau (2021) argue that early uses of perception verbs only rarely refer to mental state representations. Only through repeated exposure to epistemic complements children learn that perception verbs used in complement sentences can also refer to mental state contents. Therefore, it is possible that children's early competence with epistemic complements could have been equally or even more relevant as competence with communication complements. However, this would have made it impossible to discriminate between the effects of the epistemic verb and the effects of the complement sentences. Thus, epistemic complements may have the strongest effect because they combine two helpful linguistic features.

Theoretical Implications

De Villiers and de Villiers (2009) have proposed that acquiring complement syntax is necessary to understand false beliefs. Indeed, de Villiers and Pyers (2002) reported that each child passing the false belief tasks in their study demonstrated some productive command of complement syntax in at least one of the complement syntax measures. Other findings that have accumulated so far, however, question whether this is the case in typically developing children. For one thing, a relation between false belief understanding and complement syntax was not consistently observed in all studies (e.g., De Mulder et al., 2019; Ng et al., 2010), and some researchers even argue that certain task features cause the link between complement syntax and false belief understanding: Ruffman et al. (2003) emphasized that to solve the memory-for-false-complements task, children need to be capable of representing falsity. Thus, false belief tasks and the memory-for-false complements task share the demand of representing falsity which might drive the link between the two tasks. Results by Ng et al. (2010) also indicated that the complements' semantic content – whether they involved a "reality/desire pull" or not – was causing the relation with false belief understanding rather than the syntactic aspects of complement syntax. For another thing, Setoh et al. (2016) and Article 2 reported above-chance explicit false belief understanding in a low-demands false belief task already in two-year-olds. According to de Villiers (2005), however, comprehension of complement syntax can only be considered complete once children can correctly answer long-distance wh-questions involving a false proposition (e.g., "What did he say what he drank?"). Though, this ability is only accomplished between 3.5 and four years (de Villiers, 2005; de Villiers & de Villiers, 2000, 2009). Suppose children's comprehension of complement syntax is

only complete some time in their fourth year of life. In that case, complete mastery of complement syntax can at least not be a necessary condition for succeeding on simplified versions of false belief tasks. Additionally, Low (2010) and Grosse Wiesmann et al. (2017) reported that three- and four-year-olds' performance in spontaneous-response false belief tasks was unrelated to their complement syntax skills. Thus, some capacity to (implicitly) represent false beliefs seems possible independently of one's ability to comprehend and produce complement sentences.

Practical Implications and Suggestions for Future Research

In Articles 2 and 3, only one complement syntax task was conducted. This task utilized a sentence repetition paradigm to capture young children's early sensitivity to the correct syntax of complementation. However, to identify which aspect of complement syntax is relevant and beneficial for young children's developing Theory of Mind, the effects of different tasks tapping distinct aspects of complement syntax should be compared within the same sample of children. Moreover, the type of complement-taking verb could be systematically varied between test items, and both comprehension and production of complement syntax should be assessed as complementary measures. Possibly, receptive knowledge of complement syntax constructions (as in the complement syntax task by Burnel et al., 2021) might facilitate mental state reasoning even in children who barely produce complement sentences yet. For example, in the complement syntax task used in Articles 2 and 3, some 33-month-old children had been overwhelmed by the overall task demands. Thus, some children only repeated single words from the test sentences. For these children, a receptive measure of their complement syntax skills might have been more sensitive to their knowledge about complement structures than the productive measure used.

Further, as Fontana et al. (2018) emphasize, it is essential to distinguish between the impact of complement syntax on the development of a Theory of Mind and the impact of complement syntax on Theory of Mind when already acquired. Burnel et al. (2021), for instance, found that complement syntax was related to false belief understanding and other aspects of Theory of Mind in three- to eleven-year-old children independently of the effects of age. However, ceiling effects in both the complement syntax and the Theory of Mind measures from six years on restrict this finding. Still, it seems promising to investigate relations between Theory of Mind and complement syntax also in children who have already acquired false belief understanding, using more complex complement syntax and Theory of

Mind measures. This would shed light on the relevant question, whether Theory of Mind and complement syntax are only related temporarily in a restricted period in development (e.g., before the major mental state concepts have been mastered), or whether the usage and expression of Theory of Mind permanently depend on one's complement syntax abilities. Thus, the necessity of complement syntax skills for successful mental state reasoning could be evaluated. Also, children in the third year of life, who are only beginning to acquire mental state reasoning abilities, should be further studied to investigate the role of complement syntax in the emergence of early Theory of Mind abilities. Boeg Thomsen et al. (2021), for instance, found bi-directional longitudinal relations between complement syntax proficiency and false belief understanding in two-year-old children. Thus, the development of false belief understanding and complement syntax might mutually influence each other, since both developments are gradual processes and do not occur at once (Boeg Thomsen et al., 2021).

5.3.2 Effects of Mental State Vocabulary and Perspective-Shifting Discourse on Early Theory of Mind

The mental state training conducted in Article 3 facilitated 36-month-olds' comprehension of the seeing-knowing relation and improved their metacognition of their own ignorance. Usually, children only slowly begin to understand the relation between seeing and knowing (Pratt & Bryant, 1990; Ruffman & Olson, 1989) and to appreciate different sources of knowledge (Gopnik & Graf, 1988; O'Neill & Gopnik, 1991) in the months after their third birthday. However, repeated conversations with a more knowledgeable adult and explicit feedback on the effect of informational access on knowledge states led children to appreciate the causal link between seeing and knowing and differences between people's knowledge states earlier than in regular development. This aligns with Harris' (2005) argumentation that participation in conversation highlights different points of view, and that comprehending alternative viewpoints is relevant for mastering Theory of Mind. Also, a training study with three- and four-year-olds found effects of exposure to conversations in which others' perspectives were discussed on children's Theory of Mind (Gola, 2012). Moreover, Slaughter et al. (2007) emphasized that particularly maternal mental state talk which provides explanatory, causal, or contrastive information was related to children's Theory of Mind.

Surprisingly, the mental state training did not influence children's epistemic perspective-taking skills even though both the training and the epistemic perspective-taking tasks required reasoning about whether a person knows the content of a specific container. A possible

explanation for this result was offered in the discussion of Article 3: Within-person comparisons of knowledge about different objects and facts might be more challenging than between-person comparisons of knowledge about the same object or fact. Moreover, in the feedback during the training, between-person knowledge discrepancies were highlighted and explained more frequently than within-person knowledge discrepancies.

Article 3 also examined children's productive mental state vocabulary as a predictor of training-related improvement. Indeed, children's productive vocabulary for cognitive particles six to seven months before the training predicted their training-related improvement in the near-transfer seeing-knowing task. Particles are used to express the speaker's attitude or certainty towards what is said (Kauschke, 2012), and in Article 3, cognitive particles constituted the category with the second-lowest percentage of actively produced terms. Thus, children's productive vocabulary for rather difficult mental state terms might indicate their capability to improve their mental state reasoning beyond what they currently know. It seems, however, surprising that children's productive vocabulary for more typical cognitive terms (e.g., *know*, *think*, *guess*) did not predict their training-related improvement as well. Yet, children's mental state language was not assessed from their productions during the assessment but using a parent questionnaire (see Kristen et al., 2012, 2014). As Bartsch and Wellman (1995), as well as Shatz et al. (1983), have argued, children's early uses of cognitive terms may only fulfill rather conversational purposes and may not indicate children's true understanding of the utilized cognitive terms as referring to underlying mental states. Since parents only needed to mark all those terms that their child produces, it was not possible to know whether the respective child was using a certain cognitive term (e.g., *know* or *think*) only for conversational purposes (e.g., "Don't know, I think so.") or for truly referring to their knowledge and thoughts (e.g., "I don't know where my teddy bear is. I think I lost it at the doctor's office."). Thus, among the children credited with high productive vocabulary for cognitive terms, some children might already have developed a true understanding of the referential nature of cognitive terms, as referring to one's knowledge or beliefs, while others may still be using these terms solely as conversational devices. Productive vocabulary for cognitive terms may have been associated with training-related improvement but only in children who have already acquired the referential aspects of cognitive terms. In general, the observed effect of children's mental state language on their comprehension of mental states other than false beliefs extends the results presented in Section 1.5.4 and aligns with findings by Chiarella et al. (2013), who found

relations between 2.5-year-olds' mental state language and their visual and emotional perspective-taking as well as desire understanding.

Theoretical Implications

The above-mentioned finding that children's productive vocabulary for cognitive terms was not a predictor of training-related improvement suggests that productive vocabulary for cognitive terms might not be equated with a full, mentalistic understanding of such terms. In Article 3, productive vocabulary for cognitive particles, which are used to refer to differing degrees of certainty in knowledge, was a better indicator of children's capability to profit from a mental state training. Harris and colleagues (2017c) argued that young two-year-olds' early uses of cognitive terms already often refer to underlying epistemic states. Based on their earlier analyses of the same naturalistic speech data, Bartsch and Wellman (1995) and Shatz et al. (1983), however, had emphasized that two-year-olds rarely truly refer to cognitive states when using cognitive terms. According to these researchers, this is also reflected in the fact that children begin to use contrastives to compare different people's mental states only between ages three and four. The mental state training in Article 3 enabled three-year-old children to reason about certain aspects of knowledge that they do not usually spontaneously refer to in their naturalistic speech – such as the relation between visual access to a container and subsequent knowledge about the container's content. Even though three-year-olds learned to make correct inferences about these aspects of knowledge, the few explanation prompts included in the training showed that children still found it difficult to correctly verbally explain these inferences. Moreover, the training progress depicted in Article 3 revealed that the children did not acquire these insights at once from one training session to the next but in a stepwise manner, including setbacks in already mastered concepts. Concluding, there is empirical evidence for mentalistic uses of cognitive terms in the third year of life. Further, a linguistic intervention successfully improved three-year-olds' epistemic state reasoning. Thus, three-year-old children may be in a transitional period in which they only begin to construct a mentalistic understanding of knowledge states. This constructive process of epistemic understanding may not always become evident in children's naturalistic uses of cognitive terms because it is difficult to perfectly differentiate between conversational and truly referential uses of cognitive terms. As Montgomery (2002) argues, it is not clear whether even children themselves distinguish between different uses of the same term. Very early uses of cognitive terms may only serve conversational purposes. However, children in the third and fourth year of life,

increasingly utilize these terms for referential purposes, modeled onto how they have experienced other people apply such terms and based on their growing insights into their own and others' epistemic states.

Previous training studies which engaged children in discussions about mental states, frequently used group settings and, apart from two recent studies (Grazzani, Ornaghi, & Brockmeier, 2016; Grazzani, Ornaghi, Agliati et al., 2016), focused on children from age three on (e.g., Grazzani & Ornaghi, 2011; Guajardo & Watson, 2002; Tompkins, 2015; Tsuji, 2020). Article 3 added to this literature by showing that a conversation-based approach was also effective in promoting Theory of Mind in a one-on-one setting and with children below age three. This finding empirically corroborates Harris' (1999) argumentation that from 18 months old on – when children have fully mastered the ability to engage in joint attention with others – they can participate in information-exchanging discourse with their parents. Within these conversations with their parents, children are confronted with differences in knowledge or belief states among different people, which foster their comprehension of representational epistemic states. Moreover, the two-year-old children participating in the mental state training in Article 3 had not been taught about or trained in using complement syntax constructions, and the mental state training group did not perform significantly better on the complement syntax task than the other two groups. Thus, perspective-shifting discourse and mental state vocabulary alone were sufficient for improving the assessed mental state reasoning skills. This finding is in line with the beneficial effect of the discourse-only training in the Lohmann and Tomasello (2003) training study and the labeling training in the Sellabona et al. (2013) training study which both observed training effects of discourse without complement syntax. While linguistic determinism theory assumes that complement syntax competence is a *necessary* precondition for reasoning about counterfactual mental states (de Villiers & de Villiers, 2009), the findings of Article 3 suggest that this might not be the case for reasoning about earlier-developing mental states, such as perception or knowledge.

Practical Implications and Suggestions for Future Research

The effects of the linguistic perspective-shifting discourse intervention on two-year-olds' reasoning about epistemic states require replication, preferably using different types of training and potentially also focusing on other early developing mental states, such as desire or perception. Since older two-year-olds approach ceiling performance on desire and perception tasks, younger two-year-olds could serve as a target population for such a training study.

Suitable control conditions should be included to ensure that other confounding aspects of the intervention cannot account for the training effects. Moreover, it might be advisable to screen children for their pre-training performance in the tasks of interest and to allow only children performing, for instance, below 50% correct into the training. Previous training studies, which focused on improving children's false belief understanding, often only included children who did not yet solve the respective false belief task correctly into the training. When using a broader measure like the near-transfer seeing-knowing task, which consisted of several subtasks, it is difficult to define what passing or failing means on such a measure. Thus, setting a maximum level of 33% or 50% correct performance before the training would give all participating children sufficient room for improvement.

Parent questionnaires for assessing children's mental state vocabulary are an elegant, time-efficient, and frequently used method. In contrast, transcribing and coding each of children's utterances during the assessment, or a sub-part of the assessment, has some disadvantages: It is relatively time-consuming, and some children might be too shy or feel unwell in the unfamiliar environment and thus only talk little or do not show the whole variety of their productive mental state vocabulary. Using a parent questionnaire saves time in data coding and provides insights from the child's primary caregiver, who knows best which words the child can already actively produce. To avoid artificially high or unrealistic performance levels for certain terms, it might be helpful to have parents also report an example sentence in which the child typically produces a specific term (as in Bretherton & Beeghly, 1982). These examples could be requested either for a subset of mental state terms that have both a conversational and a referential use or for those terms that are particularly relevant for the research goal. Alternatively, using experience sampling methods, parents could be prompted daily to select all mental state terms from a checklist that their child produced during a fixed time period (e.g., during the evening hours). Additionally, the utterance in which the child used the term and the previous utterance of the interlocutor could be reported. Analyses of the effects of children's mental state language could then exclude mental state terms that the child had solely used for conversational purposes.

The positive effects of perspective-shifting conversations for children's Theory of Mind development imply that it would be beneficial for children to experience such conversations regularly in kindergartens or daycare centers. This would be particularly relevant for children coming from families in which parents have only little time to discuss mental state-related

topics with their children or in which parents are less aware that it is vital to engage young children in such conversations. Participation in mental state-related, perspective-shifting discourse with their educators, or more knowledgeable peers, might allow these children to catch up regarding their mental state reasoning skills. For example, an intervention study by Tompinks (2015) focused on three- to five-year-olds from low-income families. Tompkins (2015) found that participation in a training in which commercially available storybooks were read and afterward the story characters' mental states were discussed with the children had positive effects on their false belief understanding compared to two control groups. However, no transfer effects on emotion understanding or social skills were observed, which might have been due to the narrow focus of the storybooks on false beliefs (Tompkins, 2015). Engaging children in such conversations on a regular basis could thus have longer-lasting and broader effects on their Theory of Mind development. This would be a desirable outcome since proficient Theory of Mind skills are also relevant for children's later social and academic achievement in school: Some studies observed that better Theory of Mind skills are related to higher academic performance (Garner & Waajid, 2008), higher leadership status (hearing children in Peterson et al., 2016), or peer popularity (deaf children in Peterson et al., 2016; Slaughter et al., 2015). For instance, in a longitudinal study by Lecce et al. (2011, 2014), children's false belief understanding at 5.5 years predicted their academic achievement at 7.5 and ten years even when controlling for children's verbal ability and teacher-reported social skills. Moreover, Lockl and Schneider (2007) found that, even when controlling for general language skills, children's Theory of Mind at 3.5 and 4.5 years predicted their metamemory at 5.5 years, which is relevant for later success in school. Lecce et al. (2015) also discovered both cross-sectional and longitudinal relations between four-year-olds' cognitive Theory of Mind and their declarative metamemory. Thus, raising the awareness of educators and primary school teachers on how they can foster children in their understanding of others would greatly benefit children in later kindergarten and school classes. This goal can be accomplished, as described, by engaging children in structured book reading sessions with subsequent discussions about mental states and differences in perspectives. Further, practical training on how to conduct such sessions could be conveyed in educators' and primary school teachers' professional training.

5.4 Limitations and Directions for Future Research

The following section outlines overall limitations influencing the research presented in the three articles. Moreover, global directions for future research will be presented, which

highlight some general ideas for advancing research on Theory of Mind in the third year of life.

5.4.1 Limitations

One overall limitation affecting the research presented here was the outbreak of the Covid-19 pandemic. In March 2020, the first wave of Covid-19 hit Germany, resulting in immediate laboratory shutdowns and contact restrictions. As a consequence, the studies presented here lost some participants. This occurred because in the multi-cohort, longitudinal study, some cohorts of children had already undergone a few but not all of the measurement points, and the longitudinal design required assessing participants closely around the determined time points. Moreover, the microgenetic training needed to be terminated immediately or could not be commenced timely for a couple of children. Luckily, it was possible to continue data collection three months later by switching mostly to remote assessments via a video-conferencing tool, with children participating from home. Due to the young age groups assessed, children were not yet familiar with spending a long time in front of a computer screen. Still, online participation worked well for most children. Before combining the data from children tested remotely with the data from children tested in the laboratory, all relevant variables were tested for performance differences between the two testing environments (see Articles 1 and 3). Although no systematic performance differences were found, it might be possible that the different testing environments subtly influenced children's performance. For example, a few children were distracted by being in their home environments and having their toys close by. These issues sometimes resulted in lower compliance to participate. Since data collection for the first four measurement points of the control group had been finished before the Covid-19 pandemic, these changes in data collection only affected participants in the two training groups.

