

Children's Subjective Understanding of Conative Mental States

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades

"Doctor rerum naturalium"

der Georg-August-Universität Göttingen

im Promotionsstudiengang Behavior and Cognition (BeCog)

der Georg-August University School of Science (GAUSS)

vorgelegt von

Britta Dorothea Schünemann

aus Göttingen

Göttingen, 2021

Table of Contents

Betreuungsausschuss

Prof. Dr. Hannes Rakoczy, Abteilung Kognitive Entwicklungspsychologie, Georg-Elias-Müller- Institut für Psychologie, Universität Göttingen

Prof. Dr. Michael Waldmann, Abteilung Kognitionswissenschaft und Entscheidungspsychologie, Georg-Elias-Müller- Institut für Psychologie, Universität Göttingen

Dr. Annie Wertz, Abteilung Naturalistische soziale Kognition, Max-Planck-Institut für Bildungsforschung Berlin

Mitglieder der Prüfungskommission

Referent: Prof. Dr. Hannes Rakoczy, Abteilung Kognitive Entwicklungspsychologie, Georg-Elias-Müller- Institut für Psychologie, Universität Göttingen

Korreferentin: Dr. Annie Wertz, Abteilung Naturalistische soziale Kognition, Max-Planck-Institut für Bildungsforschung Berlin

Weitere Mitglieder der Prüfungskommission:

Dr. Claudia Fichtel, Verhaltensökologie und Soziobiologie, Deutsches Primatenzentrum, , Universität Göttingen

Prof. Dr. Julia Fischer, Abteilung Kognitive Ethologie, Deutsches Primatenzentrum, Universität Göttingen

Prof. Dr. Annekathrin Schacht, Abteilung Affektive Neurowissenschaft und Psychophysiologie, Georg-Elias-Müller- Institut für Psychologie, Universität Göttingen

Tag der mündlichen Prüfung: 07.06.2021

Table of Contents

Preliminary Note

The present thesis is a publication-based (cumulative) dissertation. It is based on two original articles that have been submitted for publication and one original research article which is currently in preparation.

Schünemann, B.*, Schidelko, L.*, Proft, M., & Rakoczy, H. (2021). *Children understand subjective (undesirable) desires before they understand subjective (false) beliefs* [Manuscript submitted for publication]. Department of Developmental Psychology. University of Göttingen. *equal contribution as first authors

This article has been submitted for publication in Journal of Experimental Child Psychology, published by Elsevier.

Schünemann, B., Proft, M., & Rakoczy, H. (2021). *Children's developing understanding of the subjectivity of intentions – a case of “advanced Theory of Mind”* [Manuscript submitted for publication]. Department of Developmental Psychology. University of Göttingen.

This article has been submitted for publication in Journal of Cognition and Development, published by Taylor & Francis.

Schünemann, B., Bleijlevens, N., Proft, M., & Rakoczy, H. (2021). *Children's meta-representational notion of intentions – Understanding the subjectivity of intentions* [Unpublished manuscript]. Department of Developmental Psychology. University of Göttingen.

This information is available in the printed version of the manuscript.

Table of Contents

Thank you

To Hannes Rakoczy for great supervision, cooperation, guidance and discussion of research ideas. To Marina Proft for getting me enthusiastic about this field and supporting me since my bachelor thesis.

To Annie Wertz and Michael Waldmann for valuable feedback and input in the thesis committee meetings and for support and encouragement throughout the whole time.

To Claudia Fichtel, Julia Fischer, and Annekathrin Schacht for agreeing to be members of the examination board.

To the whole Kindsköpfe lab. Especially, Regina Zörner, Marlen Kaufmann, Konstanze Schirmer and all student assistants and interns for making these projects possible.

To all the students who have worked with me on my projects, were a huge support and made working on these projects so much fun: Clara Hörschler, Fiona Berger, Florian Schönig, Hanna Fehlhaber, Hannah Klonk, Imke Koch, Jana Rechenburg, Johannes Ruß, Lena Linhart, Leslie-Ann Eickhoff, Lisa Ropeter, Lynn Weiss, Maike Borrmann, Maike Holland-Letz, Nora Borse, and Sophia Braunstein.

To Lisa Wenzel, Lydia Schidelko and Feride Nur Haskaraca Kizilay for making my PhD such an amazing time and for all the sweets you always brought (and in the case of Lisa almost never ate yourself).

To Natalie Bleijlevens, Isa Garbisch, Judith Keller, and Lydia Schidelko for the great projects your theses and internships became and for a really good time.

To Juliane Bräuer, Micheal Huemer, and Josef Perner for great collaborations.

To Anja, Feride, Florian, Isa, Julie, Kristin, Lydia, Marina, Mark, Michael, Natalie, and Vincent for incredible moral support and proofreading earlier drafts of this thesis. Your comments and suggestions helped a lot to improve this manuscript. And to Florian, Luna and my parents for the extended edition of moral support and enduring all my moods.

And, last but not least, to all the children who have participated in my studies.

Table of Contents

0	Summaries	1
0.1	English Summary	1
0.2	Deutsche Zusammenfassung	3
1	General Introduction	5
2	Mental States	7
2.1	Theory of Mind	7
2.2	Properties of Mental States	7
2.2.1	Intentionality (Directedness)	8
2.2.2	Propositional Attitudes	8
2.2.3	Direction of Fit and Conditions of Satisfaction	9
2.2.4	Truth Value of Propositional Attitudes	11
2.2.5	Aspectuality and Intensionality-with-an-s	12
2.2.6	Implications of the Properties of Mental States	13
3	Developmental Perspective	14
3.1	Cognitive States	14
3.1.1	The Standard Test for Subjective Understanding of Cognitive States	14
3.1.2	Ontogenetic Development of Understanding of Cognitive States	16
3.1.2.1	4-year Revolution.	16
3.1.2.2	Earlier Forms.	18
3.1.3	Theoretical Work on Theory of Mind Development	21
3.1.4	Two-System Accounts	23
3.1.4.1	Teleological Accounts.	23
3.1.4.2	The Two-System Account by Apperly and Butterfill.	26
3.2	Conative Mental States	28
3.2.1	Ontogenetic Development of Understanding of Goals	30
3.2.2	Ontogenetic Development of Understanding of Desires	31
3.2.2.1	Asymmetry View.	31
3.2.2.2	Symmetry View.	32

Table of Contents

3.2.2.3	Testing Asymmetry and Symmetry View Against Each Other	33
3.2.2.4	Empirical Evidence	34
3.2.3	Ontogenetic Development of Understanding of Intentions	37
3.2.3.1	The Complexity of the Mental State Intentions	37
3.2.3.2	Ontogenetic Development.	40
3.2.3.3	Testing for a Subjective Conception of Intentions.	42
4	Conclusion and Open Questions	44
5	Aim of Dissertation	48
6	Summary of Empirical Findings	49
6.1	Project 1: Schünemann, Schidelko, Proft, & Rakoczy (2021)	49
6.2	Project 2a: Schünemann, Proft, & Rakoczy (2021)	51
6.3	Project 2b: Schünemann, Bleijlevens, Proft, & Rakoczy (2021)	54
7	General Discussion	56
7.1	Subjective Conception of Desires	56
7.1.1	Summary of Project 1	56
7.1.2	Implications	57
7.1.3	Limitations	58
7.1.4	Conclusion	59
7.2	Subjective Conception of Intentions	59
7.2.1	Summary of Project 2	59
7.2.2	Implications	60
7.2.3	Conclusion	61
7.3	Developmental Trajectory	61
7.4	Subjectivity as a Signature Characteristic for Full-Blown Concepts	63
7.4.1	Is Subjectivity a Signature for Full-Blown Concepts?	63
7.4.2	Approaches to Test This	65
7.5	Summary of Outlook	67
7.6	Conclusion	69

Table of Contents

8	References	70
	Appendix A: Schünemann, Schidelko, Proft, & Rakoczy (2021)	i
	Appendix B: Schünemann, Proft, & Rakoczy (2021)	xxxii
	Appendix C: Schünemann, Bleijlevens, Proft, & Rakoczy (2021)	lxxx
	Curriculum Vitae	xcvi

0 Summaries

0.1 English Summary

In our everyday lives, one substantial factor is to make sense of other agents' actions. We describe and explain these actions by referring to the agents' cognitive states, such as knowledge and beliefs, and their conative states, such as their desires and intentions (Davidson, 1963). The ability to ascribe these mental states to other agents is referred to as theory of mind (Premack & Woodruff, 1978). Mental states represent reality from the agent's perspective. Accordingly, one fundamental property of cognitive and conative states is that they are subjective. To correctly ascribe these states to an agent, the ascriber has to relativize to the agent's subjective standpoint. From an ontogenetic perspective, a central question is when and how children develop a subjective conception of other agents' mental states. For cognitive states, there is extensive empirical work which has provided the basis for differentiated and detailed theoretical work (for an overview, see Rakoczy, 2017a). In contrast, much less work exists on the development of children's subjective understanding of conative states. One reason for this might be that, compared to cognitive states, testing for a subjective conception of conative states is less straightforward.

This dissertation presents ways to test for a subjective conception of desires and intentions. It is based on two projects in which these approaches were implemented. The first project tested for children's subjective understanding of desires. One way in which desires are subjective is that they can be incompatible with norms and values (e.g., the desire to destroy something). Such desires are strongly subjective because, objectively, their outcome is undesirable. I compared 2- to 4-year-olds' capacity to reason about such subjective wicked desires to their capacity to reason about objectively reasonable neutral desires and subjective (false) beliefs. Younger children were better in reasoning about subjective desires than about subjective beliefs. Also, they did not face more difficulties to reason about subjective desires than about neutral desires. This suggests that children develop a subjective understanding of desires before they develop a subjective understanding of beliefs.

The second project of this dissertation addressed children's subjective understanding of intentions in two studies. One way in which intentions are subjective is that they are aspectual (Searle, 1983): Actions are unintentional under descriptions or aspects the agent does not represent. In the first study, children observed an agent who falsely believed that a box contained only a ball but not a pen. Thus, when the agent took this box, her action was

intentional under the description “take the ball” but unintentional under the description “take the pen”. However, children younger than six falsely claimed that she intentionally took the pen. The second study addressed whether younger children’s difficulties to consider the aspectuality of intentions might have reflected performance limitations rather than competence limitations. Possibly, even younger children would have been able to consider the agent’s subjective perspective but simply failed to recognize that this was necessary to solve this task. For this reason, the second study transferred the task into a morally relevant context that emphasized the necessity to relativize to the agent’s standpoint. In this task version, the agent unintentionally performed actions that were harmful towards another agent. In this relevant context, children already appreciated the aspectuality of the agent’s intentions by the age of five. This suggests that a subjective understanding of intentions develops around one year later than a subjective understanding of beliefs.

The findings of this dissertation contribute to a comprehensive understanding of how theory of mind develops ontogenetically. They support a developmental trajectory in which the development of a subjective conception follows different courses for different mental states. In combination with existing evidence, this dissertation suggests that children first develop a subjective conception of desires before developing a subjective conception of beliefs. Only after these have developed, do children develop a subjective conception of intentions.