The remotely assessed children also participated in one final in-person appointment for some of the tasks. During these assessments, experimenters needed to wear face masks, making it more difficult for the children to recognize them from the previous remote assessments. Moreover, wearing a face mask aggravates it for young children to identify others' emotions (Gori et al., 2021; Schneider et al., 2022) and thus possibly also to understand others' positive attitudes towards them. Further, wearing a face mask muffles speech (Corey et al., 2020; Magee et al., 2020) and makes communication more difficult. Consequently, mainly

shy children who did not dare to ask for another repetition of the test question may have been hindered in showing high performance levels.

The research presented here used well-established Theory of Mind tasks to measure children's understanding of perception, knowledge, and false belief. Still, only a limited number of measures are available for the young age groups assessed. As with most Theory of Mind measures (Beaudoin et al., 2020), the utilized tasks only have one or few test trials. However, this reduces the variance that can be explained by predictors or correlates of the tasks. Further, it is problematic if a child fails to respond to the single trial, or one out of two trials, of an instrument since this results in missing data for this child. Thus, despite sampling almost 200 children for the longitudinal study, the majority of the children is missing data in at least one of the tasks. While this led to differing sample sizes for each analysis, the reported analyses still had sufficient power to detect the assumed effects.

A further limitation concerns the composition of the sample. On the one hand, the utilized language-intense tasks required including only children with a regular language development in the sample. Thus, children who had not mastered sufficient productive vocabulary by 27 months were dropped from the study. On the other hand, in the city in which the research was conducted more than 40% of all inhabitants have a migration background (Aydemir-Kengeroglu, 2019). As a result, 24% of the children in the final sample were growing up with more than one language. Without including any bilingual children in the sample, reaching the envisioned sample sizes would not have been possible in a timely manner. Only children with a good command of German, who heard German on at least 60 to 70% of their day, were kept in the sample. Still, bilingual children might not be fully comparable to monolingual children regarding their German proficiency. Researchers concur that all bilingual children have one more dominant language, in which they are more proficient, and that this dominant language can switch rapidly if external circumstances – such as the person mainly responsible for the child's care or the daycare center attended – change (Chilla, 2020; Nicoladis & Genesee, 1997; Wagner, 2014). Further, bilingual children's vocabulary size roughly equals the vocabulary size of monolingual children but only when comparing the conceptual vocabulary. This means, when considering all the words the bilingual child knows in either one of the two languages (Chilla, 2020; Nicoladis & Genesee, 1997). Thus, bilingual children's German vocabulary is usually smaller than that of monolingual German children. Lastly, all bilingual children code-mix between the two languages at some time, which is especially prominent in the

second and third year of life (Chilla, 2020; Wagner, 2014). Therefore, in the used standardized language assessments, bilingual children might have produced a sentence which contained a conceptually correct mixture of both languages but only the words produced in German were counted to measure their language skills.

5.4.2 Directions for Future Research

More research on young children's Theory of Mind before age three is highly needed. Training studies are a particularly suitable and insightful research method to investigate beneficial differential effects of certain aspects of language. However, to avoid selective drop-out from time-intense training studies, like the one in Article 3, and to achieve a higher parental willingness to finish the study, future studies should devise training methodologies that can more easily be integrated into young children's daily life while at the same time keeping a comparably high level of standardization between participants. Training studies could, for instance, be conducted in child daycare centers which the children attend anyway. An experimenter could regularly visit the daycare center and conduct training sessions with all participating children on the same days. However, this method is not well applicable to achieve a narrow age range of participants since the age range in child daycare centers is usually broad.

Another possibility could be to create interactive pre-recorded videos or apps with which children can learn about certain mental states in a playful way and get playfully feedback on their replies. Parents could administer these videos or apps to their children on predefined days using tablets or similar devices. As Kampis et al. (2021) noted in their replication of the Southgate et al. (2007) paradigm, children now grow up in different environments than ten to 15 years ago. Children are, for example, regularly exposed to digital media content. These children might find stimuli and procedures that used to be novel and modern – such as the videos in the Southgate et al. (2007) paradigm – less exciting today. This issue became, for instance, evident in Article 1 when administering the anticipatory-looking task to young children who are now used to watching animated television series. Thus, data collection methods might need to adapt to this change and increasingly utilize tasks that are more technology-based, more interactive, and better keep up children's attention. Durrleman et al. (2019), for example, reported considerable effects of a training administered using an iPad application on children's belief reasoning. The youngest children assessed were 2;9 years old, and the authors did not report issues applying such a training methodology. Training studies by Gola (2012) and San Juan and Astington (2017) utilized videos with animated characters as training

material and reported beneficial effects on three- and four-year-old children's explicit or implicit false belief understanding. A study on word-learning with two-year-olds compared children's vocabulary growth between an intervention in which children played two word-learning games on a tablet and an intervention with picture cards (Walter-Laager et al., 2016). Half of the children in each intervention were supported by an adult in their learning. Only children in the tablet group significantly improved their vocabulary (Walter-Laager et al., 2016). Both improvement and interest in the activity were higher if children were accompanied by an adult. Further, a meta-analysis by Takacs et al. (2014) found that four- to eight-year-olds' comprehension of stories improved comparably much when children interacted with storybooks enriched with multimedia elements and when children looked at traditional storybooks together with an adult. Thus, technology-based research paradigms have not only successfully been applied with young children but could also be applied more frequently to circumvent problems such as decreased interest or attention in the task.

In the training study reported in Article 3, a microgenetic design with a high number of training sessions was conducted. Analyzing the training progress of the individual children in the mental state training (see Article 3) revealed that some children acquired the conveyed concepts very fast or even started at slightly higher performance levels than the majority of the children. For high-performing children, or children who had learned all the concepts by the end of the first half of the training, the following training sessions merely constituted further repetitions of the already understood concepts. While this undoubtedly helped children consolidate their new knowledge, it is possible that it also led to decreased interest in the training sessions. Moreover, those children could possibly have acquired even more elaborate epistemic concepts given additional, more complex training tasks. Future training studies could consider including adaptive elements into the training such that children who are passing the same task several times are confronted with a slightly more advanced task during the next session. Also, the feedback children receive could be adapted on-line to the type of errors the children make (e.g., altercentric or egocentric errors). As mentioned in Article 3, some children in the complement syntax training did not seem to have understood the underlying task of repeating the content of the complement clause. In such cases, an adaptive training could also switch back to slightly easier tasks (e.g., repetition of simple sentences) to first familiarize children with the task format such that they can then profit more from the following training sessions. Implementing an adaptive training methodology would account for inter-

individual differences in how quickly children acquire certain socio-cognitive concepts. Moreover, using an adaptive method could test the limits of how far young children's early comprehension of certain concepts can be improved with a given training method.

As mentioned in Section 5.5.1, most Theory of Mind tasks only use a few items to measure a construct. This leads to reduced variance in performance and a particular risk of having missing data due to young children's unwillingness to respond to the test question. While Wellman and Liu (2004) constructed a Theory of Mind scale forming a developmental progression and consisting of tasks that assess different constructs, the scale's difficulty is too advanced for two-year-old children. Moreover, this Theory of Mind scale assesses mental state reasoning broadly by measuring understanding of desire, (false) belief, knowledge, and emotions. The near-transfer task used in Article 3 consisted of several tasks focusing on the relation between seeing and knowing and sources of knowledge based on tasks by Gopnik and Graf (1988), Pratt and Bryant (1990), Ruffman and Olson (1989), and Wimmer et al. (1988). Due to its multi-trial nature, the task allowed assessing comprehension of the seeing-knowing relation before and after the training for almost every child despite missing data in single subtasks. A percentage score was calculated and utilized if the child had valid data in more than half of the subtasks (i.e., in at least three out of five subtasks). Similarly, scales for measuring visual perspective-taking or desire understanding with several items could be constructed. While such scales or paradigms with multiple items exist, for instance, for false belief understanding (Hughes et al., 2000), emotion recognition (Pons & Harris, 2000), or understanding of non-literal communication (Happé, 1994), they have rarely been used in Theory of Mind research so far (see Beaudoin et al., 2020, Appendix II). Thus, the construction of novel scales with several items for assessing young children's mental state reasoning and the application of existing scales are advisable to mitigate the issue of missing data. This is particularly relevant when collecting data from young children who have a higher tendency to refuse to answer some of the questions than older participants.

5.5 Conclusion

Theory of Mind research has been a busy research field in the past 40 years. Human interaction and living together might barely be possible without being able to comprehend others' goals, wishes, and beliefs. Thus, developing a profound understanding of others' mental lives is an essential accomplishment for young children. This dissertation presented the findings of three empirical articles which focused on understanding the extent of children's

Theory of Mind abilities in the third year of life and on the role different aspects of language play in children's acquisition of such early Theory of Mind abilities. This research contributes the following main points to the existing literature:

First, a promising multi-trial anticipatory-looking paradigm by Grosse Wiesmann et al. (2017) was found to not reliably capture two- to four-year-olds' false belief reasoning. As reported in Article 1, a large-scale, longitudinal replication attempt only replicated above-chance belief-congruent action anticipation in one of two false belief conditions. Thus, other, novel paradigms, utilizing more modern stimuli and including carefully selected control conditions, are highly needed to better understand the extent of infants' and young children's implicit Theory of Mind. As a methodological contribution, Article 1 used an innovative method, described by Anderson and Maxwell (2016), to compare the original and the replication study's effect sizes. This supplementary analysis provided relevant additional information on top of the usual replication attempt. Further, correlative findings in Articles 1 and 2 revealed some coherence in children's early Theory of Mind around their third birthday.

Second, complement syntax skills were found to already play a role in children's Theory of Mind development around their third birthday. Article 2 observed that 33-month-olds' sensitivity to the correct syntax of complementation was predictive of their later visual and epistemic perspective-taking and metacognition of their own ignorance. The complement syntax training in Article 3 neither improved children's complement syntax skills nor their mental state reasoning. Though, children's pre-training complement syntax skills were associated with their metacognition across training and control groups. Together, Articles 2 and 3 provided some first evidence for the relevance of complement syntax skills around the third birthday and for acquiring aspects of Theory of Mind other than false belief understanding.

Third, children's early reasoning about epistemic states can already be fostered below age three using a linguistic intervention. Article 3 reported on a microgenetic training study in which children were repeatedly engaged in perspective-shifting discourse enriched with mental state terms. This training affected children's comprehension of the seeing-knowing relation and their metacognition. Moreover, children's pre-training productive vocabulary for cognitive particles was predictive of the amount of training-related improvement. Article 3 constitutes one of the first Theory of Mind training studies conducted with children below age three and requires conceptual replication. Due to the high number of and the structural similarity among training sessions, an analysis of children's progression through the training was

possible. These analyses provided a better understanding of how children come to learn about specific concepts and in which order they comprehend important substeps.

In combination, the three articles shed light on the Theory of Mind developments children undergo in the third year of life under the influence of their increasingly proficient language skills. Previous research indicates that children start to shift from an early implicit understanding of mental states to a more explicit one in this period. Undoubtedly, future research will continue to uncover how much children at this age understand about their own and other minds and how their developing language abilities assist them in making sense of others' behavior. Recently, Wellman (2018) provided a comprehensive overview of the different turns Theory of Mind research has taken and the different fields Theory of Mind research has touched upon. I can only agree with his concluding remark that "while the field began with nothing but questions; now we have many answers. But of course, answers provoke new questions. So, one key 'state of the art' accomplishment is a firmer sense of how much remains to be known" (Wellman, 2018, p. 749).

References

Abbot-Smith, K., Lieven, E., & Tomasello, M. (2001). What preschool children do and do not do with ungrammatical word orders. *Cognitive Development*, *16(*2), 679-692. https://doi.org/10.1016/S0885-2014(01)00054-5

Adrián, J. E., Clemente, R. A., & Villanueva, L. (2007). Mothers' use of cognitive state verbs in picture-book reading and the development of children's understanding of mind: A longitudinal study. *Child Development*, *78*(4), 1052-1067. https://doi.org/10.1111/j.1467-8624.2007.01052.x

Adrián, J. E., Clemente, R. A., Villanueva, L., & Rieffe, C. (2005). Parent-child picture-book reading, mothers' mental state language and children's theory of mind. *Journal of Child Language*, *32*(3), 673-686. https://doi.org/10.1017/S0305000905006963

Anderson, S. F., & Maxwell, S. E. (2016). There's more than one way to conduct a replication study: Beyond statistical significance. *Psychological Methods*, *21*(1), 1-12. https://doi.org/10.1037/met0000051

Apperly, I. A. (2012). What is "theory of mind"? Concepts, cognitive processes and individual differences. *Quarterly Journal of Experimental Psychology*, *65*(5), 825-839. https://doi.org/10.1080/17470218.2012.676055

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. https://doi.org/10.1037/a0016923

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*, 862-868. https://doi.org/10. 1111/j.1467-7687.2008.00736.x

Aslin, R. N. (2007). What's in a look? *Developmental Science*, *10*(1), 48-53. https://doi.org/10.1111/j.1467-7687.2007.00563.x

Astington, J. W. (2000). Language and metalanguage in children's understanding of mind. In J. W. Astington (Ed.), *Minds in the making: Essays in honor of David R. Olson* (pp. 267-284). Oxford, United Kingdom: Blackwell.

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology*, *35*(5), 1311-1320. https://doi.org/10.1037/0012-1649.35.5.1311

Aydemir-Kengeroglu, S. (2019). Demografie der Münchner Bevölkerung 2018: Auswertungen des Bevölkerungsbestandes zum Jahresende 2018 sowie der Bevölkerunsgbewegungen des Jahres 2018. *Münchner Statistik*, *1*, 6-11. https://stadt.muenchen.de/dam/jcr:f6921ff5-7f41-4b7d-8661-8b12d1efa57b/mb190101.pdf

Baillargeon, R., Buttelmann, D., & Southgate, V. (2018). Invited commentary: Interpreting failed replications of early false-belief findings: Methodological and theoretical considerations. *Cognitive Development*, *46*, 112-124. https://doi.org/10.1016/j.cogdev.2018.06.001

Baillargeon, R., Scott, R. M., & Bian, L. (2016). Psychological reasoning in infancy. *Annual Review of Psychology*, *67*, 159-186. https://doi.org/10.1146/annurev-psych-010213-115033

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, *14*(3), 110-118. https://doi.org/10.1016/j.tics.2009.12.006

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

Barone, P., Corradi, G., & Gomila, A. (2019). Infants' performance in spontaneous-response false belief tasks: A review and meta-analysis. *Infant Behavior and Development*, *57*, 101350. https://doi.org/10.1016/j.infbeh.2019.101350

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition*, *21*(1), 37-46. https://doi.org/10.1016/0010-0277(85)90022-8

Bartsch, K. (1996). Between desires and beliefs: Young children's action predictions. *Child Development*, *67*(4), 1671-1685. https://doi.org/10.1111/j.1467-8624.1996.tb01820.x

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

Beaudoin, C., Leblanc, É., Gagner, C., & Beauchamp, M. H. (2020). Systematic review and inventory of theory of mind measures for young children. *Frontiers in Psychology*, *10*, 2905. https://doi.org/10.3389/fpsyg.2019.02905

Beeghly, M., Bretherton, I., & Mervis, C. B. (1986). Mothers' internal state language to toddlers. *British Journal of Developmental Psychology*, *4*(3), 247-261. https://doi.org/10.1111/j.2044-835X.1986.tb01016.x

Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science*, *8*(6), 492-499. https://doi.org/10.1111/j.1467-7687.2005.00440.x

Bennett, J. (1978). Some remarks about concepts. *Behavioral and Brain Sciences*, *1*, 557-560. https://doi.org/10.1017/S0140525X00076573

Bloom, L., Rispoli, M., Gartner, B., & Hatz, J. (1989). Acquisition of complementation. *Journal of Child Language*, *16*(1), 101-120. https://doi.org/10.7916/D8FQ9WW4

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. https://doi.org/10.1016/S0010-0277(00)00096-2

Bockmann, A.-K., Sachse, S., & Buschmann, A. (2020). Sprachentwicklung im Überblick. In S. Sachse, A.-K. Bockmann, & A. Buschmann (Eds.), *Sprachentwicklung* (pp. 3-44). Berlin, Germany: Springer.

Boeg Thomsen, D., Theakston, A., Kandemirci, B., & Brandt, S. (2021). Do complement clauses really support false-belief reasoning? A longitudinal study with English-speaking 2-to 3-year-olds. *Developmental Psychology*, *57*(8), 1210-1227. https://doi.org/10.1037/dev0001012

Bowerman, M. (1979). The acquisition of complex sentences. In P. Fletcher & M. Gorman (Eds.), *Language acquisition: Studies in first language development* (pp. 285-306). Cambridge, MA: Cambridge University Press.

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. https://doi.org/10.1016/j.jecp.2016.03.004

Brandt, S., Lieven, E., & Tomasello, M. (2010). Development of word order in German complement-clause constructions: Effects of input frequencies, lexical items, and discourse function. *Language*, *86*(3), 583-610. https://doi.org/10.1353/lan.2010.0010

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. https://doi.org/10.1515/cogl.2011.013

Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology*, *18*(6), 906-921. https://doi.org/10.1037/0012-1649.18.6.906

Bretherton, I., McNew, S., & Beeghly-Smith, M. (1981). Early person knowledge as expressed in gestural and verbal communication: When do infants acquire a ‘‘theory of mind’’?. In M. Lamb & L. Sherrod (Eds.), *Social cognition in infancy* (pp. 333-373). Hillsdale, NJ: Erlbaum.