0.2 Deutsche Zusammenfassung

Die Handlungen anderer Menschen nachzuvollziehen und vorherzusehen ist von zentraler Bedeutung für unser tägliches Leben. Wir beschreiben und erklären uns die Handlungen anderer Akteur:innen, indem wir ihnen kognitive mentale Zustände, (z.B. Überzeugungen und Wissen) und konative mentale Zustände (z.B. Wünsche und Intentionen) zuschreiben (Davidson, 1963). Die Fähigkeit mentale Zustände zuzuschreiben wird als Theory of Mind bezeichnet (Premack & Woodruff, 1978). Mentale Zustände sind Repräsentationen der Realität aus der Perspektive der Akteur:innen. Um diese mentalen Zustände korrekt zuzuschreiben, muss der:die Zuschreibende folglich die subjektive Perspektive der Akteur:innen einnehmen. Aus Entwicklungsperspektive ist deshalb eine zentrale Frage, wann und wie sich ein subjektives Verständnis von kognitiven und konativen Zuständen entwickelt. Es existiert umfassende empirische Evidenz, wie sich ein subjektives Verständnis kognitiver Zustände entwickelt. Aufbauend auf dieser Evidenz existieren auch viele ausdifferenzierte theoretische Überlegungen zu diesem Entwicklungsprozess (für eine Übersicht siehe Rakoczy, 2017a). Das subjektive Verständnis konativer Zustände ist deutlich weniger genau untersucht. Dies liegt vermutlich zum Teil daran, dass es deutlich weniger offensichtlich ist, wie ein subjektives Verständnis konativer Zustände erfasst werden kann.

Diese Dissertation zeigt Möglichkeiten auf, ein solches Verständnis dennoch zu erfassen. Die Dissertation basiert auf zwei Projekten, in denen diese Ansätze umgesetzt wurden. Im ersten Projekt wurde das subjektive Verständnis von Wünschen untersucht. Wünsche sind unter anderem subjektiv, weil sie Normen und Werten widersprechen können (z.B. der Wunsch etwas zu zerstören). Objektiv ist ein solcher Wunsch nicht wünschenswert. Ein solcher Wunsch ist folglich nur aus der subjektiven Perspektive des Akteurs nachvollziehbar. In dieser Studie wurde das Verständnis zwei- bis vier-jähriger Kinder von subjektiven bösen Wünschen verglichen mit ihrem Verständnis objektiv nachvollziehbarer neutraler Wünsche und subjektiver (falscher) Überzeugungen. Jüngere Kinder zeigten ein besseres Verständnis für subjektive Wünsche als für subjektive Überzeugungen. Außerdem fiel es ihnen nicht schwerer die subjektiven Wünsche nachzuvollziehen als die neutralen Wünsche. Ein subjektives Verständnis von Wünschen scheint sich folglich schon früher zu entwickeln als ein subjektives Verständnis von Überzeugungen.

Das zweite Projekt dieser Dissertation untersuchte das subjektive Verständnis von Intentionen in zwei Studien. Dass Intentionen subjektiv sind, bildet sich unter anderem in ihrer Aspekthaftigkeit ab (Searle, 1983): Wenn Akteur:innen eine Beschreibung (Aspekt) einer

Handlung nicht repräsentieren, dann ist die Handlung unter dieser Beschreibung auch nicht absichtlich. In der ersten Studie von Projekt 2 glaubte der:die Akteur:in zum Beispiel fälschlicherweise, dass in einer Box nur ein Ball nicht aber ein Stift sei. Wenn der:die Akteur:in diese Box nahm, dann war diese Handlung absichtlich unter der Beschreibung „einen Ball nehmen“ aber nicht unter der Beschreibung „einen Stift nehmen“. Kinder berücksichtigten erst ab einem Alter von sechs Jahren, dass die Handlung nur unter der ersten Beschreibung absichtlich ist. Die zweite Studie untersuchte, inwieweit die Schwierigkeiten jüngerer Kinder in der vorherigen Studie nur eine Performanz- nicht aber eine Kompetenzlimitierung darstellen. Möglicherweise war jüngeren Kindern nicht bewusst, dass sie die Aspekthaftigkeit der Intention berücksichtigen mussten, obwohl sie dazu in der Lage gewesen wären. Aus diesem Grund wurde die Aufgabe in der zweiten Studie in einen moralisch relevanten Kontext eingebettet. Dieser stellte deutlich heraus, dass die subjektive Perspektive der Akteur:innen berücksichtigt werden musste. In dieser Adaption fügte der:die Akteur:in unabsichtlich einer anderen Person einen Schaden zu. In diesem relevanten Kontext berücksichtigten schon Fünfjährige die Aspekthaftigkeit der Intention der Akteurin/ des Akteurs. Ein subjektives Verständnis von Intentionen scheint sich also in Vergleich zu Überzeugungen erst etwa ein Jahr später zu entwickeln.

Die Ergebnisse dieser Dissertation tragen zu einem umfassenden Verständnis bei, wie sich Theory of Mind entwickelt. Sie unterstützen die Annahme unterschiedlich verlaufender Entwicklungen eines subjektiven Verständnisses der unterschiedlichen mentalen Zustände. Basierend auf existierenden Befunden und den Ergebnissen dieser Dissertation ergibt sich folgendes Bild: Kinder entwickeln zunächst ein subjektives Verständnis von Wünschen bevor sie ein solches Verständnis von Überzeugungen entwickeln. Erst danach entwickelt sich ein subjektives Verständnis von Intentionen.

1 General Introduction

‘So,’ said Dumbledore, slipping off the desk to sit on the floor with Harry, ‘you, like hundreds before you, have discovered the delights of the Mirror of Erised [...] It shows us nothing more or less than the deepest, most desperate desire of our hearts. You, who have never known your family, see them standing around you. Ronald Weasley, who has always been overshadowed by his brothers, sees himself standing alone, the best of all of them.’ [...]

‘Professor Dumbledore. Can I ask you something?’

‘Obviously, you've just done so,’ Dumbledore smiled. ‘You may ask me one more thing, however.’

‘What do you see when you look in the mirror?’

‘I? I see myself holding a pair of thick, woolen socks.’ Harry stared.

‘One can never have enough socks,’ said Dumbledore. ‘Another Christmas has come and gone and I didn't get a single pair. People will insist on giving me books.’

It was only when he was back in bed that it struck Harry that Dumbledore might not have been quite truthful. But then, he thought, as he shoved Scabbers off his pillow, it had been quite a personal question. (Rowling, 1997, pp. 229–230)

On the one hand, people's desires and intentions are decisively private. On the other hand, recognizing other people's desires and intentions is of crucial importance for all our interactions. For everyday interactions, it might not be necessary to identify people's deepest or most desperate longings. In most situations, we will ascribe much more basic desires and intentions of immediate importance, as the intention to cross a street or the desire to eat an apple. Desires and intentions are *conative states*. These are states that are directed towards bringing about certain states of the world (e.g., being the most successful of your siblings). Some of these conative states will come to us relatively obvious. We will readily accept that someone who is hungry and sees an apple will desire to eat that apple. In a similar manner, it makes much sense that someone who has lost her parents will hold the desire to be reunited with them. Other desires and intentions pose a much greater challenge and require us to consider them from the other person's subjective perspective. At first glance, everyone would say “books are good for intellectuals”. To understand that this specific intellectual, Dumbledore, does not desire books, we have to relativize to his subjective perspective. From this perspective we can see that he ascribes greater value to simple things like socks or sherbet lemon drops than to pretentious gifts. In most cases, our adult mind does not need a magic mirror to

understand desires and intentions as subjective. We understand that other people have other perspectives on reality which inform and shape their conative states. But how does such a capacity come about? Are we born endowed with the capacities of the mirror Erised (or at least a light version) or do these capacities develop only later? From research so far, we cannot tell. For the most part, research on the ontogenetic development of children's concept of conative states has neglected a fundamental property: That they are subjective. Most research on children's subjective conception of mental states has focused on *cognitive states*. These are states that aim to represent the world as it really is, as knowledge or beliefs. This dissertation aims to address the neglected topic of children's subjective understanding of conative mental states. In particular, I will look at the ontogenetic development of children's subjective conception of conative mental states: When and how do children come to appreciate that conative states are subjective and how is this capacity related to their concept of cognitive mental states?

In what follows, I will first introduce the concept of theory of mind, mental states and why mental states are subjective. I will then provide an overview about what we know about the development of a subjective understanding of cognitive mental states. I will discuss the current body of empirical and theoretical work on the conative states in particular in terms of desires and intentions. These lay the foundation for this dissertation's research questions. Then, I will present the three studies I conducted in the course of this dissertation that provide answers to these questions. I will explain and discuss what I found in this dissertation, what this indicates and how future research can build on this dissertation to achieve a comprehensive understanding of children's concept of others' mental states.

2 Mental States

2.1 Theory of Mind

We perceive other people as rational agents. They feel and experience what surrounds them. They hold beliefs about how the world is, they desire certain outcomes and have intentions how to bring about these outcomes. This gives rational agents practical reasons to act. From a first-person perspective, such a practical reasoning considers how I perceive the world to be (e.g., dark); how I want it to be (illuminated); what I want to do (turn on the light) and what action I think will achieve this outcome (flip the light switch on). These reasons then rationalize my action (flip the switch; Davidson, 1963). From a third person perspective, we can use this reasoning to predict an agent's rational action. We can also explain rational action by reversing this reasoning process to reconstruct the agent's practical reasons. Why did she flip the switch? Because she wanted the room to be illuminated and thought that turning the light switch would bring about this outcome. Thus, to explain rational action we ascribe two kinds of mental states which give an agent practical reason to perform this action: Conative state, how the agent wants the world to be (illuminated) and cognitive or epistemic states, how she believes the world to be (turning the switch will turn on the light, turning on the light will illuminate the room; Rakoczy, 2017a). The ability to ascribe such mental states to agents is called theory of mind (Premack & Woodruff, 1978). Our theory of mind is meta-representational. We are able to represent other agents' subjective representation, in other words, to meta-represent (Perner, 1991). Meta-representation allows us to make sense of rational actions from the agent's perspective. This makes theory of mind one of the most fundamental capacities for practically every social interaction we engage in (Leslie et al., 2004).

2.2 Properties of Mental States

Mental states are representations which an agent holds of the world. We hold and ascribe different kinds of mental states. Cognitive or epistemic states, such as knowledge and beliefs; conative states or pro-attitudes, such as desires and intentions and, affective states, as joy and anxiety. This dissertation will concentrate on cognitive and conative mental states. The following part will describe the fundamental properties of these states. These properties

establish why a full-blown concept of these states necessarily includes a subjective conception and how cognitive and conative states differ in terms of their logical structure¹.

2.2.1 Intentionality (Directedness)

One central property of cognitive and conative mental states is that they are about something (Searle, 1983). I cannot hold a belief without believing that something is the case or hold a desire without desiring something. This property is called *Intentionality* (lat.: *intendere* - being directed towards some target or thing). Note that Intentionality in the sense of directedness is something else than intentionality in the sense of intending something². So, the mental state “She believes that it rains” is Intentional in that the belief is directed at the state of affairs “it rains”. In the same manner, the mental state “She desires to eat the apple” is a desire directed at “eating the apple”.

2.2.2 Propositional Attitudes

Thus, there are two elements to an Intentional state: first, a certain psychological mode or attitude, the state (*S*). These psychological attitudes can be cognitive, as beliefs or knowledge, or conative, as desires and intentions. Second, an Intentional state has a certain representative content. In the case of cognitive and conative mental states, these contents generally refer to whole propositions: that it rains³. Accordingly, they are referred to as propositional contents (*p*). One might argue that some desires appear to refer only to objects: She desires the puppet. Yet, strictly speaking this desire represents the propositional content: that she has the puppet. The psychological attitude is directed towards the propositional content. For this reason, we refer to cognitive and conative mental states as *propositional attitudes* (Searle, 1983). These propositional attitudes are typically reported in the following way:

Verb (which expresses *S*) + embedded subclause (which expresses *p*): *S(p)* (belief (it rains)), reading: She believes that it rains.