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. https://doi.org/10.1016/j.jecp.2014.09.010

Brown, J. R., Donelan-McCall, N., & Dunn, J. (1996). Why talk about mental states? The significance of children's conversations with friends, siblings, and mothers. *Child Development*, *67*(3), 836-849. https://doi.org/10.1111/j.1467-8624.1996.tb01767.x

Burnel, M., Durrleman, S., Reboul, A., Carré, A., Baciu, M., & Perrone-Bertolotti, M. (2021). Theory-of-mind during childhood: Investigating syntactic and executive contributions. *Social Development*, *30*(1), 73-94. https://doi.org/10.1111/sode.12471

Burnside, K., Ruel, A., Azar, N., & Poulin-Dubois, D. (2018). Implicit false belief across the lifespan: Non-replication of an anticipatory looking task. *Cognitive Development*, *46*, 4-11. https://doi.org/10.1016/j.cogdev.2017.08.006

Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, *112*(2), 337-342. https://doi.org/10.1016/j.cognition.2009.05.006

Cannon, E. N., & Woodward, A. L. (2012). Infants generate goal-based action predictions. *Developmental Science*, *15*(2), 292-298. https://doi.org/10.1111/j.1467-7687.2011.01127.x

Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen-through 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development*, *21*(2), 315-330. https://doi.org/10.1016/S0163-6383(98)90009-1

Carpenter, M., Call, J., & Tomasello, M. (2005). Twelve-and 18-month-olds copy actions in terms of goals. *Developmental Science*, *8*(1), F13-F20. https://doi.org/10.1111/j.1467-7687.2004.00385.x

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *63*(4), i-174. https://doi.org/10.2307/1166214

Carr, A., Slade, L., Yuill, N., Sullivan, S., & Ruffman, T. (2018). Minding the children: A longitudinal study of mental state talk, theory of mind, and behavioural adjustment from the age of 3 to 10. *Social Development*, *27*(4), 826-840. https://doi.org/10.1111/sode.12315

Carruthers, P. (2013). Mindreading in infancy. *Mind & Language*, *28*(2), 141-172. https://doi.org/10.1111/mila.12014

Carruthers, P. (2016). Two systems for mindreading? *Review of Philosophy and Psychology*, *7*(1), 141-162. https://doi.org/10.1007/s13164-015-0259-y

Cheung, H., Chen, H. C., & Yeung, W. (2009). Relations between mental verb and false belief understanding in Cantonese-speaking children. *Journal of Experimental Child Psychology*, *104*(2), 141-155. https://doi.org/10.1016/j.jecp.2009.05.004

Cheung, H., Hsuan-Chih, C., Creed, N., Ng, L., Ping Wang, S., & Mo, L. (2004). Relative roles of general and complementation language in theory-of-mind development: Evidence from Cantonese and English. *Child Development*, *75*(4), 1155-1170. https://doi.org/10.1111/j.1467-8624.2004.00731.x

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. https://doi.org/10.1080/15248372.2012.689390

Chilla, S. (2020). Mehrsprachige Entwicklung. In S. Sachse, A.-K. Bockmann, & A. Buschmann (Eds.), *Sprachentwicklung* (pp. 109-130). Berlin, Germany: Springer.

Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, *9*(4), 377-395. https://doi.org/10.1016/0885-2014(94)90012-4

Clements, W. A., Rustin, C. L., & McCallum, S. (2000). Promoting the transition from implicit to explicit understanding: A training study of false belief. *Developmental Science*, *3*(1), 81-92. https://doi.org/10.1111/1467-7687.00102

Colonnesi, C., Rieffe, C., Koops, W., & Perucchini, P. (2008). Precursors of a theory of mind: A longitudinal study. *British Journal of Developmental Psychology*, *26*(4), 561-577. https://doi.org/10.1348/026151008X285660

Corey, R. M., Jones, U., & Singer, A. C. (2020). Acoustic effects of medical, cloth, and transparent face masks on speech signals. *The Journal of the Acoustical Society of America*, *148*(4), 2371-2375. https://doi.org/10.1121/10.0002279

Crivello, C., & Poulin-Dubois, D. (2018). Infants' false belief understanding: A non-replication of the helping task. *Cognitive Development*, *46*, 51-57. https://doi.org/10.1016/j.cogdev.2017.10.003

Davis, E. E., & Landau, B. (2021). Seeing and believing: The relationship between perception and mental verbs in acquisition. *Language Learning and Development*, *17*(1), 26-47. https://doi.org/10.1080/15475441.2020.1862660

De Mulder, H. N. M. (2015). Developing communicative competence: A longitudinal study of the acquisition of mental state terms and indirect requests. *Journal of Child Language*, *42*(5), 969-1005. https://doi.org/10.1017/S0305000914000543

De Mulder, H. N. M., Wijnen, F., & Coopmans, P. H. (2019). Interrelationships between Theory of Mind and language development: A longitudinal study of Dutch-speaking kindergartners. *Cognitive Development*, *51*, 67-82. https://doi.org/10.1016/j.cogdev.2019.03.006

de Villiers, J. G. (2005). Can language acquisition give children a point of view? In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 186-219). New York, NY: Oxford University Press. https://doi.org/ 10.1093/acprof:oso/9780195159912.003.0010

de Villiers, J. G., & de Villiers, P. A. (2000). Linguistic determinism and the understanding of false belief. In P. Mitchell & K. Riggs (Eds.), *Children's reasoning and the mind* (pp. 191-228). Hove, United Kingdom: Psychology Press.

de Villiers, J. G., & de Villiers, P. A. (2003). Language for thought: Coming to understand false beliefs. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 335-384). Cambridge, MA: The MIT Press.

de Villiers, J. G., & de Villiers, P. A. (2009). Complements enable representation of the contents of false beliefs: The evolution of a theory of theory of mind. In S. H. Foster-

Cohen (Ed.), *Language acquisition* (pp. 169-195). London, United Kingdom: Palgrave Macmillan. https://doi.org/10.1057/9780230240780_8

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. https://doi.org/10.1016/S0885-2014(02)00073-4

Dennett, D. C. (1978). Beliefs about beliefs. *Behavioral and Brain Sciences*, *1*, 568-570. https://doi.org/10.1017/S0140525X00076664

Devine, R. T., & Hughes, C. (2013). Silent films and strange stories: Theory of mind, gender, and social experiences in middle childhood. *Child Development*, *84*(3), 989-1003. https://doi.org/10.1111/cdev.12017

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. https://doi.org/10.1111/cdev.12237

Devine, R. T., & Hughes, C. (2018). Family correlates of false belief understanding in early childhood: A meta-analysis. *Child Development*, *89*(3), 971-987. https://doi.org/10.1111/cdev.12682

Diessel, H., & Tomasello, M. (2001). The acquisition of finite complement clauses in English: A corpus-based analysis. *Cognitive Linguistics*, *12*(2), 97-141. https://doi.org/10.1515/cogl.12.2.97

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. https://doi.org/10.1016/j.cogdev.2018.01.001

Dörrenberg, S., Wenzel, L., Proft, M., Rakoczy, H., & Liszkowski, U. (2019). Reliability and generalizability of an acted-out false belief task in 3-year-olds. *Infant Behavior and Development*, *54*, 13-21. https://doi.org/10.1016/j.infbeh.2018.11.005

Dunn, J., & Brophy, M. (2005). Communication, relationships, and individual differences in children's understanding of mind. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 50-69). New York, NY: Oxford University Press. https://doi.org/ 10.1093/acprof:oso/9780195159912.003.0003

Dunn, J., Brown, J., Slomkowski, C., Tesla, C., & Youngblade, L. (1991). Young children's understanding of other people's feelings and beliefs: Individual differences and their

antecedents. *Child Development*, *62*(6), 1352-1366. https://doi.org/10.1111/j.1467-8624.1991.tb01610.x

Durrleman, S. (2017). Grammar and Theory of Mind in autism. In J. Blochowiak, C. Grisot, S. Durrleman, & C. Laenzlinger (Eds.), *Formal models in the study of language: Applications in interdisciplinary contexts* (pp. 327-337). Cham, Switzerland: Springer. https://doi.org/ 10.1007/978-3-319-48832-5_17

Durrleman, S., Burnel, M., de Villiers, J. G., Thommen, E., Yan, R., & Delage, H. (2019). The impact of grammar on mentalizing: A training study including children with Autism Spectrum Disorder and Developmental Language Disorder. *Frontiers in Psychology*, *10*, 2478. https://doi.org/10.3389/fpsyg.2019.02478

Ensor, R., Devine, R. T., Marks, A., & Hughes, C. (2014). Mothers' cognitive references to 2-year-olds predict theory of mind at ages 6 and 10. *Child Development*, *85*(3), 1222-1235. https://doi.org/10.1111/cdev.12186

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience*, *9*(7), 878-879. https: //doi.org/10.1038/nn1729

Farrar, M. J., & Maag, L. (2002). Early language development and the emergence of a theory of mind. *First Language*, *22*(2), 197-213. https://doi.org/10.1177/014272370202206504

Farrar, M. J., Benigno, J. P., Tompkins, V., & Gage, N. A. (2017). Are there different pathways to explicit false belief understanding? General language and complementation in typical and atypical children. *Cognitive Development*, *43*, 49-66. https://doi.org/10.1016/j.cogdev.2017.02.005

Fedra, E., & Schmidt, M. F. (2019). Older (but not younger) preschoolers reject incorrect knowledge claims. *British Journal of Developmental Psychology*, *37*(1), 130-145. https://doi.org/10.1111/bjdp.12264

Fenici, M., & Garofoli, D. (2020). An associationist bias explains different processing demands for toddlers in different traditional false-belief tasks. *Human Development*, *64*, 4-6. https://doi.org/10.1159/000505208

Flavell, J. H. (1988). The development of children's knowledge about the mind. In J. W. Astington, P. L. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (pp. 244-267). Cambridge, United Kingdom: Cambridge University Press.

Flavell, J. H. (2000). Development of children's knowledge about the mental world. *International Journal of Behavioral Development*, *24*(1), 15-23. https://doi.org/10.1080/016502500383421

Flavell, J. H. (2004). Theory-of-mind development: Retrospect and prospect. *Merrill-Palmer Quarterly, 50*(3), 274-290. https://doi.org/10.1353/mpq.2004.0018

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. https://doi.org/10.1037/0012-1649.17.1.99

Flavell, J. H., Flavell, E. R., & Green, F. L. (1983). Development of the appearance-reality distinction. *Cognitive Psychology*, *15*(1), 95-120. https://doi.org/10.1016/0010-0285(83)90005-1

Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature*, *424*(6950), 769-771. https://doi.org/10.1038/nature01861

Fontana, E., Adenzato, M., Penso, J. S., Enrici, I., & Ardito, R. B. (2018). On the relationship between Theory of Mind and syntax in clinical and non-clinical populations: State of the art and implications for research. *The Open Psychology Journal*, *11*, 95-104. https://doi.org/10.2174/1874350101811010095

Frith, C. D., & Frith, U. (2012). Mechanisms of social cognition. *Annual Review of Psychology*, *63*, 287-313. https://doi.org/10.1146/annurev-psych-120710-100449

Furrow, D., Moore, C., Davidge, J., & Chiasson, L. (1992). Mental terms in mothers' and children's speech: Similarities and relationships. *Journal of Child Language*, *19*(3), 617-631. https://doi.org/10.1017/S0305000900011594

Garner, P. W., & Waajid, B. (2008). The associations of emotion knowledge and teacher-child relationships to preschool children's school-related developmental competence. *Journal of Applied Developmental Psychology*, *29*(2), 89-100. https://doi.org/10.1016/j.appdev.2007.12.001

Garnham, W. A., & Ruffman, T. (2001). Doesn't see, doesn't know: Is anticipatory looking really related to understanding of belief?. *Developmental Science*, *4*(1), 94-100. https://doi.org/10.1111/1467-7687.00153

Ghrear, S., Baimel, A., Haddock, T., & Birch, S. A. J. (2021). Are the classic false belief tasks cursed? Young children are just as likely as older children to pass a false belief task

when they are not required to overcome the curse of knowledge. *PLoS ONE*, *16*(2), e0244141. https://doi.org/10.1371/journal.pone.0244141

Gola, A. A. H. (2012). Mental verb input for promoting children's theory of mind: A training study. *Cognitive Development*, *27*(1), 64-76. https://doi.org/10.1016/j.cogdev.2011.10.003

Golan, O., Baron-Cohen, S., Hill, J. J., & Golan, Y. (2006). The "reading the mind in films" task: Complex emotion recognition in adults with and without autism spectrum conditions. *Social Neuroscience, 1*(2), 111-123. https://doi.org/10.1080/17470910600980986

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. https://doi.org/10.1111/cdev.12876

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, *59*(1), 26-37. https://doi.org/10.2307/1130386

Gopnik, A., & Graf, P. (1988). Knowing how you know: Young children's ability to identify and remember the sources of their beliefs. *Child Development*, *59*(5), 1366-1371. https://doi.org/10.2307/1130499

Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind & Language*, *7*(1-2), 145-171. https://doi.org/10.1111/j.1468-0017.1992.tb00202.x

Gori, M., Schiatti, L., & Amadeo, M. B. (2021). Masking emotions: Face masks impair how we read emotions. *Frontiers in Psychology*, *12*, 669432. https://doi.org/10.3389/fpsyg.2021.669432

Grazzani, I., & Ornaghi, V. (2011). Emotional state talk and emotion understanding: A training study with preschool children. *Journal of Child Language*, *38*(5), 1124-1139. https://doi.org/10.1017/S0305000910000772

Grazzani, I., Ornaghi, V., & Brockmeier, J. (2016). Conversation on mental states at nursery: Promoting social cognition in early childhood. *European Journal of Developmental Psychology*, *13*(5), 563-581. https://doi.org/10.1080/17405629.2015.1127803

Grazzani, I., Ornaghi, V., Agliati, A., & Brazzelli, E. (2016). How to foster toddlers' mental-state talk, emotion understanding, and prosocial behavior: A conversation-based

intervention at nursery school. *Infancy*, *21*(2), 199-227. https://doi.org/10.1111/infa.12107

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. https://doi.org/10.1016/j.cogdev.2017.08.007

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. https://doi.org/10.1111/desc. 12445

Grosso, S. S., Schuwerk, T., Kaltefleiter, L. J., & Sodian, B. (2019). 33-month-old children succeed in a false belief task with reduced processing demands: A replication of Setoh et al. (2016). *Infant Behavior and Development*, *54*, 151-155. https://doi.org/10.1016/j.infbeh.2018.09.012

Guajardo, N. R., & Watson, A. C. (2002). Narrative discourse and theory of mind development. *The Journal of Genetic Psychology*, *163*(3), 305-325. https://doi.org/10.1080/00221320209598686

Hala, S., & Chandler, M. (1996). The role of strategic planning in accessing false-belief understanding. *Child Development*, *67*(6), 2948-2966. https://doi.org/10.1111/j.1467-8624.1996.tb01897.x

Hale, C. M., & Tager-Flusberg, H. (2003). The influence of language on theory of mind: A training study. *Developmental Science*, *6*(3), 346-359. https://doi.org/ 10.1111/1467-7687.00289

Happé, F. G. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders*, *24*(2), 129-154. https://doi.org/10.1007/BF02172093

Happé, F. G. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development*, *66*(3), 843-855. https://doi.org/10.1111/j.1467-8624.1995.tb00909.x

Harman, G. (1978). Studying the chimpanzee's theory of mind. *Behavioral and Brain Sciences*, *1*, 576-577. https://doi.org/10.1017/S0140525X00076743

Harris, P. L. (1992). From simulation to folk psychology: The case for development. *Mind & Language*, *7*(1-2), 120-144. https://doi.org/10.1111/j.1468-0017.1992.tb00201.x

Harris, P. L. (1996). Desires, beliefs, and language. In P. Carruthers & P. K. Smith (Eds.), *Theories of theories of mind* (pp. 200-220). New York, NY: Cambridge University Press. https://doi.org/10.1017/CBO9780511597985

Harris, P. L. (1999). Acquiring the art of conversation: Children's developing conception of their conversation partner. In M. Bennett (Ed.), *Developmental psychology: Achievements and prospects* (pp. 89-105). Hove, United Kingdom: Psychology Press.

Harris, P. L. (2005). Conversation, pretense, and theory of mind. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 70-83). New York, NY: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195159912.003.0004

Harris, P. L., Bartz, D. T., & Rowe, M. L. (2017a). Young children communicate their ignorance and ask questions. *Proceedings of the National Academy of Sciences*, *114*(30), 7884-7891. https://doi.org/10.1073/pnas.1620745114

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. https://doi.org/10.1111/j.0963-7214.2005.00337.x

Harris, P. L., Ronfard, S., & Bartz, D. (2017b). Young children's developing conception of knowledge and ignorance: Work in progress. *European Journal of Developmental Psychology*, *14*(2), 221-232. https://doi.org/10.1080/17405629.2016.1190267

Harris, P. L., Yang, B., & Cui, Y. (2017c). I don't know and not: I don't know. *Mind & Language*, *32*(3), 283-307. https://doi.org/10.1111/mila.12143

He, Z., Bolz, M., & Baillargeon, R. (2011). False-belief understanding in 2.5-year-olds: Evidence from violation-of-expectation change-of-location and unexpected-contents tasks. *Developmental Science*, *14*(2), 292-305. https://doi.org/10.1111/j.1467-7687.2010.00980.x

Heyes, C. (2014). False belief in infancy: A fresh look. *Developmental Science*, *17*(5), 647-659. https://doi.org/10.1111/desc.12148

Hogrefe, G. J., Wimmer, H., & Perner, J. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child Development*, *57*(3), 567-582. https://doi.org/10.2307/1130337

Howard, A. A., Mayeux, L., & Naigles, L. R. (2008). Conversational correlates of children's acquisition of mental verbs and a theory of mind. *First Language*, *28*(4), 375-402. https://doi.org/10.1177/0142723708091044

Hughes, C., & Dunn, J. (1997). “Pretend you didn't know”: Preschoolers' talk about mental states in pretend play. *Cognitive Development*, *12*(4), 477-497. https://doi.org/10.1016/S0885-2014(97)90019-8

Hughes, C., & Dunn, J. (1998). Understanding mind and emotion: Longitudinal associations with mental-state talk between young friends. *Developmental Psychology*, *34*(5), 1026-1037. https://doi.org/10.1037/0012-1649.34.5.1026

Hughes, C., & Ensor, R. (2007). Executive function and theory of mind: Predictive relations from ages 2 to 4. *Developmental Psychology*, *43*(6), 1447-1459. https://doi.org/10.1037/0012-1649.43.6.1447

Hughes, C., Adlam, A., Happé, F., Jackson, J., Taylor, A., & Caspi, A. (2000). Good test-retest reliability for standard and advanced false-belief tasks across a wide range of abilities. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *41*(4), 483-490. https://doi.org/10.1111/1469-7610.00633

Jenkins, J. M., & Astington, J. W. (1996). Cognitive factors and family structure associated with theory of mind development in young children. *Developmental Psychology*, *32*(1), 70-78. https://doi.org/10.1037/0012-1649.32.1.70

Jenkins, J. M., Turrell, S. L., Kogushi, Y., Lollis, S., & Ross, H. S. (2003). A longitudinal investigation of the dynamics of mental state talk in families. *Child Development*, *74*(3), 905-920. https://doi.org/10.1111/1467-8624.00575

Johnson, C. N., & Wellman, H. M. (1980). Children's developing understanding of mental verbs: Remember, know, and guess. *Child Development*, *51*(4), 1095-1102. https://doi.org/10.2307/1129549

Kaltefleiter, L. J., Schuwerk, T., Grosse Wiesmann, C., Kristen-Antonow, S., Jarvers, I., & Sodian, B. (2022). Evidence for goal- and mixed evidence for false belief-based action prediction in two- to four-year-old children: A large-scale longitudinal anticipatory looking replication study. *Developmental Science*, *25*(4), e13224. https://doi.org/10.1111/desc.13224

Kaltefleiter, L. J., Sodian, B., Kristen-Antonow, S., Grosse Wiesmann, C., & Schuwerk, T. (2021). Does syntax play a role in Theory of Mind development before the age of 3

years? *Infant Behavior and Development*, *64*, 101575. https://doi.org/10.1016/j.infbeh.2021.101575

Kaltefleiter, L. J., Sodian, B., Schuwerk, T., & Kristen-Antonow, S. (2022, January 21). "I know what's inside because you showed me": Training 33-month-old children to attribute knowledge and ignorance to themselves and others. *Infant Behavior and Development*, *68*, 101744. https://doi.org/10.1016/j.infbeh.2022.101744

Kammermeier, M., & Paulus, M. (2018). Do action-based tasks evidence false-belief understanding in young children? *Cognitive Development*, *46*, 31-39. https://doi.org/10.1016/j.cogdev.2017.11.004

Kampis, D., Karman, P., Csibra, G., Southgate, V., & Hernik, M. (2021). A two-lab direct replication attempt of Southgate, Senju and Csibra (2007). *Royal Society Open Science*, *8*(8), 210190. https://doi.org/10.1098/rsos.210190

Kauschke, C. (2000). *Der Erwerb des frühkindlichen Lexikons: Eine empirische Studie zur Entwicklung des Wortschatzes im Deutschen.* Tübingen, Germany: Narr.

Kauschke, C. (2012). *Kindlicher Spracherwerb im Deutschen: Verläufe, Forschungsmethoden, Erklärungsansätze*. Berlin, Germany: De Gruyter. https://doi.org/10.1515/9783110283891

Kauschke, C., & Hofmeister, C. (2002). Early lexical development in German: A study on vocabulary growth and vocabulary composition during the second and third year of life. *Journal of Child Language*, *29*(4), 735-757. https://doi.org/10.1017/S0305000902005330

Kauschke, C., & Klann-Delius, G. (1997). The acquisition of verbal expressions for internal states in German. In S. Niemeier & R. Dirven (Eds.), *The language of emotions* (pp. 173-194). Amsterdam, Netherlands: John Benjamins Publishing Company.

Kidd, E., Lieven, E., & Tomasello, M. (2006). Examining the role of lexical frequency in the acquisition and processing of sentential complements. *Cognitive Development*, *21*(2), 93-107. https://doi.org/10.1016/j.cogdev.2006.01.006

Kidd, E., Lieven, E., & Tomasello, M. (2010). Lexical frequency and exemplar-based learning effects in language acquisition: Evidence from sentential complements. *Language Sciences*, *32*(1), 132-142. https://doi.org/10.1016/j.langsci.2009.05.002

Kikuno, H., Mitchell, P., & Ziegler, F. (2007). How do young children process beliefs about beliefs? Evidence from response latency. *Mind & Language*, *22*(3), 297-316. https://doi.org/10.1111/j.1468-0017.2007.00310.x

Kinderman, P., Dunbar, R. I. M., & Bentall, R. P. (1998). Theory-of-mind deficits and causal attributions. *British Journal of Psychology*, *89*, 191-204. https://doi.org/10.1111/j.2044-8295.1998.tb02680.x

Kloo, D., Kristen-Antonow, S., & Sodian, B. (2020). Progressing from an implicit to an explicit false belief understanding: A matter of executive control? *International Journal of Behavioral Development*, *44*(2), 107-115. https://doi.org/10.1177/0165025419850901

Kristen, S., Chiarella, S., Sodian, B., Aureli, T., Genco, M., & Poulin-Dubois, D. (2014). Crosslinguistic developmental consistency in the composition of toddlers' internal state vocabulary: Evidence from four languages. *Child Development Research*, *2014*, 575142. https://doi.org/ 10.1155/2014/575142

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. https://doi.org/10.1002/icd.1767

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. https://doi.org/10.1026/0049-8637.38.4.186

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. https://doi.org/10.1080/15248372.2019.1586710

Kulke, L., & Rakoczy, H. (2018). Implicit Theory of Mind: An overview of current replications and non-replications. *Data in Brief*, *16*, 101-104. https://doi.org/10.1016/j.dib.2017.11.016

Kulke, L., & Rakoczy, H. (2019). Testing the role of verbal narration in implicit theory of mind tasks. *Journal of Cognition and Development*, *20*(1), 1-14. https://doi.org/10.1080/15248372.2018.1544140

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. https://doi.org/10.1371/journal.pone.0213772

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. https://doi.org/10. 1016/j.cogdev.2017.09.001

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. https://doi.org/ 10.1177/0956797617747090

Lawson, J., Baron-Cohen, S., & Wheelwright, S. (2004). Empathising and systemising in adults with and without Asperger syndrome. *Journal of Autism and Developmental Disorders*, *34*(3), 301-310. https://doi.org/ 10.1023/b:jadd.0000029552.42724.1b

Lecce, S., Caputi, M., & Hughes, C. (2011). Does sensitivity to criticism mediate the relationship between theory of mind and academic achievement? *Journal of Experimental Child Psychology*, *110*(3), 313-331. https://doi.org/10.1016/j.jecp.2011.04.011

Lecce, S., Caputi, M., & Pagnin, A. (2014). Long-term effect of theory of mind on school achievement: The role of sensitivity to criticism. *European Journal of Developmental Psychology*, *11*(3), 305-318. https://doi.org/10.1080/17405629.2013.821944

Lecce, S., Demicheli, P., Zocchi, S., & Palladino, P. (2015). The origins of children's metamemory: The role of theory of mind. *Journal of Experimental Child Psychology*, *131*, 56-72. https://doi.org/10.1016/j.jecp.2014.11.005

Leslie, A. M. (1994). Pretending and believing: Issues in the theory of ToMM. *Cognition*, *50*(1-3), 211-238. https://doi.org/10.1016/0010-0277(94)90029-9

Lind, S. E., & Bowler, D. M. (2009). Language and theory of mind in autism spectrum disorder: The relationship between complement syntax and false belief task performance. *Journal of Autism and Developmental Disorders*, *39*(6), 929-937. https://doi.org/ 10.1007/s10803-009-0702-y

Lockl, K., & Schneider, W. (2007). Knowledge about the mind: Links between theory of mind and later metamemory. *Child Development*, *78*(1), 148-167. https://doi.org/10.1111/j.1467-8624.2007.00990.x

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. https://doi.org/10.1111/1467-8624.00597

Longobardi, E., Lonigro, A., & Laghi, F. (2016). References to mental states in mother-child conversation in the second year of life. *Journal of Child and Family Studies*, *25*(3), 756-766. https://doi.org/ 10.1007/s10826-015-0255-y

Low, J. (2010). Preschoolers' implicit and explicit false-belief understanding: Relations with complex syntactical mastery. *Child Development*, *81*(2), 597-615. https://doi.org/10.1111/j.1467-8624.2009.01418.x

Low, J., & Watts, J. (2013). Attributing false beliefs about object identity reveals a signature blind spot in humans' efficient mind-reading system. *Psychological Science*, *24*(3), 305-311. https://doi.org/10.1177/0956797612451469

Low, J., Apperly, I. A., Butterfill, S. A., & Rakoczy, H. (2016). Cognitive architecture of belief reasoning in children and adults: A primer on the two-systems account. *Child Development Perspectives*, *10*(3), 184-189. https://doi.org/10.1111/cdep.12183

Magee, M., Lewis, C., Noffs, G., Reece, H., Chan, J. C. S., Zaga, C. J., Paynter, C., Birchall, O., Rojas Azocar, S., Ediriweera, A., Kenyon, K., Caverlé, M. W., Schultz, B. G., & Vogel, A. P. (2020). Effects of face masks on acoustic analysis and speech perception: Implications for peri-pandemic protocols. *The Journal of the Acoustical Society of America*, *148*(6), 3562-3568. https://doi.org/10.1121/10.0002873

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. https://doi.org/10.2307/1127956

McLoughlin, N., Leech, K. A., Chernyak, N., Blake, P. R., & Corriveau, K. H. (2020). Conflicting perspectives mediate the relation between parents' and preschoolers' self-referent mental state talk during collaboration. *British Journal of Developmental Psychology*, *38*(2), 255-267. https://doi.org/10.1111/bjdp.12316

Meins, E., & Fernyhough, C. (1999). Linguistic acquisitional style and mentalising development: The role of maternal mind-mindedness. *Cognitive Development*, *14*(3), 363-380. https://doi.org/10.1016/S0885-2014(99)00010-6

Meristo, M., Falkman, K. W., Hjelmquist, E., Tedoldi, M., Surian, L., & Siegal, M. (2007). Language access and theory of mind reasoning: Evidence from deaf children in bilingual

and oralist environments. *Developmental Psychology*, *43*(5), 1156-1169. https://doi.org/ 10.1037/0012-1649.43.5.1156

Meristo, M., Morgan, G., Geraci, A., Iozzi, L., Hjelmquist, E., Surian, L., & Siegal, M. (2012). Belief attribution in deaf and hearing infants. *Developmental Science*, *15*(5), 633-640. https://doi.org/10.1111/j.1467-7687.2012.01155.x

Meristo, M., Strid, K., & Hjelmquist, E. (2016). Early conversational environment enables spontaneous belief attribution in deaf children. *Cognition*, *157*, 139-145. https://doi.org/10.1016/j.cognition.2016.08.023

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. https://doi.org/10.1111/j.1467-8624.2007.01018.x

Mitchell, P., & Lacohée, H. (1991). Children's early understanding of false belief. *Cognition*, *39*(2), 107-127. https://doi.org/10.1016/0010-0277(91)90040-B

Mo, S., Su, Y., Sabbagh, M. A., & Jiaming, X. (2014). Sentential complements and false belief understanding in Chinese Mandarin-speaking preschoolers: A training study. *Cognitive Development*, *29*, 50-61. https://doi.org/10.1016/j.cogdev.2013.11.001

Moll, H., & Tomasello, M. (2006). Level 1 perspective-taking at 24 months of age. *British Journal of Developmental Psychology*, *24*(3), 603-613. https://doi.org/10.1348/026151005X55370

Montgomery, D. E. (2002). Mental verbs and semantic development. *Journal of Cognition and Development*, *3*(4), 357-384. https://doi.org/10.1207/S15327647JCD3,4-01

Moore, C., Bryant, D., & Furrow, D. (1989). Mental terms and the development of certainty. *Child Development*, *60*(1), 167-171. https://doi.org/10.2307/1131082

Moore, C., Furrow, D., Chiasson, L., & Patriquin, M. (1994). Developmental relationships between production and comprehension of mental terms. *First Language*, *14*(42-43), 001-017. https://doi.org/10.1177/014272379401404201

Moore, C., Pure, K., & Furrow, D. (1990). Children's understanding of the modal expression of speaker certainty and uncertainty and its relation to the development of a representational theory of mind. *Child Development*, *61*(3), 722-730. https://doi.org/10.1111/j.1467-8624.1990.tb02815.x

Morford, J. P., & Goldin-Meadow, S. (1997). From here and now to there and then: The development of displaced reference in homesign and English. *Child Development*, *68*(3), 420-435. https://doi.org/10.1111/j.1467-8624.1997.tb01949.x

Ng, L., Cheung, H., & Xiao, W. (2010). False belief, complementation language, and contextual bias in preschoolers. *International Journal of Behavioral Development*, *34*(2), 168-179. https://doi.org/10.1177/0165025409350954

Nicoladis, E., & Genesee, F. (1997). Language development in preschool bilingual children. *Journal of Speech-Language Pathology and Audiology*, *21*(4), 258-270. https://doi.org/10.7939/R3348GW0N

Nielsen, M., & Dissanayake, C. (2000). An investigation of pretend play, mental state terms and false belief understanding: In search of a metarepresentational link. *British Journal of Developmental Psychology*, *18*(4), 609-624. https://doi.org/10.1348/026151000165887

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. https://doi.org/10.1111/j.1467-8624.1996.tb01758.x

O'Neill, D. K. (2005). Talking about "new" information: The given/new distinction and children's developing theory of mind. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 84-105). New York, NY: Oxford University Press. https://doi.org/ 10.1093/acprof:oso/9780195159912.003.0005

O'Neill, D. K., & Gopnik, A. (1991). Young children's ability to identify the sources of their beliefs. *Developmental Psychology*, *27*(3), 390-397. https://doi.org/10.1037/0012-1649.27.3.390

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. https://doi.org/10.1080/17405620601046931

Olson, D. R. (1988). On the origins of beliefs and other intentional states in children. In J. W. Astington, P. L. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (pp. 414-426). Cambridge, UK: Cambridge University Press.

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

Osterhaus, C., & Bosacki, S. L. (2022). Looking for the lighthouse: A systematic review of advanced theory-of-mind tests beyond preschool. *Developmental Review*, *64*, 101021. https://doi.org/10.1016/j.dr.2022.101021

Osterhaus, C., & Koerber, S. (2021). The development of advanced theory of mind in middle childhood: A longitudinal study from age 5 to 10 years. *Child Development*, *92*(5), 1872-1888. https://doi.org/10.1111/cdev.13627

Pan, B. A., & Snow, C. E. (1999). The development of conversational and discourse skills. In M. D. Barrett (Ed.), *The development of language* (pp. 229-249). Hove, United Kingdom: Psychology Press.

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. https://doi.org/10.1111/j.1467-7687.2008.00691.x

Paulus, M., & Sabbagh, M. A. (2018). Understanding theory of mind in infancy and toddlerhood [Special Issue]. *Cognitive Development*, *46*.

Paulus, M., Schuwerk, T., Sodian, B., & Ganglmayer, K. (2017). Children's and adults' use of verbal information to visually anticipate others' actions: A study on explicit and implicit social-cognitive processing. *Cognition*, *160*, 145-152. https://doi.org/10.1016/j.cognition.2016.12.013

Perner, J. (2010). Who took the cog out of cog*nitive* science? Mentalism in an era of anti-cognitivism. In P. A. Frensch & R. Schwarzer (Eds.), *Cognition and neuropsychology: International perspectives on psychological science* (pp. 241-261). Hove, United Kingdom: Psychology Press.