¹ The exact definition of the properties of mental states are subject to a major ongoing debate (e.g., Brandl, 1996; Jacob, 2019; Schueler, 1991). However, this debate is not the target of this dissertation. Thus, for the purpose of this dissertation, I will presume definitions as they are presented here.

² To keep the two apart I will refer to Intentionality in the sense of directedness as Intentionality with a capital ‘I’ and to intentionality in the sense of intending something as intentionality with a small ‘i’.

³ Love or hate, for instance, can also be directed towards non-propositional contents: A hates B.

2.2.3 Direction of Fit and Conditions of Satisfaction

Mental states relate to the world in a certain way. They have a certain direction of fit to the world. This direction of fit is fundamentally different for cognitive and conative mental states.

Searle describes direction of fit using the example of Cinderella:

If Cinderella goes into a shoe store to buy a new pair of shoes, she takes her foot size as given and seeks shoes to fit (shoe-to-foot direction of fit). But when the prince seeks the owner of the shoe, he takes the shoe as given and seeks a foot to fit the shoe (foot-to-shoe direction of fit). (Searle, 1983, p. 8)

An analogous logic applies to Intentional mental states. Our mind (mental states) stands in a certain relation to the world. But just as in Searle's Cinderella example, the direction of this fit is opposite for cognitive and conative mental states. Cognitive mental states aim to match the world, to bring their propositional content in line with reality. In other words, they aim at being accurate. They have a *mind-to-world direction of fit*. Anna's belief that it rains aims to be accurate in that it fits with the weather as it is in reality. In contrast, conative mental states have a *world-to-mind direction of fit*. They aim at adjusting the world to their propositional content. My desire to have ice-cream aims to adjust the reality in which I do not have ice-cream to the desired state in which I have ice-cream¹ (Searle, 1983; Smith, 1987).

The direction of fit of Intentional mental states brings about certain *conditions of satisfaction* that aim to be satisfied. As a result of the opposite directions of fit, cognitive and conative mental states also have different conditions of satisfaction. According to their mind-to-world direction of fit, a cognitive state holds the condition of satisfaction that the state's propositional content fits the reality. The cognitive state will be true if it matches the reality, but false if it does not. Hence, we evaluate cognitive states against the normative standard of truth. Take the example of Anna: She believes that it rains. Her belief will only be true, if in reality it rains. If the sun shines, the conditions of her belief will not be satisfied. In consequence, her belief will be false.

¹ There are also some cognitive and conative states, though, that are not directly aimed at matching the world or make the world match. I can hold the fantasy to be at the French Riviera or wish for snow to be purple (Paul, 2020, p. 70).

This is different for conative mental states. These aim at being fulfilled or realized. If a conative state does not fit the world, this does not mean it is false. It only shows that the world still has to be modified to fit the conative state. Conative mental states differ in their conditions of satisfaction. Desires and intentions are both pro-attitudes that aim to bring about a certain outcome. However, desires aim at changing a certain state of the world. The conditions of satisfaction of a desire are that this state's propositional content is brought about. When this is the case, the desire is fulfilled. Intentions aim at bringing about a certain action that functions as a means to a certain end. They can be either realized or not realized. Thus, in contrast to desires, intentions involve the commitment to actually perform the action in "the right way", as it was intended (Astington & Gopnik, 1991; Bratman, 1987; Searle, 1983). Furthermore, intentions are causally self-referential. They are only realized if the intended state of affairs is brought about by the exact action that is specified in the content of the intention and if the intention itself causes this action. If an action is not caused by the intention itself, the intention is not realized even if the action looks similar and achieves the intended outcome. My desire to have an apple is fulfilled when I have an apple, irrespective of how I have obtained this apple. This is different for intentions. Let us assume I form the intention to go to the supermarket and buy an apple. For the intention to be realized, I have to perform the particular action my intention commits me to and it has to be my intention that causes me to perform this action. The intention will not be realized if my neighbor comes by and brings me an apple. Also, it will not be realized if it is not my intention that causes this action. This is the case in deviant causal chains (Harman, 1976; Searle, 1983): Suppose, I have formed this intention in the morning. In the afternoon I go for a walk. It starts to rain and I take shelter from the rain in the nearby supermarket. I am hungry so I make use of this current situation and buy some food. Among this food is an apple. In this case, I have performed the intended action but not because of my original intention. Thus, my intention has not been realized. It has to be my intention to go to the supermarket and buy an apple that makes me go to the supermarket and buy an apple. Only then, my intention's condition of satisfaction is met, and my intention can be evaluated as realized.

As we can see, cognitive and conative states differ fundamentally in their directions of fit and conditions of satisfaction. Cognitive states have to be evaluated against the normative standard of being true because they aim at fitting the world. Conative states aim to change the world so that it fits their content. They cannot be false or true. Desires are evaluated against

the normative standard of being fulfilled, intentions against the standard of being realized (in a causally self-referential way).

2.2.4 Truth Value of Propositional Attitudes

A regular statement is true if the content is congruent with reality. The statement “Anna is standing in the rain that is heavy.” is only true if it is really raining, the rain is really heavy, and Anna is really standing in the rain. This is different for mental states. The report “Anna believes that it rains” can be true even if it does not rain (Josef Perner, 1988, pp. 26–27). Why is that the case? Mental states have different directions of fit. These result in different conditions of satisfaction. Beliefs can be true or false, desires fulfilled, and intentions realized. But only because a mental state does not satisfy the conditions of satisfaction, this does not mean that an agent cannot hold it. Take the example “Anna believes that it rains”. As we know, this mental state has an attitude (belief) that is directed towards the propositional content (it rains). The conditions of satisfaction of this belief will be fulfilled if it really rains. If this is in fact the case the belief will be true. However, if it does not rain the belief will be false. Yet, even if the conditions of satisfaction of a mental state are not satisfied, an agent can still hold this state. Consider the following scenario: The weather forecast has predicted that it would rain all day long. Anna has not looked out of the window but based on the forecast *she believes that it rains*. If she had looked out of the window, she would have seen that in reality it does not rain but that there is bright sunshine. Hence, although the belief is false, it makes sense from her subjective perspective to hold this false belief. Thus, when we ascribe the propositional attitude “She believes that it rains” to Anna, the truth value of this report does not depend on whether it rains or not. The truth value of the reported propositional attitude depends on whether Anna believes that it rains. Likewise, the truth value of reported conative mental states depends on whether the agent’s desire or intention really represents its content in this way. Accordingly, to correctly ascribe a propositional attitude the ascriber has to meta-represent how the agent represents the content of her mental state (De Villiers & De Villiers, 2000).

This provides us with an answer to why a full-blown concept of mental states must include subjectivity. To reason about an agent’s mental state, it is not sufficient to compare the content against some normative standard on an objective level. Imagine, we would reason about Anna’s belief on an objective level. We could only rely on the fact that it does not rain. Thus, when Anna leaves the house, we would incorrectly predict that she will not take the umbrella. We have to represent how Anna represents the content of her belief from her subjective perspective. If we take her subjective perspective, we can represent that because of the forecast

she mis-represents the reality and believes that it rains. This allows us to arrive at the correct prediction that Anna will take the umbrella when she leaves the house. Thus, to explain and predict rational action, the ascriber has to relativize to the agent's subjective standpoint and meta-represent her representation of the content.

2.2.5 Aspectuality and Intensionality-with-an-s

Thus, ascribing propositional attitudes requires relativizing to the agent's subjective standpoint. This relativization commits us to one fundamental feature of propositional attitudes: that they are aspectual (Anscombe, 1957; Nelson, 2019; Searle, 1983). They only represent their content under some specific descriptions or aspects. Consider the famous example of Greek mythology: Oedipus and Yocasta. One way to refer to Yocasta is under the description or aspect *beautiful woman*. Another description is *Oedipus' mother*. Thus, *beautiful woman* and *Oedipus' mother* are coreferential terms. If I make a regular statement about Yocasta I can substitute these descriptions without changing the truth value of the statement:

The beautiful woman has brown eyes.

Oedipus' mother has brown eyes.

Under the premise that Yocasta really has brown eyes, and that *the beautiful woman* really is *Oedipus' mother* both statements are true. We call these contexts that allow for the substitution of coreferential terms *extensional*. This is substantially different when we report an agent's mental states. In the myth of Oedipus and Yocasta, Oedipus does not know that Yocasta, the beautiful woman, is also his mother. He falls in love with the beautiful woman and marries her. So how can we report Oedipus' belief?

Oedipus believes that he married *the beautiful woman*.

This is definitely true. But what happens if we substitute the description in our report:

Oedipus believes that he married *his mother*.

This is not true. Oedipus would not agree to this report of his belief. When the ascriber reports a propositional attitude, this is an *intensional* (with an s) context. Such an intensional context commits the ascriber to the description under which the content is represented, in this case *beautiful woman*. Thus, in contrast to extensional contexts, in intensional contexts the substitution of coreferential terms may change the truth value of the sentence. In the case of Oedipus, we know that he is ignorant of the second description. But even in cases where it is

not explicitly stated which descriptions the agent represents, we cannot simply substitute coreferential terms. We cannot be sure that the agent represents this description.

Ascribing mental states creates an intensional context. This applies equally for cognitive and conative mental states. We can say Oedipus desires to marry the beautiful woman, but we cannot say that he desires to marry his mother. Likewise, we can say that he intends to marry the beautiful woman but not his mother. Coreferential terms are not restricted to identities and proper names (Clark Kent = Superman), but can also be terms as predicates (75% empty = 25% full) or descriptions of actions (raise your hand = indicate that you have a question; Bermúdez, 2020; Goldman, 1970). That mental states are aspectual becomes relevant in all kinds of everyday situations: You intended to eat the tasty cookie but not to eat Lisa's last cookie. You believed that you threw away a random piece of paper but not your shopping list.

The phenomenon of aspectuality in intensional contexts thus adds to the question why a full-blown concept of mental states must appreciate that mental states are subjective. To correctly ascribe mental states, we need to relativize to the agent's subjective standpoint to consider under what description she represents the propositional content.

2.2.6 Implications of the Properties of Mental States

The previous sections have provided answers to the two questions I raised at the beginning. First, how do cognitive and conative states differ in terms of their logical structure? Cognitive and conative states have different directions of fit. Cognitive states aim to fit the world and conative states aim to align the world with the state they represent. This brings about different conditions of satisfaction and normative implications. As I will describe later, this is of substantial importance for the way we measure children's subjective understanding of the different kinds of states.

Second, why does a full-blown concept of mental states have to include a subjective conception? To correctly ascribe a mental state, the ascriber has to represent how the agent represents the content of her state from her subjective perspective. Moreover, it is necessary to relativize to the agent's subjective standpoint to identify under which descriptions she represents the propositional content.

3 Developmental Perspective

The following part will give an overview on how a subjective conception of mental states develops ontogenetically and theories on why it develops this way. Most empirical work targets the subjective conception of cognitive states. Probably as a consequence of this more extensive empirical basis, theoretical work on the development of theory of mind is also more specific and differentiated for cognitive states. I will first describe this work. Then, I will discuss why it is at first glance counterintuitive that research on theory of mind development has focused on cognitive states and describe possible reasons behind this focus. I will give an overview of research on conative states. In the end, I will propose plausible developmental trajectories that comprehensively consider cognitive and conative mental states.

3.1 Cognitive States

In what follows I will give an overview about the ontogenetic development of children's understanding of other agents' cognitive states. I will first describe the standard test for children's subjective conception of cognitive states. The next section will describe empirical work on forms of epistemic (knowledge and belief related) reasoning at earlier ages: Children's early developing ability to consider that an agent only knows about what she has perceptual access to. And evidence as well as counterevidence that already in infancy children can solve implicit versions of the false belief task. I will then introduce theoretical work on children's theory of mind development that builds on this rich empirical foundation.