Perner, J., & Roessler, J. (2012). From infants' to children's appreciation of belief. *Trends in Cognitive Sciences*, *16*(10), 519-525. https://doi.org/10.1016/j.tics.2012.08.004

Perner, J., & Ruffman, T. (2005). Infants' insight into the mind: How deep?. *Science*, *308*(5719), 214-216. https://doi.org/10.1126/science.1111656

Perner, J., & Wimmer, H. (1985). "John *thinks* that Mary *thinks* that…" Attribution of second-order beliefs by 5-to 10-year-old children. *Journal of Experimental Child Psychology*, *39*(3), 437-471. https://doi.org/10.1016/0022-0965(85)90051-7

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. https://doi.org/10.1111/j.2044-835X.1987.tb01048.x

Perner, J., Ruffman, T., & Leekam, S. R. (1994). Theory of mind is contagious: You catch it from your sibs. *Child Development*, *65*(4), 1228-1238. https://doi.org/10.1111/j.1467-8624.1994.tb00814.x

Perner, J., Sprung, M., Zauner, P., & Haider, H. (2003). *Want that* is understood well before *say that*, *think that*, and false belief: A test of de Villiers's linguistic determinism on German-speaking children. *Child Development*, *74*(1), 179-188. https://doi.org/10.1111/1467-8624.t01-1-00529

Peterson, C. C., & Siegal, M. (2000). Insights into theory of mind from deafness and autism. *Mind & Language*, *15*(1), 123-145. https://doi.org/10.1111/1468-0017.00126

Peterson, C. C., & Slaughter, V. (2003). Opening windows into the mind: Mothers' preferences for mental state explanations and children's theory of mind. *Cognitive Development*, *18*(3), 399-429. https://doi.org/10.1016/S0885-2014(03)00041-8

Peterson, C. C., O'Reilly, K., & Wellman, H. M. (2016). Deaf and hearing children's development of theory of mind, peer popularity, and leadership during middle childhood. *Journal of Experimental Child Psychology*, *149*, 146-158. https://doi.org/10.1016/j.jecp.2015.11.008

Phillips, A. T., & Wellman, H. M. (2005). Infants' understanding of object-directed action. *Cognition*, *98*(2), 137-155. https://doi.org/10.1016/j.cognition.2004.11.005

Pillow, B. H. (1989). Early understanding of perception as a source of knowledge. *Journal of Experimental Child Psychology*, *47*(1), 116-129. https://doi.org/10.1016/0022-0965(89)90066-0

Pons, F., & Harris, P. L. (2000). *Test of Emotion Comprehension–TEC*. Oxford, United Kingdom: Oxford University Press.

Poulin-Dubois, D., & Yott, J. (2018). Probing the depth of infants' theory of mind: Disunity in performance across paradigms. *Developmental Science*, *21*(4), e12600. https://doi.org/10.1111/desc.12600

Poulin-Dubois, D., Azar, N., Elkaim, B., & Burnside, K. (2020). Testing the stability of theory of mind: A longitudinal approach. *PloS ONE*, *15*(11), e0241721. https://doi.org/10.1371/journal.pone.0241721

Poulin-Dubois, D., Rakoczy, H., Burnside, K., Crivello, C., Dörrenberg, S., Edwards, K., Krist, H., Kulke, L., Liszkowski, U., Low, J., Perner, J., Powell, L., Priewasser, B., Rafetseder, E., & Ruffman, T. (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. https://doi.org/10.1016/j.cogdev.2018.09.005

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. https://doi.org/10.1016/j.cogdev.2017.10.004

Pratt, C., & Bryant, P. (1990). Young children understand that looking leads to knowing (so long as they are looking into a single barrel). *Child Development*, *61*(4), 973-982. https://doi.org/10.1111/j.1467-8624.1990.tb02835.x

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*, 515-526. https://doi.org/10. 1017/S0140525X00076512

Priewasser, B., Fowles, F., Schweller, K., & Perner, J. (2020). Mistaken Max befriends Duplo girl: No difference between a standard and an acted-out false belief task. *Journal of Experimental Child Psychology*, *191*, 104756. https://doi.org/10.1016/j.jecp.2019.104756

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. https://doi.org/10.1016/j.cogdev.2017.08.002

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. https://doi.org/10.1016/j.cognition.2018.06.017

Rohlfing, K. J. (2019). *Frühe Sprachentwicklung.* Tübingen, Germany: Narr.

Rubio-Fernández, P. (2019). Publication standards in infancy research: Three ways to make Violation-of-Expectation studies more reliable. *Infant Behavior and Development*, *54*, 177-188. https://doi.org/10.1016/j.infbeh.2018.09.009

Rubio-Fernández, P., & Geurts, B. (2013). How to pass the false-belief task before your fourth birthday. *Psychological Science*, *24*(1), 27-33. https://doi.org/10.1177/0956797612447819

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. https://doi.org/10.1007/s13164-015-0290-z

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. https://doi.org/10.1073/pnas.1701286114

Ruffman, T. (2014). To belief or not belief: Children's theory of mind. *Developmental Review*, *34*(3), 265-293. https://doi.org/10.1016/j.dr.2014.04.001

Ruffman, T., & Olson, D. R. (1989). Children's ascriptions of knowledge to others. *Developmental Psychology*, *25*(4), 601-606. https://doi.org/10.1037/0012-1649.25.4.601

Ruffman, T., & Perner, J. (2005). Do infants really understand false belief? Response to Leslie. *Trends in Cognitive Sciences*, *9*(10), 462-463. https://doi.org/10.1016/j.tics.2005.08.001

Ruffman, T., Garnham, W. A., 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. https://doi.org/10.1006/jecp.2001.2633

Ruffman, T., Perner, J., & Parkin, L. (1999). How parenting style affects false belief understanding. *Social Development*, *8*(3), 395-411. https://doi.org/10.1111/1467-9507.00103

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. https://doi.org/10.1111/1467-8624.00435

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. https://doi.org/10.1016/S0885-2014(03)00002-9

San Juan, V., & Astington, J. W. (2012). Bridging the gap between implicit and explicit understanding: How language development promotes the processing and representation of false belief. *British Journal of Developmental Psychology*, *30*(1), 105-122. https://doi.org/10.1111/j.2044-835X.2011.02051.x

San Juan, V., & Astington, J. W. (2017). Does language matter for implicit theory of mind? The effects of epistemic verb training on implicit and explicit false-belief understanding. *Cognitive Development*, *41*, 19-32. https://doi.org/10.1016/j.cogdev.2016.12.003

Schick, B., de Villiers, P., de Villiers, J. G., & Hoffmeister, R. (2007). Language and theory of mind: A study of deaf children. *Child Development*, *78*(2), 376-396. https://doi.org/10.1111/j.1467-8624.2007.01004.x

Schneider, D., Bayliss, A. P., Becker, S. I., & Dux, P. E. (2012). Eye movements reveal sustained implicit processing of others' mental states. *Journal of Experimental Psychology: General*, *141*(3), 433-438. https://doi.org/10.1037/a0025458

Schneider, D., Lam, R., Bayliss, A. P., & Dux, P. E. (2012). Cognitive load disrupts implicit theory-of-mind processing. *Psychological Science*, *23*(8), 842-847. https://doi.org/10.1177/0956797612439070

Schneider, J., Sandoz, V., Equey, L., Williams-Smith, J., Horsch, A., & Graz, M. B. (2022). The role of face masks in the recognition of emotions by preschool children. *JAMA Pediatrics*, *176*(1), 96-98. https://doi.org/10.1001/jamapediatrics.2021.4556

Scholl, B. J., & Leslie, A. M. (2001). Minds, modules, and meta-analysis. *Child Development*, *72*(3), 696-701. https://doi.org/10.1111/1467-8624.00308

Schuwerk, T., Jarvers, I., Vuori, M., & Sodian, B. (2016). Implicit mentalizing persists beyond early childhood and is profoundly impaired in children with autism spectrum condition. *Frontiers in Psychology*, *7*, 1696. https://doi.org/10.3389/fpsyg.2016.01696

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

Schuwerk, T., Priewasser, B., Sodian, B., & Perner, J. (2018). The robustness and generalizability of findings on spontaneous false belief sensitivity: A replication attempt. *Royal Society Open Science*, *5*(5), 172273. https://doi.org/10.1098/rsos.172273

Schuwerk, T., Vuori, M., & Sodian, B. (2015). Implicit and explicit theory of mind reasoning in autism spectrum disorders: The impact of experience. *Autism*, *19*(4), 459-468. https://doi.org/10.1177/1362361314526004

Scott, R. M. (2014). Post hoc versus predictive accounts of children's theory of mind: A reply to Ruffman. *Developmental Review*, *34*(3), 300-304. https://doi.org/10.1016/j.dr.2014.05.001

Scott, R. M. (2017). The developmental origins of false-belief understanding. *Current Directions in Psychological Science*, *26*(1), 68-74. https://doi.org/10.1177/0963721416673174

Scott, R. M., & Baillargeon, R. (2009). Which penguin is this? Attributing false beliefs about object identity at 18 months. *Child Development*, *80*(4), 1172-1196. https://doi.org/10.1111/j.1467-8624.2009.01324.x

Scott, R. M., & Baillargeon, R. (2017). Early false-belief understanding. *Trends in Cognitive Sciences*, *21*(4), 237-249. https://doi.org/10.1016/j.tics.2017.01.012

Scott, R. M., Roby, E., & Setoh, P. (2020). 2.5-year-olds succeed in identity and location elicited-response false-belief tasks with adequate response practice. *Journal of Experimental Child Psychology*, *198*, 104890. https://doi.org/10.1016/j.jecp.2020.104890

Scott, R. M., Setoh, P., & Baillargeon, R. (2017). Reply to Rubio-Fernández et al.: Different traditional false-belief tasks impose different processing demands for toddlers. *Proceedings of the National Academy of Sciences*, *114*(19), E3751-E3752. https://doi.org/10.1073/pnas.1703665114

Sellabona, E. S., Sánchez, C. R., Majoral, E. V., Guitart, M. E., Caballero, F. S., & Ortiz, J. S. (2013). Labelling improves false belief understanding: A training study. *The Spanish Journal of Psychology*, *16*, e6. https://doi.org/10.1017/sjp.2013.8

Senju, A., Southgate, V., Miura, Y., Matsui, T., Hasegawa, T., Tojo, Y., Osanai, H., & Csibra, G. (2010). Absence of spontaneous action anticipation by false belief attribution in children with autism spectrum disorder. *Development and Psychopathology*, *22*(2), 353-360. https://doi.org/10.1017/ S0954579410000106

Senju, A., Southgate, V., Snape, C., Leonard, M., & Csibra, G. (2011). Do 18-month-olds really attribute mental states to others? A critical test. *Psychological Science*, *22*(7), 878-880. https://doi.org/10.1177/ 0956797611411584

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. https://doi.org/10.1073/pnas.1609203113

Setoh, P., Scott, R. M., & Baillargeon, R. (2020). Reply to Fenici and Garofoli: Why would toddlers act on low-level associations only when processing demands are reduced? *Human Development*, *64*, 7-9. https://doi.org/10.1159/000506805

Shatz, M., Wellman, H. M., & Silber, S. (1983). The acquisition of mental verbs: A systematic investigation of the first reference to mental state. *Cognition*, *14*(3), 301-321. https://doi.org/10.1016/0010-0277(83)90008-2

Shwe, H. I., & Markman, E. M. (1997). Young children's appreciation of the mental impact of their communicative signals. *Developmental Psychology*, *33*(4), 630-636. https://doi.org/10.1037/0012-1649.33.4.630

Siegal, M., & Beattie, K. (1991). Where to look first for children's knowledge of false beliefs. *Cognition*, *38*(1), 1-12. https://doi.org/10.1016/0010- 0277(91)90020-5

Slade, L., & Ruffman, T. (2005). How language does (and does not) relate to theory of mind: A longitudinal study of syntax, semantics, working memory and false belief. *British Journal of Developmental Psychology*, *23*(1), 117-141. https://doi.org/10.1348/026151004X21332

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. https://doi.org/10.1111/cdev.12372

Slaughter, V., Peterson, C. C., & Mackintosh, E. (2007). Mind what mother says: Narrative input and theory of mind in typical children and those on the autism spectrum. *Child Development*, *78*(3), 839-858. https://doi.org/10.1111/j.1467-8624.2007.01036.x

Snow, C. E., Pan, B. A., Imbens-Bailey, A., & Herman, J. (1996). Learning how to say what one means: A longitudinal study of children's speech act use. *Social Development*, *5*(1), 56-84. https://doi.org/10.1111/j.1467-9507.1996.tb00072.x

Sodian, B., Kristen-Antonow, S., & Kloo, D. (2020). How does children's theory of mind become explicit? A review of longitudinal findings. *Child Development Perspectives*, *14*(3), 171-177. https://doi.org/10.1111/cdep.12381

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. https://doi.org/10.1111/cdev.12533

Song, H. J., & Baillargeon, R. (2008). Infants' reasoning about others' false perceptions. *Developmental Psychology*, *44*(6), 1789-1795. https://doi.org/10.1037/a0013774

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, *18*(7), 587-592. https://doi.org/10.1111/j.1467-9280.2007.01944.x

Surian, L., & Geraci, A. (2012). Where will the triangle look for it? Attributing false beliefs to a geometric shape at 17 months. *British Journal of Developmental Psychology*, *30*(1), 30-44. https://doi.org/10.1111/j.2044-835X.2011.02046.x

Surian, L., & Leslie, A. M. (1999). Competence and performance in false belief understanding: A comparison of autistic and normal 3-year-old children. *British Journal of Developmental Psychology*, *17*(1), 141-155. https://doi.org/10.1348/026151099165203

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science*, *18*(7), 580-586. https://doi.org/10.1111/j.1467-9280.2007.01943.x

Symons, D. K. (2004). Mental state discourse, theory of mind, and the internalization of self-other understanding. *Developmental Review*, *24*(2), 159-188. https://doi.org/10.1016/j.dr.2004.03.001

Tager-Flusberg, H. (2000). Language and understanding minds: Connections in autism. In S. Baron-Cohen, H. Tager-Flusberg & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience* (pp. 124-149). New York, NY: Oxford University Press.

Tager-Flusberg, H., & Joseph, R. M. (2005). How language facilitates the acquisition of false-belief understanding in children with autism. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 298-318). New York, NY: Oxford University Press. https://doi.org/ 10.1093/acprof:oso/9780195159912.003.0014

Takacs, Z. K., Swart, E. K., & Bus, A. G. (2014). Can the computer replace the adult for storybook reading? A meta-analysis on the effects of multimedia stories as compared to sharing print stories with an adult. *Frontiers in Psychology*, *5*, 1366. https://doi.org/10.3389/fpsyg.2014.01366

Tardif, T., & Wellman, H. M. (2000). Acquisition of mental state language in Mandarin- and Cantonese-speaking children. *Developmental Psychology*, *36*(1), 25-43. https://doi.org/10.1037/0012-1649.36.1.25

Taumoepeau, M., & Ruffman, T. (2006). Mother and infant talk about mental states relates to desire language and emotion understanding. *Child Development*, *77*(2), 465-481. https://doi.org/10.1111/j.1467-8624.2006.00882.x

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. https://doi.org/10.1111/j.1467-8624.2007.01126.x

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. https://doi.org/10.1111/j.2044-835X.2011.02067.x

Tomasello, M. (2018). How children come to understand false beliefs: A shared intentionality account. *Proceedings of the National Academy of Sciences*, *115*(34), 8491-8498. https://doi.org/10.1073/pnas.1804761115

Tompkins, V. (2015). Improving low-income preschoolers' theory of mind: A training study. *Cognitive Development*, *36*, 1-19. https://doi.org/10.1016/j.cogdev.2015.07.001

Tompkins, V., Benigno, J. P., Kiger Lee, B., & Wright, B. M. (2018). The relation between parents' mental state talk and children's social understanding: A meta-analysis. *Social Development*, *27*(2), 223-246. https://doi.org/10.1111/sode.12280

Träuble, B., Marinović, V., & Pauen, S. (2010). Early theory of mind competencies: Do infants understand others' beliefs? *Infancy*, *15*(4), 434-444. https://doi.org/10.1111/j.1532-7078.2009.00025.x

Tsuji, H. (2020). Perspective-shifting discourse training to improve young Japanese children's understanding of theory of mind. *Infant and Child Development*, *29*(4), e2178. https://doi.org/10.1002/icd.2178

Valle, A., Massaro, D., Castelli, I., & Marchetti, A. (2015). Theory of mind development in adolescence and early adulthood: The growing complexity of recursive thinking ability. *Europe's Journal of Psychology*, *11*(1), 112-124. https://doi.org/10.5964/ejop.v11i1.829

Wagner, L. (2014). Sprachentwicklung und Sprachentwicklungsstörungen im Kontext von Mehrsprachigkeit. *Kinder- und Jugendmedizin*, *14*(5), 331-335. https://doi.org/10.1055/s-0038-1629232

Walter-Laager, C., Brandenberg, K., Tinguely, L., Moschner, B., Schwarz, J., & Pfiffner, M. R. (2016). Interesse von Kleinkindern an unterschiedlichen Sprachlernmedien. *Frühe Bildung*, *5*(1), 40-49. https://doi.org/10.1026/2191-9186/a000243

Wang, B., Low, J., Jing, Z., & Qinghua, Q. (2012). Chinese preschoolers' implicit and explicit false-belief understanding. *British Journal of Developmental Psychology*, *30*(1), 123-140. https://doi.org/10.1111/j.2044-835X.2011.02052.x