3.1.1 The Standard Test for Subjective Understanding of Cognitive States

Testing for a subjective understanding of cognitive states is very straightforward. Cognitive states have a mind-to-world direction of fit. A cognitive state will not meet its conditions of satisfaction if it does not represent the reality accurately. For instance, a belief will be false if it represents reality inaccurately. Accordingly, to appreciate that an agent can hold a belief even though it is false, the ascriber has to take the agent's subjective perspective. Thus, the agent has to acknowledge two facts: A state can mis-represent reality (represent it to rain, which is false because the sun shines). And another agent can have a different representation of reality than the ascriber or other agents (Anna has watched the weather forecast and believes that it rains. Susi has looked out of the window and believes that it does not rain). Accordingly, when the ascriber ascribes a (mis-)representation to an agent she has to meta-represent the agent's representation of reality (Rakoczy, 2017a). This provides a straightforward approach to test for a subjective conception of cognitive states: If someone is able to appreciate that an agent can

hold a belief which is false, she has proven her capacity to reason about cognitive states on a subjective level.

Interestingly, it was a study on apes' understanding of conative states that initiated the development of a paradigm that tested for subjective understanding of cognitive states in humans. In Premack and Woodruff's (1978) famous study, the chimpanzee Sarah watched videos of a human actor who faced a certain problem (e.g., she could not reach a banana). Sarah then had to choose the correct solution (a stick) among several pictures that proposed solution which would not solve the problem. The authors interpreted Sarah's success on this task as evidence for her capacity to understand the actor's purpose. Yet, replies to the article reduced Sarah's performance to associative processes. They pointed out that this task cannot dissociate between Sarah's representation of the problem and Sarah's representation of the actor's representation of the problem. However, these critiques also revealed how a subjective conception of other agents' states could actually be tapped. They proposed a framework for a task that requires the ascriber to predict the action a rational agent has reason to do because of her subjective representation: The ascriber learns that the agent believes that p and desires q . p actually deviates from the ascriber's own representation of reality. From p and q , the ascriber can infer that the agent will x and therefore anticipate x (Bennett, 1978; Dennett, 1978).

Building on this logic, Wimmer and Perner (1983) then developed a task that could be applied to young children: the *change-of-location* task. Since then, this task has become a standard task and the litmus test for children's appreciation of other agents' subjective cognitive states. The rationale of this task is as follows: Agent A puts an object O in box 1 and leaves. In her absence, O is moved to box 2. A then returns. Children have to predict where A will look for O. To correctly predict that A will look for O in the initial location, box 1, children have to ascribe the false belief (O is in box 1) to A. They have to meta-represent her misrepresentation. Another classic paradigm that follows a similar logic but a different procedure, is the *unexpected-content* task (Hogrefe et al., 1986). In this task, children see a prototypical box (e.g., a Smarties package). They are asked what they think is inside the box (Smarties). The experimenter then reveals that there are actually pencils inside the box. To test for their subjective understanding, children are then asked what their mother/father (A) outside will say is in the box if she/he sees the box for the first time. In another adaptation, children were asked what they themselves initially thought was in this box (Gopnik & Astington, 1988). Both types of tasks revealed that children considered A's or their former self's misrepresentation only by the age of four.

The standard false belief tasks faces skepticism going in both directions. Children's capacity might be underestimated: Children only fail these tasks because of some extraneous demands, as linguistic or inhibitory demands (e.g., Baillargeon et al., 2010; Carruthers, 2013; Leslie, 2005). However, they also faced skepticism that children's capacity is overestimated. Such skepticism doubts that children really use proper propositional attitude concepts to solve these tasks (e.g., Fabricius et al., 2010; Lalonde & Chandler, 2002). Irrespective of this criticism, the two paradigms became and still are the standard litmus test for a subjective understanding of other agents' cognitive states. The following section will give an overview on when children come to solve these kinds of tasks and what early forms of reasoning about cognitive states precedes a subjective understanding.

3.1.2 Ontogenetic Development of Understanding of Cognitive States

3.1.2.1 4-year Revolution.

The initial findings, that children solve the false belief task only by age four, turned out to be an intriguingly consistent pattern. By that age, children correctly predict that the agent A will look for the object O in the initial location as well as that A will say there are smarties in the box and that their former selves thought so, too. Children younger than four fail all three test questions. Interestingly, they do not fail because they answer randomly. Rather, they constantly give the reality-congruent answer: that A will look for O where it really is (box 2) and that A will think and their former selves would have thought that there are pencils in the smarties box. This pattern kept showing up over numerous studies and different adaptations. Children solve the false belief tasks irrespective of which type of task they receive, the nature of A and O (puppet or toy; real or shown on a video/picture) and whether O is still present in the end (see Wellman et al., 2001 for a meta-analysis). Moreover, the capacity to solve these tasks develops approximately around the age of four to five years in cultures all over the world (Callaghan et al., 2005; Shahaeian et al., 2011; Wellman et al., 2001). The fairly clear picture that emerges from these results is the following: Children can reason about an agent's action based on their false beliefs by the age of four. Thus, by this age they understand that an agent can misrepresent reality and that this reality diverges from their own representation. They have developed a subjective conception of cognitive mental states (Rakoczy, 2017a).

As we have seen, propositional attitudes are not only subjective in that they can be true even though their content does not fulfill its conditions of satisfaction. They are also subjective in that they create intensional (with an s) contexts. Thus, to demonstrate a genuine subjective

understanding, children have to be aware that aspects or descriptions of the propositional content cannot be substituted by coreferential terms. First studies indicated that children developed a concept of the aspectuality of beliefs only around the age of six to seven. In one of these studies, children were told about an agent who was aware of one description of a person X (e.g., the thief who stole her watch), but unaware of another description Y (the man with curly red hair). They found that even 6-year-old children did not object to the substitution of the coreferential terms X and Y: The agent thinks that she must find the man with the curly red hair (Russell, 1987). Apperly and Robinson (1998, 2001) developed a less verbal task. There are two objects: a die and an eraser. The child learns that the die is also an eraser. An agent appears. Because the agent only arrives now, she does not know that die is also an eraser. The agent needs an eraser. Only by age seven, children correctly predicted that she will definitely take the eraser-eraser and not the die-eraser. Younger children chose randomly between both options (for different approaches finding similar results, see Apperly & Robinson, 2003; Hulme et al., 2003; Sprung et al., 2007).

At first glance, these results suggest that younger children do not yet appreciate that beliefs represent their content only under certain descriptions. Thus, a notion of cognitive states that accounts for this property appears to develop only by the age of six or seven. However, more recent evidence attributes children's difficulties in these tasks to the substantial extraneous demands of these tasks. Even tasks that did not have substantial linguistic demands posed high demands in terms of reference resolution. If children are asked to point out the eraser, they have to consider the following: If someone knows that there is an eraser and a die-eraser, she will refer to the die-eraser as "die-eraser" and not "eraser". Thus, the term "eraser" has to refer to eraser-eraser and not to die-eraser. Before children master this communicational challenge¹, they do not grasp to which object the term "eraser" refers and choose randomly (Rakoczy et al., 2015). And indeed, in a task that reduced linguistic task demands and at the same time did not make reference resolution necessary, children succeeded at the same age they solve standard false belief tasks. The task followed the basic procedure of the change-of-location task. The agent observed that a pen was put in box 1. She leaves and the child learns that the pen is also a rattle. When she returns the experimenter moves it "only as the rattle" to box 2 (she hides the pen in her hands but rattles it while moving it). Already by age four,

¹ This is even difficult for adults (Keysar et al., 2003).

children correctly predicted that the agent would look for the pen in box 1 (Rakoczy et al., 2015; see also Rakoczy & Oktay-Gür, 2020). The findings of this fundamentally simplified task suggest that the belief concept of 4-year-olds does already appreciate under which descriptions beliefs are held (Rakoczy, 2017b).

So, it seems that children develop some form of subjective understanding of cognitive mental states around the age of four. This includes the meta-representational capacity to represent an agent's mis-representations. There is also evidence that they understand that an agent represents this content only under a particular description. But what happens before children develop such a subjective conception of cognitive mental states?

3.1.2.2 Earlier Forms.

Already from early on, children engage in interactions that appear to require at least some form of epistemic perspective taking, for instance joint attention. This clearly suggests that they cannot be completely unaware of other agents' epistemic states. From early on children seem to be capable to reason about what an agent has perceptual access to and what she knows and does not know because of this access. I will give examples of empirical work that support this claim. But before, let us turn to the question why children might be able to reason about knowledge before they can reason about beliefs. Consider the following analogy. Anna lives on an island for quite some time. She knows every corner and has drawn a perfectly accurate map of the island. Her friend Max visits her. He has never been on that island before and decides to explore it on his own. Every evening he tells Anna where he has been. Anna draws a second map from Max' perspective. At the beginning, the Max-map is completely empty. Every evening she copies from her map the parts Max has been to. One day, Max' explanations do not fit Anna's own map. He admits he might have got lost but he is still convinced of his description of what he saw where. For Anna's Max-map this means she cannot simply fill empty parts by copying from her map. Instead, she has to listen carefully and map everything how Max describes it. This analogy shows two things: First, for the earlier entries Anna did not need to take Max' perspective. She could simply apply the objective logic: If he has been at X, he has seen Y. Only for this last entry, she had to construct the island from Max' subjective perspective. Second, one can easily see how much more demanding this last entry is compared to the earlier entries.

The same applies to ascribing knowledge and belief. Ascribing whether someone knows something or not, because she has (no) perceptual access, only requires children to

consider *what* the agent can or cannot perceive (called “level I perspective-taking”; Flavell et al., 1978, 1981). If the agent has not seen something, she has only partial representation. Ascribing beliefs requires children to understand *how* the agent sees the same thing they saw depending on her viewpoint (e.g., the same digit can be seen as a “6” or a “9”; called “level II perspective-taking”; Flavell et al., 1986). The agent can thus represent reality in another way or mis-represent reality. As we have seen, ascribing diverging (mis-) representations requires the ascriber to relativize to the agent’s subjective standpoint. Note however, that a proper understanding of knowledge would also consider how (under which description) an agent represents the content. Yet, it is possible to decide whether an agent represents or does not represent something by simply tracking on an objective level what she had access to.

This difference in complexity is mirrored in children’s ontogenetic development. A notion of what an agent knows develops early in childhood, significantly earlier than what she believes (Peterson et al., 2012; Wellman & Liu, 2004). This early competency becomes visible in numerous different contexts. Already in infancy, children use pointing to provide information to other agents (Liszkowski et al., 2006) and even consider which information the agent already has when doing so (Liszkowski et al., 2008). They can also consider an agent’s knowledge states to predict actions. If an agent always reaches for A but not for B, they will expect her to also reach for A in the future. In contrast, if the agent cannot see (i.e., is ignorant of) B they will not predict that she will reach for A (Luo & Johnson, 2009). Likewise, they expect that if a third agent who knows about an agent’s preference for A will give A to the agent. Yet, if the third agent does not know about this preference they do not expect her to give A (Vouloumanos et al., 2014). Even in more explicit tasks, children begin to reason about an agent’s knowledge before her belief. They appreciate how perceptual access determines an agent’s knowledge state (Pratt & Bryant, 1990) and consider this state when they choose whom to ask for information already at the age of three (Pillow, 1989). Also, even children who fail classical belief tasks or tasks on the aspectuality of beliefs succeed on the knowledge control question: They can state correctly that the agent does not know that O is in box 2 or that the agent is ignorant of a certain description (e.g., Fabricius et al., 2010; Oktay-Gür et al., 2018; Perner et al., 2015; Rakoczy et al., 2015; Sprung et al., 2007). Even on a linguistic level, children start using the word “know” earlier than the word “think” (Bartsch & Wellman, 1995). Drawing on such evidence, it seems that an objective form of reasoning about cognitive states

(namely partial representations) appears to develop before children can reason subjectively about mis-representations¹.