Wang, L. U., & Leslie, A. M. (2016). Is implicit theory of mind the 'Real Deal'? The own-belief/true-belief default in adults and young preschoolers. *Mind & Language*, *31*(2), 147-176. https://doi.org/10.1111/mila.12099

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, *311*(5765), 1301-1303. https://doi.org/10.1126/science.1121448

Warnell, K. R., & Redcay, E. (2019). Minimal coherence among varied theory of mind measures in childhood and adulthood. *Cognition*, *191*, 103997. https://doi.org/10.1016/j.cognition.2019.06.009

Watson, A. C., Painter, K. M., & Bornstein, M. H. (2001). Longitudinal relations between 2-year-olds' language and 4-year-olds' theory of mind. *Journal of Cognition and Development*, *2*(4), 449-457. https://doi.org/10.1207/S15327647JCD0204_5

Welch-Ross, M. K. (1997). Mother-child participation in conversation about the past: Relationships to preschoolers' theory of mind. *Developmental Psychology*, *33*(4), 618-629. https://doi.org/10.1037/0012-1649.33.4.618

Wellman, H. M. (1978). Knowledge of the interaction of memory variables: A developmental study of metamemory. *Developmental Psychology*, *14*(1), 24-29. https://doi.org/10.1037/0012-1649.14.1.24

Wellman, H. M. (2011). Developing a Theory of Mind. In U. Goswami (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development* (2nd ed., pp. 258-284). New York, NY: Blackwell-Wiley. https://doi.org/ 10.1002/9781444325485

Wellman, H. M. (2018). Theory of mind: The state of the art. *European Journal of Developmental Psychology*, *15*(6), 728-755. https://doi.org/10.1080/17405629.2018.1435413

Wellman, H. M., & Gelman, S. A. (1998). Knowledge acquisition in foundational domains. In D. Kuhn & R. S. Siegler (Eds.), W. Damon (Series Ed.), *Handbook of child psychology: Vol 2. Cognition, perception, and language* (pp. 523-573). New York, NY: Wiley

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, *75*(2), 523-541. https://doi.org/10.1111/j.1467-8624.2004.00691.x

Wellman, H. M., & Woolley, J. D. (1990). From simple desires to ordinary beliefs: The early development of everyday psychology. *Cognition*, *35*(3), 245-275. https://doi.org/10.1016/0010-0277(90)90024-E

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. https://doi.org/10.1111/1467-8624.00304

Wellman, H. M., Lopez-Duran, S., LaBounty, J., & Hamilton, B. (2008). Infant attention to intentional action predicts preschool theory of mind. *Developmental Psychology*, *44*, 618-623. https://doi.org/10. 1037/0012-1649.44.2.618

Wellman, H. M., Phillips, A. T., Dunphy-Lelii, S., & LaLonde, N. (2004). Infant social attention predicts preschool social cognition. *Developmental Science*, *7*, 283-288. https://doi.org/10.1111/j.1467- 7687.2004.00347.x

Wimmer, H., & Mayringer, H. (1998). False belief understanding in young children: Explanations do not develop before predictions. *International Journal of Behavioral Development*, *22*(2), 403-422. https://doi.org/10.1080/016502598384441

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. https://doi.org/10.1016/0010-0277(83)90004-5

Wimmer, H., Hogrefe, G. J., & Perner, J. (1988). Children's understanding of informational access as source of knowledge. *Child Development*, *59*(2), 386-396. https://doi.org/10.2307/1130318

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*, 1-34. https://doi.org/10.1016/S0010- 0277(98)00058-4

Yamaguchi, M., Kuhlmeier, V. A., Wynn, K., & VanMarle, K. (2009). Continuity in social cognition from infancy to childhood. *Developmental Science*, *12*, 746-752. https://doi.org/10.1111/j.1467-7687.2008.00813.x

Yazdi, A. A., German, T. P., Defeyter, M. A., & Siegal, M. (2006). Competence and performance in belief-desire reasoning across two cultures: The truth, the whole truth and nothing but the truth about false belief?. *Cognition*, *100*(2), 343-368. https://doi.org/10.1016/j.cognition.2005.05.004

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. https://doi.org/10.1016/j.cogdev.2019.05.002

Yott, J., & Poulin-Dubois, D. (2012). Breaking the rules: Do infants have a true understanding of false belief? *British Journal of Developmental Psychology*, *30*(1), 156-171. https://doi.org/10.1111/j.2044-835X.2011.02060.x

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. https://doi.org/10.1080/15248372.2015.1086771

Appendix

Appendix A: Translations of the Example Sentences

English Translations of the German Example Sentences

German Example Sentence	English Translation	Page Number
Ich glaube, der Osterhase **versteckt** morgen Schokolade.	I think the Easter Bunny **will hide** chocolate tomorrow.	p. 33
Ich glaube, *dass* der Osterhase morgen Schokolade **versteckt**.	I think *that* the Easter Bunny **will hide** chocolate tomorrow.	p. 33

German Translations of the English Example Sentences

English Example Sentence	German Translation	Page Number
I think that the sun will shine tomorrow.	Ich denke, dass die Sonne morgen scheint.	p. 31
You don't know where the chocolate is.	Du weißt nicht, wo die Schokolade ist.	p. 32
We asked if they want soup for dinner.	Wir haben gefragt, ob sie Suppe zum Abendessen möchten.	p. 32
I suggest that we leave before it begins to rain.	Ich schlage vor, dass wir gehen, bevor es anfängt zu regnen.	p. 32
I promise that I will help you with this work.	Ich verspreche, dass ich dir mit dieser Arbeit helfen werde.	p. 32
We better leave before it begins to rain.	Wir gehen besser bevor es anfängt zu regnen.	p. 32
Look what I can do.	Schau was ich kann.	p. 32
I am pretending I am princess.	Ich tue so als wäre ich eine Prinzessin.	p. 33
Look Mum, I flipped over.	Schau, Mama, ich bin umgekippt.	p. 35
Anna thought that zebras were purple.	Anna dachte, dass Zebras lila seien.	p. 37
Anna thinks that zebras are black and white.	Anna denkt, dass Zebras schwarzweiß sind.	p. 36
Anna is crying because zebras are purple.	Anna weint, weil Zebras lila sind.	p. 37
Anna believed that Peter thought that Lisa was home.	Anna glaubte, dass Peter dachte, dass Lisa zuhause wäre.	p. 37
(Jack thinks) his food is in the fridge.	(Jack denkt,) sein Essen ist im Kühlschrank.	p. 37
Jack said the dog is in the kitchen.	Jack sagte, der Hund ist in der Küche.	p. 37
He thought he found his ring but it was really a bottle cap.	Er dachte, er hätte seinen Ring gefunden, aber es war tatsächlich ein Flaschenverschluss.	p. 38
What did he say what he drank?	Was hat er gesagt, was er getrunken hat?	p. 190
Don't know, think so.	Weiß nicht, glaub schon.	p. 193
I don't know where my teddy bear is. I think I lost it at the doctor's office.	Ich weiß nicht, wo mein Teddybär ist. Ich glaube ich habe ihn in der Arztpraxis verloren.	p. 193

Appendix B: Procedures of the Experimental Tasks

Procedure of the Low-Demands False Belief Task (Articles 1 and 2)

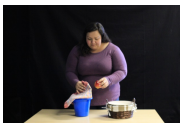
Material:

- Picture book with the story, practice, and test trial pictures bound in the correct order

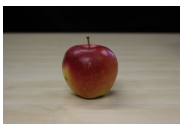
Procedure:
Place the bound storybook on the table in front of the child and stand next to the book. Always turn the page to reveal the next picture and then recite the corresponding story line.

Story Trial 1:
"This is a story about Lilly. Look, this is Lilly."

Story Trial 2:
"Lilly finds an apple in a bucket."



Practice Trial 1:
"Where is Lilly's apple?"




Story Trial 3:
"Lilly puts her apple in a basket for later."

Story Trial 4:
"Then, she goes outside to play with a ball."

Practice Trial 2:
"Where is Lilly's ball?"


Story Trial 5:
"While Lilly is outside, her brother Peter finds the apple and takes it away."

Story Trial 6:
"Lilly is hungry. She comes back inside to look for her apple."

Test Trial:
"Where will Lilly look for her apple?"


Procedure of the Explicit False Belief Tasks (Article 1)

Contents False-Belief:

Material:

- Typical Smarties tube (without Smarties)
- Little toy pig that fits into the Smarties tube, hidden inside the Smarties tube
- Playmobil figurine called Lukas, hidden underneath the table or in the experimenter's pocket

Procedure:

Show the child the Smarties tube and ask: "Look what I have! What do you think is in this Smarties tube?"
Child should reply with 'Smarties'. If not, make the child reply with 'Smarties' by asking, for instance, "Does this look as if there are Smarties inside?" or "What kind of box is this? What is normally inside such a box?" or "Should there be Smarties inside or books?"

Then, slowly start opening the tube and say: "Ok, then let's take a look … oh, there is actually a PIG inside the tube!".
Take out the pig and show it to the child. After a few seconds, put the pig back into the tube and close the tube's lid.

Then ask the child: "Okay ... what was in the box?"

If the child fails this question, show the content again, close the box again, and ask the question again until the child replies correctly (i.e., with pig).

Fetch Lukas and put him next to the Smarties tube on the table such that Lukas is facing the child. Then say: "This is Lukas. Lukas has never ever seen what's in this box!"

Then, move Lukas towards the box such that Lukas is now facing the tube and say: "So, here comes Lukas."

Then ask:

Test question:
"So ... what does Lukas think what is in the box? Smarties or a pig?"

Control question:
"Has Lukas ever looked into this box?"

Explicit False-Belief:

Material:

- Laminated sheet with a picture of a backpack and a picture of a closet
- Playmobil figurine called Paul, hidden underneath the table or in the experimenter's pocket

Procedure:

Take the sheet with the pictures and put it on the table. Show the child the figurine Paul and say: "This is Paul." Place Paul next to the sheet such that Paul is approximately in the middle between the two objects. Then say: "Paul is looking for his gloves. Paul's gloves could either be in his backpack (point at the picture) or in his closet (point at the picture)."

[In case the child makes their own assumption at this point (e.g., "I think they are in his back-pack/closet"), shortly take up the assumption: "Oh that's a good idea … BUT … really…" or: "Oh, that's a good idea … AND … really …"]

Otherwise, just continue with the normal procedure.

Say: "So ... really, Paul's gloves are in his backpack (point at the picture and make a short pause) – but Paul thinks that his gloves are in his closet (point at the picture and make a short pause)."

Now ask:

Test question:
"So, where will Paul (point at Paul) now look for his gloves? In his backpack or in his closet?"

Control question:
"And where are Paul's gloves really? In his backpack or in his closet?"

Procedure of the False Belief Picture Book Task (Article 2)

Material:

- Picture book „Can't sleep" by K. Moerbeek at 33 months and picture book "Who's peeking at me?" by K. Moerbeek at 36 months
- Playmobil figurine called Jonas (hidden underneath the table or in the experimenter's pocket)

Procedure:
First, ask the parents: "Does [name of the child] know this book?"
Then, turn to the child:
"Look what I have here! A book! I want to look at the book with you."
Show the child the cover page of the book.

Open up the first page of the book.
On each new page of the book, read the text in the book. Then, point at the eye depicted on the right-hand side of the double page and ask:
"Look, what's that?"

Wait for the child's reply, if the child does not reply or says something else than eye, say:
"Oh, I think that's an eye! Let's take a look on the next page."
If child assumes it's an eye, say:
"Yeah, I also think that's an eye! Let's take a look on the next page."

Turn the page and point at one of the eyes of the new animal and say:
"Look, it was the eye of the [name of the animal]!"

However, after turning the last page, say:
"Oh, look, that's not an eye! That's a dot! It's a dot on the snake/butterfly!"

Turn back to the penultimate page, point at the eye, and ask:
Test question 1: "So, what did you think this was before we turned the page? An eye or a dot?"

Then, do **not** turn the page and ask:
Control question 1: "And what is this really? An eye or a dot?"

Now, close the book and fetch the figurine Jonas. Place Jonas in front of the book and ask:
"Look, this is Jonas. Jonas has never ever looked inside this book!"

Open the book at the penultimate page (i.e., at the picture of the vulture/hippopotamus) and ask:
Test question 2: "If we show Jonas this picture, what will Jonas think this is? An eye or a dot?"

Then, do **not** turn the page and ask:
Control question 2: "And what is this really? An eye or a dot?"

Procedure of the Complement Syntax Task (Articles 2 and 3)

Material:

- A puppet theatre, laptop, loud speakers
- 3 hand puppets:
  Duck Emma and frog Ole (operated by **E2** hiding behind the puppet theatre)
  Mouse Ida (operated by **E1** sitting next to the puppet theatre)

Procedure:

Familiarization:
E1 introduces the mouse Ida and introduces the game:
E1: "Look, this is the mouse Ida." As mouse Ida: "Hello!" "Mouse Ida has two friends, frog Ole and duck Emma. And even though mouse Ida has biiig ears, she doesn't hear so well. Thus, she never hears what her friends tell her ..."
E1: "Shall the mouse call her friend frog Ole?" As mouse Ida: "Hello Ole! Where are you? Come here!"

Training phase:
The frog (played by E2) shows up: "Hello!"
Mouse Ida: "Huhh?! What did he say?"
Wait for child to reply with "Hello".

If the child replies with "Hello", immediately switch to the two practice stories.

If the child does not respond with "Hello" or refuses to do so, use the following prompts:

- After 5 sec.: Mouse: "What did he say? I did not understand what Ole said."
- After another 5 sec.: Mouse: "Ole, could you repeat that again, because I didn't understand you!"
  Frog: "Hello!"
  Mouse: "Huuhh, what did he say?"
- After another 5 sec.: E1: "Come on, help the mouse and tell her what the frog said! Otherwise, the mouse will be really sad ..."
  Mouse: "Pleeeaase tell me what the frog said!"
- If the child still does not react/respond try with this:
  Frog: "Blablabla!" , Mouse: "What did he say?"
  Frog: "Quak, quak!", Mouse: "What did he say?"
- If the child finally replies, give positive feedback through the mouse: "Thank you soo much!!"
- If the child still refuses to reply, the task needs to be terminated.

As soon as the child is willing to repeat something the frog said:

Start with the practice stories:
Frog: "Peter says the frog is fat!"
Mouse: "What did Peter say?"
Wait for the reply. If the child does not reply, use one of the below prompts.

Frog: "Grandma says that the rabbit dances!"
Mouse: "What did Grandma say?"
Wait for the reply. If the child does not reply, use one of the below prompts.

If the child does not repeat the frog's stories, use the following prompts:
E1: "Look, I will show you how to do that. Now, I first help the mouse. Frog, repeat the story for me!"
Frog repeats the corresponding story.
Mouse: "What did Peter/Grandma say?"
E1 to mouse: "The frog is fat." or "that the rabbit dances."
E1 to child: "Now it's your turn!"

After the child has repeated the two practice stories:
E1: "Oh, I think the frog is really tired now and he goes to sleep. But maybe duck Emma will show up now!" As mouse Ida: "Bye Frog! Emma! Come here!"

Test phase:
The duck Emma shows up played by E2.
E1: "Now duck Emma will tell us some stories that she has heard. You can help the mouse again and tell her the stories. The stories are quite long and sometimes duck Emma mixes them up a little bit. But then you just tell the mouse the story correctly."

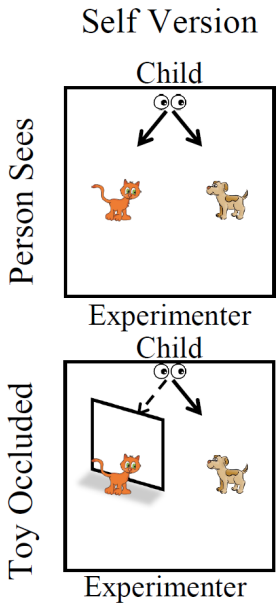
The pre-recorded sentences are played to the child via the laptop and the loud speakers. E2 moves the duck and operates the laptop. For the 33-month-olds, only 16 sentences are played, for the 36-month-olds, all 24 sentences are played.
After each sentence, the mouse turns to the child and says: "What did Anna say?"

Information:
The first two sentences are learning exercises, but they are already counted for final performance. These two sentences can be re-played up to twice if the child refuses to reply.
For all other sentences: **at maximum one repetition**, if the child does not reply. Unless the child has obviously not understood the sentence acoustically (because the child was talking or a sibling/the parent made a noise), then another repetition is fine.

Control questions:
They are asked after the child has repeated the last sentence or earlier if the task needed to be stopped at an earlier point.

E1: "Does the duck always say everything correctly?"
E1: "And you, did you repeat the stories correctly then?"

Procedure of the Visual and Epistemic Perspective-Taking Tasks (Articles 2 and 3)



Visual Perspective-Taking Self:

Material:

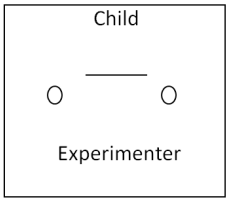
- Two stuffed animals (a dog and a cat)
- A thin cardboard partition wall

Procedure:
While asking the below control questions, point at each animal.
If children do not know what the animal is called, tell them and ask again.
If children use their own word for the animal, use that word in the later test questions.

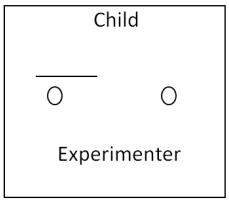
Control question 1: "What's that?" [point at dog]

Control question 2: "What's that?" [point at cat]

Arrange the animals on the table (see picture).



E: "Now, I move this one here."
Move the partition wall in front of the cat such that
the child cannot see the cat anymore but can still see the dog
(see picture)-



E: "Now I will ask you what you can see.
Say 'yes' if you can see it and say 'no' if you cannot see it."

Prompt: If the child later does not say 'yes' or 'no' to Test question 1 or Test question 2, repeat the above sentence once, and ask the test question again.

Test question 1: "So, can you see the cat?"

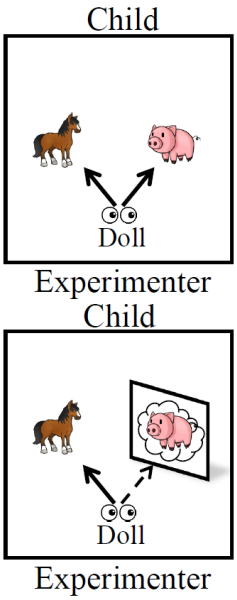
Test question 2: "So, can you see the dog?"

Visual Perspective-Taking Other:

Material:

- Two stuffed animals (a pig and a horse)
- One hand puppet (Martin)
- A thin cardboard partition wall

Other Version

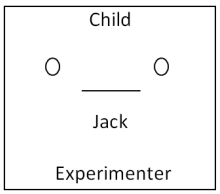


Procedure:
While asking the below control questions, point at each animal.
If children do not know what the animal is called, tell them and ask again.
If children use their own word for the animal, use that word in the later test questions.

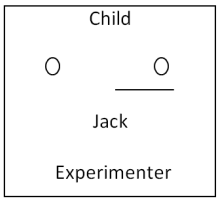
Control question 1: "What's that?" [point at pig]

Control question 2: "What's that?" [point at horse]

Arrange the animals on the table (see picture).



E: "This is Martin."
Take out puppet Martin and place him as in the picture:



E: "Now, I move this one here."
Move the partition wall in front of the pig such that Martin cannot see the pig anymore but can still see the horse (see picture).

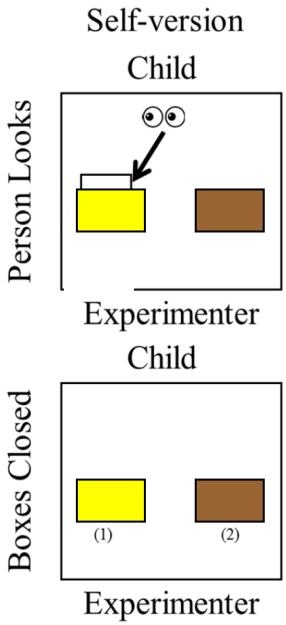
E: "Now I will ask you what Martin can see.
Say 'yes' if he can see it and say 'no' if he cannot see it."

Prompt: If the child later does not say 'yes' or 'no' to Test question 1 or Test question 2, repeat the above sentence once, and ask the test question again.

Test question 1: "So, can Martin see the pig?"

Test question 2: "So, can Martin see the horse?"

Epistemic Perspective-Taking Self:



Material:

- One yellow box, one brown box (each with lid)
- A one Euro coin in the yellow box

Procedure:
E: "I will show you what's in the yellow box, but not what's in the brown box."
Open the yellow box and show the child what's inside. Let the child take out the coin, hold it, and put it back.

Leave the box open.

Control question: "What's in the box?"

If the child does not know what it is, tell the child and ask the question again.

Close the box.

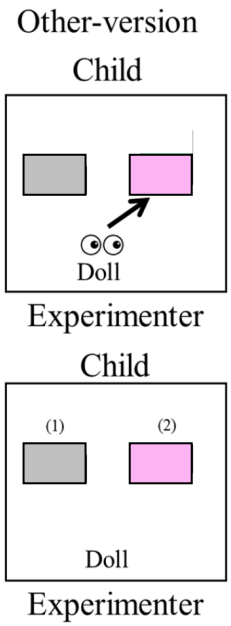
E: "I will now ask you whether you know what's in each of these boxes.
Say 'yes' if you know it and 'no' if you do not know it."

Prompt: If the child later does not say 'yes' or 'no' to Test question 1 or Test question 2, repeat the above sentence once, and ask the test question again.

Test question 1: "Do you know what's in the yellow box?"

Test question 2: "Do you know what's in the brown box?"

Epistemic Perspective-Taking Other:



Material:
- One grey box, one pink box (each with lid)
- One hand puppet (Kathi)

Procedure:
E: "That's Kathi!"
Show the puppet to the child.

E: "I will show Kathi what's in the pink box, but not what's in the grey box."
Open the pink box and let Kathi peek inside, do NOT let the child peek inside!

Kathi: "Ok, I see what's in the box!"

Close the box again.

E: "I will now ask you whether Kathi knows, what's in each of the boxes.
Say 'yes' if she knows it and 'no', if she does not know it."

Prompt: If the child later does not say 'yes' or 'no' to Test question 1 or Test question 2, repeat the above sentence once, and ask the test question again.

Test question 1: "Kathi looked inside the pink box. Does Kathi know what's in the pink box?"

Test question 2: "Kathi did not look inside the grey box. Does Kathi know what's in the grey box?"

Procedure of the Metacognition of Own Ignorance Task (Articles 2 and 3)

Material:

- Black-and-white pictures of familiar and unfamiliar objects bound in a book, one object per page

Procedure:

Take out the book and open it.
E: "We will take a look at some pictures now."

2 control trials: (book and bird)

Turn to the first page with the first control item.
E: "What's that?" Wait for the child's reply or reaction.

Turn to the next page with the second control item.
E: "What's that?" Wait for the child's reply or reaction.

12 test trials: 6 familiar and 6 unfamiliar objects

Always turn to the next page with the next picture and ask for the child's reaction or reply.
E: "What's that?"
If the child fails to reply, the question can be repeated once with the same wording, otherwise, turn to the next page.

Randomizations: (xx = one of the unfamiliar objects)

trial	order A	order B
control 1	book	book
control 2	bird	bird
1	car	car
2	xx	socks
3	socks	xx
4	dog	xx
5	xx	dog
6	spoon	spoon
7	xx	chair
8	xx	xx
9	xx	xx
10	chair	shoe
11	xx	xx
12	shoe	xx

Procedure of the Pretest and Posttest (Article 3)

Material:

- For pretest and posttest each, five different containers and five different small objects that fit into the containers

Procedure:

Scenario 1: The relation between seeing and knowing:

E2 is waiting outside the room with the two puppets Kathi and Martin**.**

E1 **hides a red spoon** in the **purple paper basket** together with the child.

E1: "Look, we hide the spoon. Now we fetch Kathi and Martin."

After calling Kathi and Martin in, E2 brings the puppets Kathi and Martin into the room. E1 holds the paper basket closed.

E1: "I will now show Kathi what's in the basket, but I will not show Martin. Martin looks away."

Kathi **looks** inside the basket and Martin **does not** (Martin turns away such that the child realizes that Martin really is not looking).

Test question: E1: "Does Martin know what's in the paper basket?"

Control question: E1: "Has Martin looked inside the paper basket?"

E2 leaves the room with the two puppets.

Scenario 2: The relation between telling and knowing:

E1 **hides** a **green die in a pink bowl** together with the child.

E1: "So, now we hide the die here. And now we fetch Kathi and Martin again."

After calling Kathi and Martin inside, E2 brings the puppets Kathi and Martin into the room.

E1 covers the bowl with their hands such that nobody can look inside.

E1: "I will now tell Martin what's in the bowl, but I will not tell Kathi. Kathi looks away and she will not listen."

E1 **whispers into** Martin's ear: "There is a die inside!" (Kathi turns away such that the child realizes that Kathi really is not looking or listening).

Test question: E1: "Does Martin know what's in the bowl?"

Explanation prompt: E1: "Why?"

Control question: E1: "Have I told Martin what's inside?"

E2 leaves the room with the two puppets.

Scenario 3: Contrasts between different knowledge states:

E1: "Soon E2 will come into the room and show you something. But, I will not look and I will not listen!" E2 is called into the room by E1 and enters the room with an **orange plastic cup with blue dots and a green spinning top hidden in the cup.**

E2 shows the child the content of the cup: "Shh! There is a spinning top inside!" while E1 is turning away, having their eyes closed, and having their ears covered with their hands.

E2 leaves the room again and E1 turns back to the child.

Test question 1: E1: "Do I know what's inside?"

Test question 2: E1: "Do you know what's inside?"

Control question 1: E1: "Have I looked inside?"

Control question 2: E1: "Have you looked inside?"

Scenario 4: Challenging another person's knowledge claims:

E1 and the child hide a **pink bouncy ball with glitter** in a **red watering can with yellow dots**.

E2 is called into the room and E1 covers the opening of the watering can with their hands.

Prompt: E2: "I know what's in there!"

Test question: E2: "Do you think I know that?"

Control question: E2: "Have I looked in there?"

E2 stays in the room afterwards.

Scenario 5: Differences between previous and current states of knowledge:

E1 shows the child a **purple flower pot** but does not reveal the content (**dark purple spoon**).

E1: "What do you think is in the flower pot?" E1 waits for the child's assumption. If the child does not make one, provide a suggestion yourself. The child should confirm that they assume the same. Then, E1 shows the child the true content of the flower pot and puts it back.

Test question 1: E1: "What did you previously think was in the flower pot?"

Test question 2: E1: "What do you now think is in the flower pot?"

Control question: E1: "Have you looked inside the flower pot before?"

Containers and objects of the posttest:

Scenario 1: a green match box car in a yellow bowl

Scenario 2: a pink spoon in a yellow paper bag

Scenario 3: a memory card with a rocket in a small cardboard house with orange dots

Scenario 4: a yellow plastic egg in a blue cloth bag

Scenario 5: a light blue matchbox car in a green watering can

Appendix C: Coding Schemes

Coding Scheme of the Low-Demands False Belief Task (Articles 1 and 2)

Participant
Code of the Participant

Included_Task	
value	description
0	Not included in sample if 1. Child failed to answer practice questions. 2. Child pointed incorrectly in one or both of the practice questions. 3. Experimenter asked “in the bucket or in the basket” or asked “here or here” or anything else pointing out the two options in the test question. 4. Child replied with “at Peter”, “it is gone” or similar. 5. Parents or siblings interfered. 6. Ambiguous reply (e.g., pointing at one thing but saying the other thing). 7. Experimenter made a different mistake. 8. Child did not reply or said “I don’t know”.
1	Included in sample.

FirstPracticeQuestion	
value	description
No value	Child failed to point at either of the pictures.
0	Child pointed at the picture of the banana.
1	Child pointed at the picture of the apple.

SecondPracticeQuestion	
value	description
No value	Child failed to point at either of the pictures.
0	Child pointed at the picture of the rattle.
1	Child pointed at the picture of the ball.

TestQuestion	
value	description
No value	Child failed to point at either of the pictures.
0	Child pointed at the picture of the bucket.
1	Child pointed at the picture of the basket.

Coding Scheme of the Explicit False Belief Tasks (Article 1)

Participant
Code of the participant

Included_Task_ContentsFB	
value	description
0	Not included in sample if 1. Child did not reply with 'Smarties' in Content_Belief. 2. Child did not reply with 'Pig' in Memory_Control. 3. Child did not reply to one or both of the test questions. 4. Child replied ambiguously. 5. Parents or siblings interfered. 6. Experimenter made a mistake. 7. Child did not know what 'Smarties' are.
1	Included in sample.

FB_Smarties_Content_Belief
Verbal utterance of the child. Should be 'Smarties' at some point! Please note if child had no content_belief.

FB_Smarties_Memory_Control
Verbal utterance of the child. Should be 'Pig(let)' at some point! Please note if child repeatedly did not pass memory_control.

FB_Smarties_Other_Belief	
value	description
No value	Child failed to reply to the question or said something else.
0	Pig(let)
1	Smarties

FB_Smarties_Other_Control	
value	description
No value	Child failed to reply to the question or said something else.
0	Yes
1	No

FB_Smarties_Successs	
value	description
No value	Child failed to reply to one or both of the previous questions.
0	0 in one or both of the previous questions.
1	1 in both previous questions.

Included_Task_ExplicitFB	
value	description
0	Not Included in sample if 1. Child did not reply to one or both of the test questions. 2. Child replied ambiguously. 3. Parents or siblings interfered. 4. Experimenter made a mistake.
1	Included in sample.

FB_Paul_Search	
value	description
No value	Child failed to reply to the question or said something else.
0	Backpack
1	Closet

FB_Paul_Reality	
value	description
No value	Child failed to reply to the question or said something else.
0	Closet
1	Backpack

FB_Paul_Success	
value	description
No value	Child failed to reply to one or both of the previous questions.
0	0 in one or both of the previous questions.
1	1 in both of the previous questions.

FB_SumScore	
value	description
No value	No value in FB_Smarties_Success and/or FB_Paul_Success.
X	Sum of FB_Smarties_Success and FB_Paul_Success if values in both variables.

Coding Scheme of the False Belief Picture Book Task (Article 2)

Participant
Code of the participant

Included_Task_SelfPart	
value	description
0	Not included in sample if 1. Child failed to reply to questions. 2. Child replied ambiguously in test/control questions (did not clearly decide for "eye" or "dot"). 3. Parents or siblings interfered. 4. Child replied with "butterfly" or "snake" in the test question of the self part. 5. Experimenter made a mistake (e.g., showed the real picture while asking the control question).
1	Included in sample.

Included_Task_OtherPart	
value	description
0	Not included in sample if 1. Child failed to reply to questions. 2. Child replied ambiguously in test/control questions (did not clearly decide for "eye" or "dot"). 3. Parents or siblings interfered. 4. Child replied with "butterfly" or "snake" in the test question of the other part. 5. Experimenter made a mistake (e.g., showed the real picture while asking the control question).
1	Included in sample.

PB_Self_TestQuestion	
value	description
No value	Child failed to reply to the question or said something else.
0	Dot
1	Eye

PB_Self_ControlQuestion	
value	description
No value	Child failed to reply to the question or said something else.
0	Eye
1	Dot (or: butterfly/snake)

PB_Self_Passers	
value	description
No value	Child failed to reply to test and/or control question of the self part.
0	0 in PB_Self_TestQuestion and/or PB_Self_ControlQuestion
1	1 in PB_Self_TestQuestion and PB_Self_ControlQuestion

PB_Other_TestQuestion	
value	description
No value	Child failed to reply to the question or said something else.
0	Dot
1	Eye

PB_Other_ControlQuestion	
value	description
No value	Child failed to reply to the question or said something else.
0	Eye
1	Dot (or: butterfly/snake)

PB_Other_Passers	
value	description
No value	Child failed to reply to test and/or control question of the other part.
0	0 in PB_Other_TestQuestion and/or PB_Other_ControlQuestion
1	1 in PB_Other_TestQuestion and PB_Other_ControlQuestion

Coding Scheme of the Complement Syntax Task (Articles 2 and 3)

Step 1:
After children's verbal productions from the task have been transcribed, the transcripts need to be coded on a sentence-by-sentence basis.

Complete	
value	description
-1	One-word reply OR Complement sentence did not contain the verb and/or the object OR Child repeated one of the previous sentences (i.e., one with a different content) OR Child said something completely unrelated to the sentence/task.
0	Complement sentence contained both the verb and the object, BUT the subject of the complement clause was missing and/or the complementizer "dass" if the original sentence had contained a "dass".
1	Complement sentence was complete, this means, it contained the subject, verb, and object of the original clause and for a sentence with a complementizer also the complementizer "dass". Note: It is sufficient if the child used only the article instead of the full subject (e.g., "der Clown fängt Kugeln" -> "der fängt Kugeln").

Literal	
Only code for repetitions with **Complete = 0** or **Complete = 1**.	
value	description
1	Complement sentence was repeated as in the audio recording, that is: - Verb and object were in the same order as in the original complement sentence (it does not matter where the subject was placed). - No "dass" was added to a complement clause which originally did not have a complementizer. - No second verb was included into the complement clause (e.g., "dass der Opa kriegt Blumen kriegt" or "dass der Opa hat Blumen kriegt"). - If a "dass" was left out from a complement clause with "dass", this also counts as literal = 1 because it is unclear whether the child deliberately left out the "dass" (e.g., "Anna sagt, dass der Opa kriegt Kekse" was repeated as "der Opa kriegt Kekse").
0	Complement sentence was modified, that is: - Order of verb and object was switched (e.g., "der Opa Kekse liebt" was modified into "der Opa liebt Kekse"). - A "dass" was added to a complement sentence without a "dass" (e.g., "der Opa Blumen kriegt" was modified into "dass der Opa Blumen kriegt"). - Both the order of verb and object were switched AND a "dass" was added to the complement sentence. - A complement sentence with two verbs was created (see example above). - If a "dass" was left out from a complement clause with "dass", this also counts as literal = 1 because it is unclear whether the child deliberately left out the "dass". EXCEPT: The child repeated matrix clause and the complement clause. Thus, it is clear that the child deliberately left out the "dass".

	Example 1: “Anna sagt dass der Arzt holt Bücher.” was modified into “Anna sagt der Arzt holt Bücher.” Example 2: “Anna sagt dass der Clown Kugeln fängt.” was modified into “Anna sagt der Clown Kugeln fängt.”