Theory of mind research has also looked into the possibility of an early potentially subjective form of reasoning about mis-representations. This research applied non-verbal or at least significant less verbal adaptations of the standard false belief task. At first, evidence from these tasks suggested an early implicit form of subjective reasoning. Yet, more recent evidence cannot replicate these results and raises substantial skepticism about these tasks' reliability and validity (for an overview, see Kulke et al., 2018).

In these implicit tasks, children were confronted with a change-of-location paradigm. But in contrast to the traditional task, children were not asked to predict the agent's behavior. Instead, their eye movement was tracked and analyzed for signs of anticipation and violation of expectation. Children observed a change-of-location procedure. When A reappeared, even infants correctly looked at the initial location where the agent believed the object was, apparently anticipating that A will go there (e.g., Clements & Perner, 1994; Southgate et al., 2007; Surian & Geraci, 2012). To measure for violation of expectation, children also saw the change-of-location procedure. But after A reappears, she approaches one of the two locations. Infants looked longer at the scene when A approached the belief-incongruent (the current) location than when she approached the belief-congruent (the initial) location. Longer looking times are interpreted to indicate that children are surprised, because their expectation has been violated (e.g., Onishi & Baillargeon, 2005; Surian et al., 2007). Also, a third interactive approach revealed a relatively early competency. Here, the change-of-location paradigm was embedded in a helping scenario: A returns after O has been moved to box 2. A approaches box 1 and tries to open it. However, she is unable to do so and asks the child for help. Instead of opening box 1, children opened box 2 and gave O to A. Interestingly, they helped to open box 1 in a control condition in which A had observed the change. Children seemed to have understood that in the first case the agent mis-represents O's location and thus tries to open the wrong box. In contrast, when she correctly represents that A is in box 2, they seem to expect

¹ Though see Phillips and Norby (2019) who postulate that children's early reasoning capacities demonstrate not only a grasp of knowledge in the sense of perceptual access, but a proper understanding of knowledge.

A to really want to open the empty box 1 (e.g., Buttellmann et al., 2009; Knudsen & Liszkowski, 2012; Southgate et al., 2010).

These results set forward a controversy about the nature of this early implicit belief reasoning (Christensen & Michael, 2016). It was interpreted as evidence that standard false belief tasks underestimated performance due to their high linguistic and cognitive demands (Carruthers, 2013; Leslie, 2005; Scott & Baillargeon, 2017). Yet, there were also different forms of criticism. More skeptic positions interpreted this early capacity as a certain sensitivity to beliefs but not the subjective understanding measured via standard verbal false belief tasks (Apperly & Butterfill, 2009; Low et al., 2016; Priewasser et al., 2017). Deflationary accounts explained children's responses by low-level behavioral cues as direction of attention or preference for novelty (Heyes, 2014; Josef Perner & Ruffman, 2005; Sirois & Jackson, 2007).

However, as impressive as these results were, they turned out to hardly replicate. This appears to be the case for all approaches (e.g., anticipatory looking: Kampis et al., 2020; Kulke et al., 2017; Kulke & Hinrichs, 2021; Grosse Wiesmann et al., 2018; violation of expectation: Grosse Wiesmann et al., 2017; Poulin-Dubois & Yott, 2018; Powell et al., 2018; helping: Dörrenberg et al., 2018; Grosse Wiesmann et al., 2017; Wenzel et al., 2020). Moreover, there is evidence that suggests that not even adults show belief-congruent anticipatory looking (Kulke et al., 2019). Currently, the topic is subject to a multi-lab project. Labs of both positions cooperate to systematically test the task's validity and reliability (Schuwerk et al., 2021). Thus, at the moment, we cannot tell whether infants can reason implicitly about beliefs. Yet, based on the number of non-replications these findings should be considered with caution.

In conclusion, the following picture emerges: Already in infancy, children develop a concept of cognitive states that allows them to reason objectively about an agent's partial representations (what she knows/does not know). That infants can reason in an implicit manner about an agent's mis-representations seems rather unlikely. Children develop a subjective conception of cognitive states around the age of four. At that age, they acknowledge that agents' representations can diverge and can mis-represent reality. First evidence suggests that at the same age they can also consider that beliefs are subjective in that they are aspectual.

3.1.3 Theoretical Work on Theory of Mind Development

This rich empirical evidence has provided the basis for extensive and differentiated theoretical work on how children develop the ability to meta-represent other agents' representations from the agents' subjective perspective (for an overview, see Rakoczy, 2017a). While this work is

not restricted to cognitive states, most theories are much less specified and differentiated for conative states. This might be the effect of the asymmetry in empirical evidence which is much more extensive for cognitive states.

The nativist perspective suggests that theory of mind is innate or, at least, present from very early on. Such accounts suggest a modular form of theory of mind that allows infants to explain action by automatically ascribing mental states (Carruthers, 2013; Fodor, 1992; Leslie, 2005). In explicit but not implicit false belief tasks this competence is masked by extraneous performance factors such as the high amount of information that needs to be processed (e.g., Baillargeon et al., 2010; Carruthers, 2013). What allows children to solve standard false belief tasks at age four is only the capacity to master these extraneous demands.

Nativist accounts stand in contrast to accounts that propose a conceptual change. For instance, simulation theories suggest that we build on our own mental states to simulate other agents' mental states. Simulation accounts differ in the nature of this simulation they argue for. One proposed form of simulation operates on introspection. The ascriber introspects what she thinks and feels in this context and uses this introspection to simulate the agent's mental states (Goldman, 1993). Alternatively, it has been argued that the ascriber does not introspect but pretends to view the world from the agent's perspective. Then, she uses her own processes of thinking in an offline mode to simulate the agent's mental states (Gordon, 1986). Children's competence in ascribing mental states to other agents increases with their experience in such perspective taking. Another theoretical perspective are theory-theories. These assume that we think about mental states as theoretical terms. For instance, we have a theory about how a belief is related to the reality the believer perceives; how it is related to other mental states; and how it is related to a certain output. Thus, theory of mind development takes the form of theory revision. It requires children to formulate and revise theories or to abstract regularities they encounter in everyday life (e.g., Gopnik & Wellman, 1994).

Another line of theoretical work reflects the contrast of early and later developing competencies that empirical work has revealed. This work postulates two distinct systems. The following section will introduce two-system accounts. Two formulations of two-system accounts will be described in detail. These propose systems that engage in different levels of perspective taking.

3.1.4 Two-System Accounts

The idea of two forms of reasoning about cognitive mental states is even older than the standard false belief task. As mentioned before, Flavell and colleagues (1981) proposed to distinguish two different levels of visual perspective taking: level I (what) and level II (how). Three-year-olds could easily tell *what* an agent could and could not see of a partly covered picture of a turtle. Yet, they could not tell *how* an experimenter who sat opposite from them would see a turtle on a picture that was put between them: As standing on its feet or lying on its back. It later turned out that this distinction between level-I and level-II perspective taking corresponded to children's reasoning about epistemic states. Level-I perspective taking was sufficient to reason about what an agent knows but level-II was necessary to represent how an agent (mis-)represent the world (Flavell et al., 1986).

In related manners, theoretical work has proposed two systems¹ that account for early and late developing forms of theory of mind. The accounts differ in their motivation (e.g., knowledge - belief, implicit - explicit theory of mind, or capacities in other species) as well as in their exact elaboration (e.g., Apperly & Butterfill, 2009; Doherty, 2006; Gordon, 1986; Newen & Wolf, 2020; Perner, 1991; Perner & Roessler, 2010; Tomasello, 2018). Yet, they all share the following logic: One early developing system relies on relatively basic notions to explain rational action. Only the later developing system functions on a sophisticated concept. The following part will describe two very comprehensive two-system accounts.

3.1.4.1 Teleological Accounts.

Perner and Roessler introduced an account that differentiates between an early developing objective teleological² (goal-oriented) form of reasoning and a later developing capacity to reason on a subjective level about mental states (Perner et al., 2018; Perner & Esken, 2015; Perner & Roessler, 2010, 2012). The account builds on the basic idea that in a wide range of situations, it is sufficient to predict and explain rational action based on an objective teleological reasoning strategy. This simple form of reasoning interprets actions as directed towards bringing about a certain goal. Agents will perform actions they objectively have good

¹ Newen and Wolf (2020) actually propose a third intermediate system.

² Ancient Greek.: τέλος – 'end', 'aim', or 'goal'. Already Aristotle introduced the idea that actions can be explained teleologically as directed towards achieving something good (Charles, 2012).

reason to perform, because they will achieve an objectively good goal. When I aim to satisfy my hunger I have good reason to go to the cafeteria and eat something (Perner et al., 2018, p. 99). This objective teleological reasoning schema functions on the following features: First, it takes a teleological stance. An action is directed towards achieving a certain goal state. Second, this goal has a positive value. It has to be objectively desirable. Even if the positive value is minimal. Third, if doing X brings about the desired goal state and Y is the best way to do X there is normative reason to do so. If a baby needs to be fed, feeding her (X) brings about this goal state. If the mother is nearby and can breastfeed her (Y) we would expect her to do so. If the mother is not there, but the father who has access to the bottle therefore can bottlefeed the baby (Y') we would expect him to do that (Perner & Esken, 2015, p. 75). Fourth, the reasons behind an action are objective facts not some form of representations. It is a fact that eating food satisfies hunger and that food can be obtained in a cafeteria. So based on these facts, I have good reason to go to the cafeteria. And fifth, these facts are publicly available. They do not only give me reason to act (and go to the cafeteria) but potentially they also give you reason to act. If you go to the cafeteria anyways, the publicly available fact that my hunger can be satisfied by food provides you with good reason to bring some food for me, too.

Perner and Roessler (2010) propose that infants start as pure teleologists. In an early step, infants then come to incorporate the explanatory role of knowledge, recognition and perception: If an agent is ignorant of a reason-giving fact or does not recognize it as a reason, this reason cannot give her reason to act. Whether or not someone is aware of a reason-giving fact is publicly available (e.g., she was not there and therefore has not perceived that her baby cried). Equipped with this teleological reasoning schema, young children can reason about rational actions without taking the agent's subjective perspective. But even though the explanatory power of this schema is extensive it has a substantial limitation: The possibility of mis-representations. If I mis-represent that the cafeteria is still open, because I have confused the days, I will have, subjectively, good reason to go there. Although this is all rational, we cannot reason about this solely on objective facts. We have to regard the agent as someone who acts for normative reasons but from her own perspective. Accordingly, these reasons will build on *her* (mis-) representations and what *she* finds desirable. We have to engage in teleology-in-perspective.