Corrected	
For sentences which originally had an incorrect word order in the complement clause. Only code for sentences with **Complete = 0** or **Complete = 1** and with **Literal = 0.**	
value	description
1	The modification of the complement sentence (see literal = 1) resulted in an accurate verb-object order in the complement clause, that is: -> either: “**dass” + (subject) + object + verb** -> or**: (subject) + verb + object** by making one of the following modifications: - Order of verb and object was switched. - A “dass” was added to a complement clause which originally did not contain a “dass” without additionally changing the order of verb and object. - If both matrix clause and complement clause were repated: The “dass” was left out from a repetition without additionally changing the order of verb and object.
0	The complement sentence was modified but the modification did not result in a correction of the sentence because: - Order of verb and object were switched, but additionally the “dass” was added or removed. - Modification resulted in having a complement sentence with two verbs. - The resulting sentence did not have the same structure as the original clause anymore (e.g., “Clown soll Bälle fangen.” or “Kugeln zu fangen der Clown.”).

Corrected	
For sentences which originally had a correct word order in the complement clause. Only code for sentences with **Complete = 0** or **Complete = 1** and with **Literal = 0.**	
value	description
-1	Complement sentence was worsened/turned into an incorrect sentence due to making one of the following modifications: - Order of verb and object were switched, resulting in an incorrect order. - A “dass” was added to a complement clause which originally did not include a complementizer WITHOUT additionally switching the order of verb and object. - The “dass” was left out from a complement sentence which originally included a complementizer AND the matrix sentence was also repeated but the order of object and verb was not switched. - Modification resulted in having an incorrect sentence with two verbs. - The resulting sentence was for some other reason not a grammatically correct sentence anymore.
No value	If literal = 0, but none of the above cases happened, leave the variable empty for correct clauses.

Step 2:
Aggregate raw data across all valid sentences of a participant.

Participant
Code of the participant

CS_33/36_NonLiteralRepetitionsTotal
Count how many sentences there are with complete = 0/1 **and** literal = 0.

CS_33/36_NonLiteralRepetitionsCorrectSentences
Count how many green sentences there are with complete = 0/1 **and** literal = 0.

CS_33/36_NonLiteralRepetitionsIncorrectSentences
Count how many red sentences there are with complete = 0/1 **and** literal = 0.

CS_33/36_CorrectedSentences
Count how many red sentences there are with complete = 0/1 **and** corrected = 1 (literal **has to be 0).**

CS_33/36_FalsifiedSentences
Count how many green sentences there are with complete = 0/1 **and** corrected = -1 (literal **has to be 0).**

CS_33/36_RepeatedSentences_CorrectSentences
Count how many green sentences there are with complete = 0/1.

CS_33/36_RepeatedSentences_IncorrectSentences
Count how many red sentences there are with complete = 0/1.

Step 3:
Calculate the final scores.

CS_33/36_Quotient1 (original literality score by Grosse Wiesmann et al., 2017)
= CS_33/36_NonLiteralRepetitionsIncorrectSentences / CS_33/36_NonLiteralRepetitionsTotal

CS_33/36_Quotient2 (original correction score by Grosse Wiesmann et al., 2017)
= CS_33/36_CorrectedSentences / CS_33/36_NonLiteralRepetitionsTotal

CS_33/36_Score1 (new literality score)
= (CS_33/36_NonLiteralRepetitionsIncorrectSentences / CS_33/36_RepeatedSentences_Incor-rectSentences) – (CS_33/36_NonLiteralRepetitionsCorrectSentences / CS_33/36_Repeated-Sentences_CorrectSentences)

CS_33/36_Score2 (new correction score)
= (CS_33/36_CorrectedSentences / CS_33/36_RepeatedSentences_IncorrectSentences) – (CS_33/36_FalsifiedSentences / CS_33/36_RepeatedSentences_CorrectSentences)

Coding Scheme of the Visual and Epistemic Perspective-Taking Tasks (Articles 2 and 3)

Participant
Code of the participant

Included_VisualPT_Self	
value	description
0	Not included in sample if 1) Child looked behind the wall, but **E1 did not ask again.** 2) Child failed to reply to both control questions (i.e., no answer). 3) Child did not reply to test questions. 4) Parents or siblings interfered. 5) Experimenter made a mistake.
1	Included in sample.
Included_VisualPT_Other	
value	description
0	Not Included in sample if **1)** Child moved the wall (such that Martin can see both animals), and **E1 did not ask again.** 2) Child failed to reply to both control questions (i.e., no answer). 3) Child did not reply to test questions. 4) Parents or siblings interfered. 5) Experimenter made a mistake.
1	Included in sample.
Included_EpistemicPT_Self	
value	description
0	Not Included in sample if 1) Child failed to reply to the control question (i.e., no answer, does not know what money is). 2) Child did not reply to test questions. 3) Parents or siblings interfered. 4) Experimenter made a mistake.
1	Included in sample.
Included_EpistemicPT_Other	
value	description
0	Not included in sample if 1) Child did not reply to test questions. 2) Parents or siblings interfered. 3) Experimenter made a mistake (e.g., question was formulated incorrectly, e.g. "Kathi knows what's in the box" instead of "Kathi looked into the box").
1	Included in sample.

VisualPT_Self_TQ1:	
value	description
No value	No reply.
0	Incorrect reply, i.e., saying “yes“ or nodding.
1	Correct reply, i.e., saying “no“ or shaking head.
VisualPT_Self_TQ2:	
value	description
No value	No reply.
0	Incorrect reply, i.e., saying “no“ or shaking head.
1	Correct reply, i.e., saying “yes“ or nodding.

VisualPT_Other_TQ1:	
value	description
No value	No reply.
0	Incorrect reply, i.e., saying “yes” or nodding.
1	Correct reply, i.e., saying “no” or shaking head.
VisualPT_Other_TQ2:	
value	description
No value	No reply.
0	Incorrect rely, i.e., saying “no” or shaking head.
1	Correct reply, i.e., saying “yes” or nodding.

EpistemicPT_Self_TQ1:	
value	description
No value	No reply.
0	Incorrect reply, i.e., saying “no” or shaking head.
1	Correct reply, i.e., saying “yes” or nodding. Also acceptable: child says “money”.
EpistemicPT_Self_TQ2:	
value	description
No value	No reply.
0	Incorrect reply, i.e., saying “yes” or nodding. Or, child says: “money”, “nothing”, or any other object.
1	Correct reply, i.e., saying “no” or shaking head.

EpistemicPT_Other_TQ1:	
value	description
No value	No reply.
0	Incorrect reply, i.e., saying “no” or shaking head.
1	Correct reply, i.e., saying “yes” or nodding.
EpistemicPT_Other_TQ2:	
value	description
No value	No reply.
0	Incorrect reply, i.e., saying “yes” or nodding.
1	Correct reply, i.e., saying “no” or shaking head.

VisualPT_Self:	
value	description
No value	No value in VisualPT_Self_TQ1 and/or VisualPT_Self_TQ2
0	0 points in VisualPT_Self_TQ1 and/or VisualPT_Self_TQ2
1	1 point in VisualPT_Self_TQ1 and VisualPT_Self_TQ2
VisualPT_Other:	
value	description
No value	No value in VisualPT_Other_TQ1 and/or VisualPT_Other_TQ2
0	0 points in VisualPT_Other_TQ1 and/or VisualPT_Other_TQ2
1	1 point in VisualPT_Other_TQ1 and VisualPT_Other_TQ2
VisualPT_Sum	
value	description
No value	No value in VisualPT_Self and/or VisualPT_Other
X	Sum of VisualPT_Self and VisualPT_Other

EpistemicPT_Self:	
value	description
No value	No value in EpistemicPT_Self_TQ1 and/or EpistemicPT_Self_TQ2
0	0 points in EpistemicPT_Self_TQ1 and/or EpistemicPT_Self_TQ2
1	1 point in EpistemicPT_Self_TQ1 and EpistemicPT_Self_TQ2
EpistemicPT_Other:	
value	description
No value	No value in EpistemicPT_Other_TQ1 and/or EpistemicPT_Other_TQ2
0	0 points in EpistemicPT_Other_TQ1 and/or EpistemicPT_Other_TQ2
1	1 point in EpistemicPT_Other_TQ1 and EpistemicPT_Other_TQ2
EpistemicPT_Sum	
value	description
No value	No value in EpistemicPT_Self and/or EpistemicPT_Other
X	Sum of EpistemicPT_Self and EpistemicPT_Other

Coding Scheme of the Metacognition of Own Ignorance Task (Articles 2 and 3)

Participant
Code of the participant

Included_Task	
value	description
0	Not included in sample if: 1. Child did not reply correctly to at least one of the two control items. 2. Child failed to reply to the test items.
1	Included in sample.

Control_Item1	
value	description
0	Book was not correctly named.
1	Book was correctly named.
Examples for correct answers:	
Book:	Picture book, book, other type of book

Control_Item2	
value	description
0	Bird was not correctly named.
1	Bird was correctly named.
Examples for correct answers:	
Bird:	Bird, specific types of birds (duck, swan, sparrow, eagle, ...)

Word_Misapp	
value	description
0	None of the real items were falsely named by using an incorrect word.
X	Number of real items for which an incorrect word was used / misapplied (e.g., “socks“ used for picture of “shoes”).

Word_Inven	
value	value
0	None of the real items were named with an invented word.
X	Number of real items for which an invented word was used (e.g., “Babi“ used for picture of “shoes“).

Word_Hmm	
value	value
0	Child did not say “Hmm“ confronted with any of the real items.
X	Number of real items for which the child said “Hmm”.

Word_ShakeHead	
value	value
0	Child did not shake head confronted with any of the real items.
X	Number of real items for which the child shook head.

Word_No	
value	value
0	Child did not say “no” confronted with any of the real items.
X	Number of real items for which the child said “no”.

Word_IDK	
value	value
0	Child did not say “I don’t know” or “don’t know” confronted with any of the real items.
X	Number of real items for which the child said “I don’t know” or “don’t know”.

Word_Mom	
value	value
0	Child did not look at parent confronted with any of the real items.
X	Number of real items for which the child looked at parent.

Word_Exp1	
value	value
0	Child did not look at experimenter 1 confronted with any of the real items.
X	Number of real items for which the child looked at experimenter 1.

Word_NoAnswer	
value	value
0	Child never provided no answer confronted with any of the real items.
X	Number of real items for which the child did not provide any answer.

Fake_Misapp	
value	value
0	None of the fictitious items were falsely named by using an incorrect word.
X	Number of fictitious items for which an existing word was misapplied (e.g., “socks” used for fictitious item).

Fake_Inven	
value	value
0	None of the fictitious items were named with an invented word.
X	Number of fictitious items for which an invented word was used (e.g., “Babi” used for fictitious item).

Fake_Hmm	
value	value
0	Child did not say “Hmm” confronted with any of the fictitious items.
X	Number of fictitious items for which the child said “Hmm”.

Fake_ShakeHead	
value	value
0	Child did not shake head confronted with any of the fictitious items.
X	Number of fictitious items for which the child shook head.

Fake_No	
value	value
0	Child did not say “no” confronted with any of the fictitious items.
X	Number of fictitious items for which the child said “no”.

Fake_IDK (= Metacognition_ExplicitScore)	
value	value
0	Child did not say “I don’t know” or “don’t know” confronted with any of the fictitious items.
X	Number of fictitious items for which the child said “I don’t know” or “don’t know”.

Fake_Mom	
value	value
0	Child did not look at parent confronted with any of the fictitious items.
X	Number of fictitious items for which the child looked at parent.

Fake_Exp1	
value	value
0	Child did not look at experimenter 1 confronted with any of the fictitious items.
X	Number of fictitious items for which the child looked at experimenter 1.

Fake_NoAnswer	
value	value
0	Child never provided no answer confronted with any of the fictitious items.
X	Number of fictitious items for which the child did not provide any answer.

FakeSign **(Signs of Uncertainty:** “Hmm”, shake head, “no”, “I don’t know”, look at mother/experimenter)	
value	description
0	For none of the fictitious items, the child used one of the signs of uncertainty.
X	Number of fictitious items for which the child used at least one of the signs of uncertainty (max. = 6).
RealSign	
value	description
0	For none of the real items, the child used one of the signs of uncertainty.
X	Number of real items for which the child used at least one of the signs of uncertainty (max. = 6).
FakePerfQuality (= Metacognition_QualitativeScore) **Qualitative score where each sign of uncertainty is rated with a different amount of points for the fictitious items.** saying “Hmm” = 1 point; shake head or saying “no” = 2 points; saying “I don’t know” = 3 points all other signs = 0 points If the child used several signs of uncertainty, always use the most valuable one per item.	
value	description
X	Sum score across all six fictitious items based on the above rules (max. = 18).

Coding Scheme of the Pretest and Posttest (Article 3)

Participant
Code of the participant

Subtask1_TQ	
value	description
No value	No reply or naming the object or saying ‘I don’t know’.
0	Saying ‘yes’ or nodding.
1	Saying ‘no’ or shaking head.

Subtask1_CQ	
value	description
No value	No reply or naming the object or saying ‘I don’t know’.
0	Saying ‘yes’ or nodding.
1	Saying ‘no’ or shaking head.

Subtask2_TQ	
value	description
No value	No reply or naming the object or saying ‘I don’t know’.
0	Saying ‘no’ or shaking head.
1	Saying ‘yes’ or nodding.

Subtask2_Explanation
Reaction / Answer of the child.

Subtask2_CQ	
value	description
No value	No reply or naming the object or saying ‘I don’t know’.
0	Saying ‘no’ or shaking head.
1	Saying ‘yes’ or nodding.

Subtask3_TQ1	
value	description
No value	No reply or naming the object or saying ‘I don’t know’.
0	Saying ‘yes’ or nodding.
1	Saying ‘no’ or shaking head.

Subtask3_TQ2	
value	description
No value	No reply or naming the object or saying ‘I don’t know’.
0	Saying ‘no’ or shaking head.
1	Saying ‘yes’ or nodding.

Subtask3_CQ1	
value	description
No value	No reply or naming the object or saying ‘I don’t know’.
0	Saying ‘yes’ or nodding.
1	Saying ‘no’ or shaking head.

Subtask3_CQ2	
value	description
No value	No reply or naming the object or saying ‘I don’t know’.
0	Saying ‘no’ or shaking head.
1	Saying ‘yes’ or nodding.

Subtask4_Prompt
Reaction / Answer of the child.

Subtask4_TQ	
value	description
No value	No reply or naming the object or saying ‘I don’t know’.
0	Saying ‘yes’ or nodding.
1	Saying ‘no’ or shaking head.

Subtask4_CQ	
value	description
No value	No reply or naming the object or saying ‘I don’t know’.
0	Saying ‘yes’ or nodding.
1	Saying ‘no’ or shaking head.

Subtask5_TQ1	
value	description
No value	No reply or saying an entirely different object (neither suggestion nor actual content) or saying ‘I don’t know’.
0	Object that is actually in the container.
1	Object that was suggested prior to looking into the container.

Subtask5_TQ2	
value	description
No value	No reply or saying ‘I don’t know’.
0	Object that was suggested prior to looking into the container or any other object that is not in the container.
1	Object that is actually in the container.

Subtask5_CQ	
value	description
No value	No reply or naming an object or saying ‘I don’t know’.
0	Saying ‘yes’ or nodding.
1	Saying ‘no’ or shaking head.

Pretest/Posttest_Score1	
value	description
No value	No value in Subtask1_TQ and/or Subtask1_CQ
0	0 in Subtask1_TQ and/or Subtask1_CQ
1	1 in Subtask1_TQ and Subtask1_CQ

Pretest/Posttest_Score2	
value	description
No value	No value in Subtask2_TQ and/or Subtask2_CQ
0	0 in Subtask2_TQ and/or Subtask2_CQ
1	1 in Subtask2_TQ and Subtask2_CQ

Pretest/Posttest_Score3_Other	
value	description
No value	No value in Subtask3_TQ1 and/or Subtask3_CQ1
0	0 in Subtask3_TQ1 and/or Subtask3_CQ1
1	1 in Subtask3_TQ1 and Subtask3_CQ1

Pretest/Posttest_Score3_Self	
value	description
No value	No value in Subtask3_TQ2 and/or Subtask3_CQ2
0	0 in Subtask3_TQ2 and/or Subtask3_CQ2
1	1 in Subtask3_TQ2 and Subtask3_CQ2

Pretest/Posttest_Score4	
value	description
No value	No value in Subtask4_TQ and/or Subtask4_CQ
0	0 in Subtask4_TQ and/or Subtask4_CQ
1	1 in Subtask4_TQ and Subtask4_CQ

Pretest/Posttest_Score5	
value	description
No value	No value in Subtask5_TQ1 and/or Subtask5_TQ2
0	0 in Subtask5_TQ1 and/or Subtask5_TQ2
1	1 in Subtask5_TQ1 and Subtask5_TQ2

Pretest/Posttest_totalscore	
value	description
No value	If values in less than 3 out of the 5 subtasks.
Mean(Pretest/Posttest_Score1, Pretest/Posttest_Score2, Pre-test/Posttest_Score3_Other, Pre-test/Posttest_Score3_Self, Pre-test/Posttest_Score4, Pre-test/Posttest_Score5)	Mean of the six scores (only calculated if scores in at least 3 out of the 5 subtasks, ok if only other or self part of subtask 3 has valid data).