Teleology-in-perspective transfers the teleological schema to the agent's perspective. We create experiential records for each agent. If an event occurs (the baby cries) within the agent's informational field, we can enter the fact in her record. If the event did not occur within

her informational field, it will not be entered. Thus, for the standard false belief task, we enter in A's record: O is in box 1. But we do not enter: O is transferred to box 2. To predict A's action, we need to deliberate what someone who has A's experiential record would have good reason to do. In A's record there is no information that O has been transferred to box 2. Hence, we can predict A's action by the following counterfactual argument: "If O had not been transferred to box 2, one would have good reason to look for O in box 1". Thus, to engage in teleology-in-perspective and reason about mental states subjectively requires the capacity to reason counterfactually. Support for this theory comes from a large body of evidence that finds close relations between false belief tasks and counterfactual reasoning capacities (for meta-analytical overviews see Rafetseder et al., 2021; Rafetseder & Perner, 2018). This relation alone could also be caused by shared extraneous task demands. However, more recent evidence finds that children's counterfactual reasoning capacities were directly mirrored in their answers to false belief tasks. Older children and adults who have a mature counterfactual reasoning capacity keep all facts fixed except for the changed fact. They stick to the nearest possible world. In contrast, younger children also change other facts in their reasoning (Leahy et al., 2014). Rafetseder and colleagues (2021) engaged children in variations of false belief tasks. These could only be solved if children stuck to the nearest possible world and changed only the one fact the agent had not observed. While older children applied this strategy, younger children failed these tasks in the same way they failed respective counterfactual reasoning tasks: They changed facts the agent actually knew. This is interpreted as evidence that counterfactual reasoning does not only support belief reasoning but is the substantial foundation on which subjective belief reasoning builds.

So, according to Perner and Roessler's teleological account children first reason on a purely teleological reasoning schema. This allows them to explain rational action on an objective level. Later, they develop the capacity to reason about rational action on a subjective level. This requires them transfer their teleological reasoning to the agent's perspective and consider from her perspective what she has reason to do.

In a related approach, Gergely and Csibra (2003) propose a teleological stance to explain why already infants have some concept of rational action but fail more sophisticated practical reasoning. In early habituation studies, even infants appeared to understand how actions, goals, and constraints were related. Children in these studies observed a circle that encountered some form of obstacle which it had to surround. In the next step, the obstacle was removed, or it became apparent that there was no goal to achieve. Infants were surprised when

the circle acted irrational and continued to perform the now unnecessary surrounding action (Csibra et al., 1999, 2003; Gergely et al., 1995). To explain why young children reason correctly about the relation of action, goal state and situational constraints but fail false belief tasks, this account proposed two different stances: A teleological stance which is already present in infancy and a later developing mental stance (Gergely & Csibra, 2003). Both stances comply with the principle of rationality but on different levels. The teleological stance assumes that there is a certain goal state, certain situational constraints and an action that can achieve this goal state. This gives the rational agent practical reason to act. In contrast, the mental stance considers how the agent represents this reality. Her desire represents her goal state. Her belief represents the situational constraints from her perspective. And her intentions represent by which action she intends to bring about the outcome she desires. The teleological stance is sufficient to explain why an agent changes her action to achieve the goal state if the constraints have changed. Yet, it does not suffice to succeed on the false belief task: Here, the goal state is “obtaining O”. The situational constraints show that “O is in box 2”. The most rational action to achieve “having O” is “to go to box 2”. Thus, applying the teleological stance, A has reason to go to box 2. In contrast, the mentalistic stance can consider that A believes that O is in box 1. Together with A’s desire to obtain A, this gives A good reason to go to box 1.

Summarizing, both teleological accounts propose an early teleological form of practical reasoning. Yet, this teleological reasoning is limited to the objective reality. Only a later developing form of reasoning allows to reason about this reality from the agent’s perspective.

3.1.4.2 The Two-System Account by Apperly and Butterfill.

Apperly and Butterfill’s account was motivated by the surprising evidence of an early developing implicit concept of beliefs (Apperly, 2010; Apperly & Butterfill, 2009; Butterfill & Apperly, 2013; Low et al., 2016). While the recent skepticism on this evidence challenges some aspects of their theory, their differentiated proposal of the nature of theory of mind remains of great informative value. In their main logic, Apperly and Butterfill appeal to the well-established separate systems for number cognition. Infants (and non-human animals) have a limited but useful basic concept of numbers. They can automatically identify numbers of up to three to four items and can distinguish between sets of items if their ratio is sufficiently large (they can differentiate sets of 100:300 but not 20:25). A later developing system then provides a full-blown concept of numbers (Carey, 2009; Feigenson et al., 2004). Relating this to theory of mind, Apperly and Butterfill propose two distinct systems that can account for the tension between flexibility and efficiency of mental state ascription: In certain situations, we are very

efficient. We reason automatically about an agent's beliefs but do not elaborate any further about her perspective. Just as in automatic number cognition, where on an automatic implicit level we only take an approximate approach. In other situations, we ascribe mental states very flexibly. When we explicitly elaborate an agent's mental state, we consider all information available. The two systems then take effect respectively. The first system relies on simple representational capacities. These provide efficiency. But this efficiency comes at the expense of flexibility. The second system relies on more complex capacities, which make it less efficient but therefore more flexible. Ontogenetically, the first system is already present early in infancy. The second system emerges only later around the age of four, when children's meta-representational skills as well as their language and executive functions are sufficiently pronounced. The first system remains operating along with the second. This allows adults and children from the age of four to access both systems depending on the situation.

The first efficient system is limited in the following way: It cannot process proper propositional attitudes. Instead, it relies on belief-like states that serve as proxies for beliefs. These are relational attitudes which are limited to level-I-perspective taking. They only track *what* an agent can or cannot perceive (see Flavell et al., 1981): When an agent A encounters an object O at a certain location L, she registers O to be at L until she encounters it at another location. Based on this information, the first system can predict correctly: When A has the goal to obtain O she will go to the location where she has registered O (which is L). This suffices to correctly guide children's eye movements in implicit false belief tasks. In contrast, explicit tasks require more sophisticated elaboration. For this reason, they activate the second system. The second system operates on proper propositional attitudes. Accordingly, it can also engage in reasoning about *how* an agent has perceived something and, thereby, whether her she represents reality differently.

This structure reveals certain signature limits of the first system: First, it cannot track how an agent has perceived reality. Hence, it cannot resolve under which description the agent represents something. Second, the first system cannot consider the interaction of different mental states. There is indeed some evidence which supports this distinction between system 1 and system 2. Children below the age of four, older children and adults received an implicit and an explicit aspectual false belief task. According to Apperly and Butterfill's account, the younger children should use system 1 to solve both tasks, because they have not developed the second system, yet. The older children and adults should use system 1 to solve the implicit tasks but system 2 to solve the explicit task. And indeed, younger children failed to consider

under which description the agent represents an object in the implicit task (see also Fizke et al., 2017) and the explicit task. Older children and adults solved the explicit task but failed the implicit task version (Low & Watts, 2013). In another implicit task, older children and adults were impacted in their reaction times if a present agent had different perceptual access in *what* she saw. This suggests that they automatically processed the agent's perspective in addition to their own. If the agent differed in *how* she saw something this did not impact reaction times. Thus, it seems that participants indeed used system 1 to solve this task. And that this system is unable to process and therefore to consider how other agent's perceive reality (Surtees et al., 2012).

Apperly and Butterfill's two-system account aims to resolve why children younger than four fail explicit false belief tasks while even infants succeed in implicit versions. We currently do not know whether this paradox pattern of performance which gives reason to their account does indeed exist. Nevertheless, this approach could also account for an early developing objective and a later developing subjective reasoning capacity. The early developing system 1 that operates on level-I-perspective taking is limited to objective reasoning. Only the later developing system 2 which can engage in a level-II-perspective taking allows for subjective reasoning. Both systems remain in function throughout life and older children and adults still rely on the objective reasoning strategy when they have to reason fast and efficient.

3.2 Conative Mental States

Such rich empirical and differentiated theoretical work on children's subjective conception, as it exists for cognitive states, is still lacking for conative mental states. There is ample research on children's objective reasoning about conative states. However, in contrast to research on cognitive states, emphasis has not been put on identifying when children really take the agent's perspective to ascribe her conative states. This is surprising considering the significant role conative states play in most interactions. If we look at explanatory practices, we can see a certain asymmetry. Broadly speaking, we can explain most rational actions by appealing only to the pro-attitudes an agent holds and can neglect what she knows and believes (Steglich-Petersen & Michael, 2015). Consider the following example. Paul observes that Julia opens the cookie jar. If we ask Paul why Julia does this, he will most likely reply "Because she wants cookies". This seems legit and is most likely how all of us would reason in such a situation. But at closer inspection, this explanation of Julia's action is elliptical. It misses one essential component, the cognitive state on which Julia's action is based: "Because she wants cookies and thinks that the cookies are in that jar." (Rakoczy et al., 2007). However, in most situations

we can take the agent's cognitive states as common ground and rely on reality-congruent beliefs as a default. Paul knows and believes that the cookies are in the cookie jar and so does Julia. As long as there is no obvious reason to assume that there is no common ground, there is no need to even consider cognitive states. It would even seem strange, if Paul would give us this extra information. In contrast, this extra information will not seem strange, if there was reason to assume that Julia holds only partial or false representations of the reality. For example, when she has missed that someone has put the cookies in another box. However in everyday life, we will encounter a shared common ground much more often than such diverging representations.

So, typically, we can make sense of an action referring only to the agent's pro-attitudes (especially her desires). Yet typically, the reverse is not possible. In most cases, we cannot make sense of an action only referring to the agent's cognitive states¹. The belief "She believes that there are cookies in the cookie jar" by itself does not predict or explain an action. We need some information about what this action aims to achieve. Could we just go for common ground here? "Paul wants cookies, so Julia wants cookies"? In most cases this will not work because desires are personal. Paul and Julia will have a common epistemic ground when they had perceptual access² to the same events. But there is no reason to assume that they have the same desires. Even if both of them desired the cookies, Julia would most likely rather desire that she has the cookies and *not* Paul. Especially, if there are only a few. One could think of a suitable elliptical reasoning form that assumes objective desirability as common ground, as proposed by Perner and Roessler's teleological account (2010): "In general, cookies are good". Taking this as common ground, Paul could explain Julia's action in terms of her belief. Yet, while beliefs and knowledge are often shared, goals, desires and intentions diverge much more often. Thus, in most situations conative mental states have the higher explanatory value in that they are sufficient to make sense of an action on their own.

If conative mental states are so important, why does theory of mind research put so much focus on subjective conception of cognitive mental states? One significant reason for that

¹ An exception are cases, where the desire but not the belief is completely clear. For example, playing hide and seek. To explain why she looks behind the tree, I will most likely take her desire (finding people) for granted and refer to her belief (Because she believes that someone hides there).

² Setting aside the possibility of other factors that might have influenced their perception, as for instance color-blindness.

might be the challenge to tap such a capacity for conative states empirically. How to tap a subjective conception of mental cognitive states is quite straightforward. Thus, to address children's meta-representational capacities in general, it makes sense to start off with beliefs as a rather clear test case (see Apperly, 2010). The reason why testing for a subjective conception of conative states is more challenging lies in their logical structure. As cognitive states aim to fit the reality, they can be either true or false. To prove the capacity to represent an agent's subjective representation, the child has to demonstrate that she acknowledges that an agent can hold a belief that is false. In contrast, conative states aim to adapt the world to their propositional content. In consequence, only because they do not fit the world, they are not true or false. This might also be the reason, why the discussion about Premack and Woodruff's paradigm to measure theory of mind via conative states brought forward paradigms for cognitive states. Nevertheless, although conative states do not allow for the same straightforward approach as cognitive states, there are certain constellations that require the ascriber to relativize to the agent's subjective perspective.

In what follows I will give an overview how children's understanding of conative states develops. I will present how it is possible to tap a subjective conception of conative states, and what we already know about such a subjective conception of conative states.

3.2.1 Ontogenetic Development of Understanding of Goals

A first notion of pro-attitudes becomes apparent in infants' capacity to make sense of actions in terms of their goals. They appear to grasp that actions are directed towards a certain outcome, the goal. This notion of goals seems to go beyond an appreciation of mere targets. (If I have the goal to kick the ball, "kicking the ball" is the goal and "ball" is the target; Butterfill, 2020). Already in their first months of life, infants show a capacity to track other agents' goals. In a very elegant design, Woodward (1998) had children observe an agent who could choose between a ball and a teddy. The agent always reached for the teddy. In the test phase, the position of the toys was switched. If children did indeed track the agent's goal they should have been surprised if the agent suddenly reached for the ball. If, however, they only tracked the trajectory of the agent's movement they should have been surprised if the agent suddenly reached for the other location. Children as young as six months showed the goal tracking reaction. Moreover, they only showed this goal tracking reaction for human hands which qualify as an agent but not for a mechanical claw (Cannon & Woodward, 2012; but see Ganglmayer et al., 2019 for a non-replication). Infants can even ascribe a goal to an action that

has not even been successfully completed (e.g., reaching for something but not grasping it; Wellman & Brandone, 2009).

Also, they adapt their interpretation of goal-directed actions to their own experiences. When 3-month-olds had experienced how they can use sticky mittens to grasp for sticky toys they correctly tracked the goal of an agent wearing sticky mittens. In contrast, infants who had not made such experience did not track the agent's goal (Sommerville et al., 2005).

As previously mentioned, infants do not only identify actions as goal directed, but they also seem to understand agency as rational. From early on, they expect agents to perform the most rational action to bring about their goal. In such studies, infants were habituated to an agent who always performed a curvilinear movement to surround an obstacle to reach her goal. Once this action was not rational anymore (because the goal had changed or the obstacle had been removed) they were surprised if the agent still performed the curvilinear movement (Csibra et al., 1999, 2003; Gergely et al., 1995).

So, infants seem to be well capable of tracking an agent's goal. They understand actions as directed towards a certain goal, even if the goal is never achieved, and understand that agents will perform rational actions. Yet, this first notion is far from a subjective conception and full-blown concept of conative states. In what follows, I will look separately at existing work of desires and intentions. I will summarize what we already know from an ontogenetic perspective about children's understanding of desires and intentions in general and about their subjective conception in particular.

3.2.2 Ontogenetic Development of Understanding of Desires

3.2.2.1 *Asymmetry View.*

An understanding of desires appears to develop somewhat later. But it is predominantly assumed that children's concepts of desires still develops fundamentally earlier than their concept of beliefs. This view finds support in a large body of evidence. In their famous study, Wellman and Woolley (1990) found that 2-year-olds can reason competently about desires in terms of desire-dependent actions and emotions. Children correctly predicted that an agent would continue searching until her desire (to find the rabbit) was fulfilled but that she would cease searching once she has found it. Moreover, they expected the agent to be happy when she finds the rabbit but not if she finds something else (the dog). Even if the 2-year-old herself would have preferred to find the dog (for similar results see, Bartsch & Wellman, 1995; Hadwin & Perner, 1991; Rakoczy et al., 2007; Wellman & Banerjee, 1991; Wellman & Bartsch, 1988;

Wellman & Liu, 2004). Children can even appreciate desires they themselves cannot even relate to. Infants as young as 18 months understood that other agents can have preferences which they do not share: When an experimenter had expressed a preference for broccoli, they would give her broccoli although they themselves preferred crackers (Repacholi & Gopnik, 1997; but see Ruffman et al., 2018 for failed replication attempts).

Also on a theoretical level, certain properties of desires provide reasons for an asymmetrical development of desires and beliefs. The first is based on the higher explanatory value of desire (Steglich-Petersen & Michael, 2015). As described earlier, this is reflected in the possibility to explain an agent's action elliptically: referring only to the agent's desire but not to her belief. In consequence, we will most likely refer to desires more frequently than to beliefs in everyday life. Accordingly, desires have not only a primacy with regard to their explanatory power but also their frequency. One can easily imagine how this higher frequency might be reflected in children's acquisition of concepts. Young children might gather more experience in ascribing desires than beliefs and, therefore, might also become proficient in ascribing desires earlier (Rakoczy et al., 2007). Second, the differences in logical structure and the normative implications they entail might cause the asymmetric development. As beliefs have a mind-to-world direction of fit they are evaluated against the standard of truth. In contrast, desires have a world-to-mind direction of fit and are evaluated against mostly future fulfillment. Thus, to ascribe beliefs the ascriber has to immediately evaluate whether the belief is true (as is normally the case) or false. This brings along the requirement to inhibit the default ascription of true beliefs. In contrast, desires are not fulfilled by default. Even if they are fulfilled, this will only happen in the future and the question of fulfillment does not arise while ascribing the desire. In consequence, desires are less demanding with regard to inhibition and executive function (Rakoczy, 2010).

3.2.2.2 *Symmetry View.*

The described work clearly shows that there is an early desire reasoning capacity. Yet, it does not resolve in how far this early desire reasoning capacity reflects a genuine subjective understanding of desires. At first glance, this desire reasoning appears to be subjective in at least some sense. Young children can ascribe desires that diverge from their own. However, substantial skepticism that this requires genuine perspective taking comes from the teleological account by Perner and Roessler (2010, 2012; see also Perner et al., 2018). They assume that just as for cognitive states, children can engage in some form of practical reasoning about desires. But this is restricted to objective teleological reasoning. Thus, just as they can reason

objectively about knowledge they can reason on the basis of objective desirability. Based on the premises “She wants to find the rabbit”, “Finding the rabbit is objectively good”, “She has not found the rabbit yet”, “Searching for the rabbit achieves finding the rabbit” they can conclude the rational action “She will continue to search for the rabbit”. Following the teleological account, this only requires children to interpret the agent’s action as directed towards an objectively desirable goal. This form of reasoning does not require taking the agent’s subjective perspective. Still, it allows children to correctly predict and explain an agent’s action in a wide range of cases. It even allows to predict actions based on person-relative goals: Different agents will pursue different choices (the experimenter chooses broccoli while the child will choose crackers). This only requires children to appreciate that different things are desirable for different agents (broccoli for adults and crackers for children, just like being in the air is good for humans but bad for fish). Perner and Roessler argue that a genuine subjective desire understanding is only given by teleology-in-perspective. What follows, is that subjective understanding of desires and beliefs share the same underlying core competency of subjective reasoning. Thus, both concepts should develop in tandem around the age of four.

3.2.2.3 Testing Asymmetry and Symmetry View Against Each Other

If children’s early desire reasoning does not prove their subjective conception, how can we test for such a subjective conception? What kind of desires do unambiguously require subjective reasoning? Desires are strongly subjective in one substantial way: They can be incompatible. A desire can be incompatible with the content of another desire or general values and norms. One way in which desires can be incompatible is between two agents: A wants to win the race (p) and B wants to win the race (q). If A wins the race this will have the immediate consequence that B cannot win the race anymore. Here person-relative objective desirability alone is not sufficient. p is not only objectively desirable for A and not for B; it prevents B from achieving her own goal q. Accordingly, to reason about both, p and q, we need to relativize to A’s and B’s subjective standpoint: p is good from A’s perspective but bad from B’s perspective. A will, thus, do everything to achieve p and B will do everything to prevent p (and vice versa).

In a similar logic, desires can be mutually incompatible within one agent: the so-called Ulysses conflict. Ulysses’ desires were incompatible in that he wanted to hear the Sirens sing (p), but at the same time he did not want to approach them and die (q), which is the inevitable consequence of hearing their song. Thus again, both desires, p and q, are mutually exclusive but here they are held by the same agent (Choe et al., 2005). Fulfilling p prevents as a direct consequence fulfilling q. Ascribing both desires, p and q, requires the ascriber to relativize to

different subjective standpoints. When Ulysses was thinking about the adventure, he wanted to hear the Sirens (i.e., he desires p). When he thought about it reasonably, he desired not to approach them (q). In the end, Ulysses came up with a clever solution and has his crew tie him up the mast. In most everyday situations however, desires remain mutually exclusive. You cannot tidy up and go to the cinema at the same time or smoke a cigarette and fulfill your desire to quit smoking.

A different way in which desires can be incompatible is in relation to objective norms and values, as is the case for wicked desires. If A wants to hurt B, the outcome will be “B suffers”. On an objective level, this outcome will be bad and objectively not desirable. Yet, from A’s subjective perspective it will be good. Thus, to reason about wicked desires the ascriber has to relativize to A’s subjective standpoint: objectively it is bad, but subjectively, from A’s perspective, it is good.

3.2.2.4 Empirical Evidence

Research on children’s subjective reasoning of desires has mainly focused on desires which are interpersonally incompatible. However, existing evidence on when children come to develop a subjective understanding of desires is inconclusive. In line with asymmetric accounts, some studies find that even young children are already proficient in ascribing incompatible desires (e.g., Fizke et al., 2014; Proft et al., 2021; Rakoczy, 2010; Rakoczy et al., 2007). However, other findings are in line with the teleological account and found this capacity to emerge only later when children can also reason subjectively about beliefs (e.g., Lichtermann, 1991; Moore et al., 1995; Priewasser et al., 2013). This mixed evidence comes with a lack of clarity on a conceptual and methodological level (and is plausibly, to a certain degree, also caused by it). Methodologically, several different approaches have been used to tap children’s subjective understanding of mutually incompatible desires: asking directly for A’s and B’s desires, asking children to ascribe emotions to A and B after p but not q was fulfilled, or engaging the children in competitive games in which their desires conflict with their opponents’ desires. These different measures have yielded different results which suggest different developmental trajectories. Conceptually, it is not completely clear what constellation of desires really counts as mutual exclusivity in the sense that it requires subjective reasoning.

One line of research employed desires which were incompatible in that fulfilling p made fulfilling q impossible. In one such study, 3-year-olds played against a puppet. Both had the task of finishing a jigsaw puzzle. They could either take parts from a blue or a red box. From

which box both could take a part was decided by a card that was drawn from the stack. At the very end, the puppet needed the card to be red, and the child needed the card to be blue. Children failed to state that the puppet wanted a red card while they wanted a blue card (Moore et al., 1995). Likewise, in a different approach, children learned about two agents who held incompatible desires. A and B sat in one boat. A wanted the boat to go right while B wanted the boat to go left. The river then took the boat to the right side, whereby A's desire was fulfilled and B's was not. While children had no difficulties in stating that A would be happy about this outcome, they had trouble stating whether B was happy, too (Lichtermann, 1991). These results indicate that young children fail to ascribe incompatible desires. However, children's difficulties to ascribe such incompatible desires disappeared in modified versions of these tasks (Rakoczy et al., 2007). In these modified versions, children were not told about the incompatible desires, but the agents stated their desires themselves. Moreover, to establish that both agents strongly cared about the outcome, the agents quarreled about the outcome. When 3-year-olds received these simplified task versions, they succeeded for both contexts in reporting the diverging desires and in ascribing the incompatible desire-dependent emotions to both agents. Thus, on a methodological level it is not clear which of these tasks and contexts allow to tap children's genuine understanding of desires.

On a conceptual level, it is not clear whether such tasks really establish a mutual exclusivity between desires. In the tasks described before fulfilling p (the boat goes right) made fulfilling q (the boat goes left) impossible. This does not preclude that children simply reason about the desires in a subsequent manner: First, what is objectively good for someone who states A's desire and then, what is objectively good for someone who states B's desire (Priewasser et al., 2013). To ensure that children really engage in subjective reasoning, the task must be such that children can only succeed if they process both desires simultaneously. To this end, a more recent line of evidence asked children to choose between two different actions: action X that helps to fulfill p but not q, and action Y that helps to fulfill p and hinders B in fulfilling q. For example, in a mayoral election, candidate A and B both want to win. To fulfill her desire p (A wins), A can either win over neutral people (X) or she can try to convince people who before would have voted for B (Y). X only helps A to win but it neither helps nor harms B. In contrast, Y helps A to win and directly hinders B from winning. Thus, when A considers only her own desire to win, X and Y are equally effective. X might even be easier. In contrast, when A also considers that B wants to win, she should engage in Y as it does not only win her votes but also costs B votes.

Recent studies engaged children in competitive games in which they had to choose between actions of the structure X and Y. In one such study, triads of children played a game with the goal to be the first to fill her own strand with beads. Children could either draw beads from the middle (X) or take a bead from another child's strand (Y). Results showed that only children who solved false belief tasks engaged in Y-moves, indicating a symmetric development of subjective desire and belief reasoning (Priewasser et al., 2013). Nevertheless, also in this task children's bad performance might originate from methodological issues. In a task with a similar payoff structure, younger children showed no difficulties in stating the incompatible desires as well as the according emotions while they still failed to engage in Y-moves. This clearly puts the earlier results into question. Possibly, that children did not steal other children's beads might simply reflect that they are reluctant to harm the other players (Proft et al., 2021). So, the results on children's reasoning about incompatible desires remain inconclusive due to the substantial conceptual and methodological disagreement.

In contrast, value-incompatible desires are clearly subjective on a conceptual level. Yet, existing empirical work raises methodological concerns. So far, studies have measured children's understanding of wicked desires via emotion ascriptions. Children learned about an agent who held a wicked desire (e.g., to push someone off the swing). They were then asked to rate the agent's emotion once her desire was fulfilled. Only by the age of four, children correctly ascribed positive emotions to the agent whose desire had been fulfilled. Younger children falsely ascribed negative emotions to the agent (Yuill, 1984; Yuill et al., 1996). At first glance, these results support that a subjective conception of desires develops only later, at the same age when children come to reason subjectively about beliefs. Yet, while it is conceptually clear that wicked desires are strongly subjective, it is unclear how suitable emotion ascription is to capture children's ability to ascribe wicked desires. In general, ascribing emotions according to mental states appears to be more demanding and to develop later than ascribing the state itself (Harris et al., 1989). In the case of wicked desires, it becomes even more complicated: Children have to pair positive emotions on the desirer's side with the negative outcome the "victim" experiences. In comparison, this is not necessary for neutral desires where positive emotions are paired with neutral or even slightly positive outcomes. At closer inspection, "positive emotions" is not even the normatively correct answer. By the age of ten, children did not ascribe purely positive emotions. Instead, they ascribed mixed emotions: "She is happy, because she has fulfilled her desire, but also experiences some form of remorse for doing something bad." (Yuill et al., 1996). So conceptually, value-incompatible

desires clearly require subjective reasoning. Existing evidence finds that children incorrectly ascribe negative emotions to an agent who has fulfilled his wicked desire until age four. However, it is unclear in how far this really reflects an inability to reason about such subjective desires.

In conclusion, previous work has shown that children are capable of some form of desire reasoning. However, it is unclear whether this early notion of desires also allows children to represent desires from the agent's perspective. Evidence on children's capacity to reason about subjective desires is still inconclusive. What is missing is a study that builds on the conceptually unambiguous case of wicked desires and uses a task which is not confounded by extraneous task demands.

3.2.3 Ontogenetic Development of Understanding of Intentions

3.2.3.1 The Complexity of the Mental State Intentions

The development of children's concept of intentions mirrors how complex and multifaceted this state is. In the following section, I will first describe the properties that make the state of intention so complex. I will then turn to what we already know about the ontogenetic development of intention understanding in general and children's subjective conception in particular. As described before, desires and intentions differ in some substantial properties. While desires are pro-attitudes that are tied to states of the world, intentions are pro-attitudes that are tied to actions (Astington, 2001). The conditions of satisfaction of desires are met when the desire's content is brought about. How it is brought about is not relevant for a desire. In contrast, the conditions of satisfaction of intentions are more complex: Intentions commit us to a particular action (Bratman, 1987). This action has to be the action as it is represented in the propositional content of the intention (Astington & Gopnik, 1991; Bratman, 1987). Also, intentions are causally self-referential. An intention is only realized if the intention itself causes the intentional action and the action is performed in the "right way" (Searle, 1983). These strong ties between intention and action provide intentions with a high potential for predicting actions. However, this does not make intentions isomorphic. They stand in a many-to-many relation. The same intention might be realized by different actions (I can greet someone by waving my hand or saying hello) and in the same manner physically similar actions can have different underlying intentions (I can run for exercise or to be on time to catch the train; Baird & Baldwin, 2001; Searle, 1983).

Moreover, intentions are complex as they are multifaceted in themselves. Anscombe (1957) distinguished three different guises of intentions that are fundamental to a full-blown concept of intentions: *intentional action*, in contrast to accidental behavior; *intention in acting*, which refers to the reasons behind an action; and *expression of intention for the future*, which refers to actions that are planned to be performed in the future (as cited in Astington, 2001). In a similar approach, Searle (1983) distinguishes between *prior intentions* and *intention-in-action*. The two guises are related to action in the following way: The intention-in-action presents the action (e.g., move the arm in an upward movement). The prior intention represents the intention-in-action (to perform this action). The intention-in-action is satisfied if the experience of acting intentionally way causes the action in a certain way: It has to be my experience that my arm moves upwards, that causes this action. If I do not experience that I move my arm, I failed to realize my intention-in-action, because something else caused it to go up (Searle, 1983, p. 88). The prior intention is satisfied if it causes the action as represented in the intention-in-action.

Interestingly, Malle and Knobe (1997) found a quite similar differentiation directly reflected in the adult folk concept of intentions. In an empirical approach, they found that adult participants distinguished between intentions and intentionality in their judgements and their descriptions. According to participants' responses, intentions represent an action in terms of the agent's belief and desire (I will turn to this belief-desire structure immediately). This intention refers to actions that are planned for the future. Intentionality refers to this action when it is performed. To perform an action intentionally, the agent must not only hold the according intention but also have the skill to perform this action and the awareness that she is actually performing the action. In this way, this intuitive folk psychological concept can even account for complex cases, as deviant causal chains. A typical example for a deviant casual chain is described by Searle (1983, p. 82): A man forms the prior intention to kill his uncle because he wants to inherit his fortune. In his car, he starts to plan how to commit this murder. He becomes so agitated over his planning that he runs over a pedestrian. Coincidentally, this pedestrian happens to be his uncle. Searle explains this case by causal self-referentially. The man did not kill his uncle intentionally because his action was not caused by his intention. The folk concept of intention would come to the same conclusion on such deviant causal chains by referring to skill and awareness: He did not run over the man by using his skill to drive his car in this way and was not aware that he was running over his uncle (Astington, 2001).

Another complex factor of intentions is their function in practical reasoning. To explain rational action, the ascriber relies on the practical syllogism that refers to some form of conative and cognitive states. According to belief-desire theory of intentional action, the ascriber can explain an action by referring to the premise of the agent's desire and the premise of her belief: If an agent wants Y and believes that action X brings about Y, this agent will perform X¹. As beliefs and desires in combination cause intentional action, the ascriber can directly conclude the agent's action from these states (Davidson, 1963). However, this belief-desire structure has been criticized for oversimplifying intentional action. Beliefs and desires alone do not commit us to an intentional action: That I desire to obtain an apple and believe that if I go in the kitchen and take an apple from the shelf, will bring about that I obtain an apple², provides me with a reason to do so. However, it does not commit me to doing so. Desires can be easily reconsidered. I can simply change my mind and decide to go for a cookie. In contrast, once we have formed a prior intention this commits us to carry out this intention. Bratman (1987) argues that intentions are inert. When we reconsider intentions, this requires us to reconsider not only this intention as part of the former background of desires and beliefs but also current desires and beliefs. Thus, by default, we do not reconsider prior intentions but retain them. Moreover, prior intentions impact later formed intentions to achieve internal coherence. I cannot intend to eat the chocolate bar now and to give the chocolate bar to my neighbor as a birthday present later. It appears oversimplified that intentional action can be directly concluded from desires and beliefs. Instead desires and beliefs have to be weighed against a background of prior intentions and plans. From this we can conclude an intention that commits us to perform a certain action (Bratman, 1987).

So, on the one hand, intentions are very complex and multifaceted states. On the other hand, intentions are of fundamental importance for all interactions and probably "the state we impute most" (Premack & Woodruff, 1978, p. 515). They are "the simplest and most obvious mental state [but at the same time] the most difficult to understand completely" (Astington,

¹ Depending on the situation, the ascriber will have to refer to more premises regarding the agent's beliefs: the agent believes that she is able to perform Y, that she will perform X because she desires Y, or that this basic act (e.g., moving the hand with the switch) will generate the higher act (turning the light switch on).

²The same holds for the belief that I am able to go to the kitchen and take the apple or the belief that my desire to obtain the apple will cause me to go in the kitchen and take the apple.

2001, p. 88). This tension between complexity and importance also seems to be reflected in the ontogenetic development of this complex concept. Children appear to reason with whatever notion of intention they can yet grasp. Already from early on, they reason based on a rather limited but still fairly powerful notion of intentions, which is sufficient in a wide range of situations. However, a sophisticated full-blown concept seems to develop only gradually and not before late preschool age.

3.2.3.2 Ontogenetic Development.

A first step towards reasoning about intentions is goal ascription. Children's remarkable capacities to identify actions as goal-directed (as described earlier) reflect a first appreciation of features and structures which are highly relevant for intentions. However, there is disagreement in how far these early capacities reflect a concept of intentions. Very rich interpretations take them to demonstrate a genuine understanding of intentions. These accounts propose that the human brain comes equipped with a system (or systems) to detect and interpret intentional action (Baron-Cohen, 1995; Premack, 1990). More conservative accounts refrain from interpreting infants' early capacities as ascription of intention. They rather suggest that it shows that children can identify intention-relevant features, which provides the basis to develop a genuine understanding of intentions (Wellman & Phillips, 2001; Woodward et al., 2001). Very skeptic positions reduce children's performance in goal-ascription tasks to reactions to low-level perceptual cues, such as physical and temporal regularities. They grant that such a capacity to detect structures is the basis for a later concept of intentions, but they argue that this capacity is also fundamental to all kinds of other cognitive capacities (Baird & Baldwin, 2001; Povinelli, 2001).

Over the course of infancy, children do not only identify actions as goal-directed but also begin to distinguish intentional action from accidental behavior. Already by the age of nine months, children seem to distinguish in their reactions between intentional and accidental behavior. Infants reacted impatiently when the experimenter intentionally withheld a toy from them because she was unwilling to give the toy to the child. They showed much less impatience when the experimenter accidentally withheld it because she was too clumsy (Behne et al., 2005; and see Marsh et al., 2010 who found a similar capacity already in 6-month-olds). Older infants identified intentional action based on simple linguistic markers. They observed an agent who performed two actions on an object. Both actions produced an outcome, but one action was accompanied by an accidental marker, "Whoops!", and the other by an intentional marker, "There!". Infants in this study preferred to imitate the action that was accompanied by an

intentional marker (Carpenter et al., 1998; Olineck & Poulin-Dubois, 2005). Infants are not only selective in what they imitate, they even modify unsuccessful actions in their imitations to achieve the desired outcome. Infants observed an agent who repeatedly failed to achieve her goal, for example, pull apart a dumbbell. When they later had the chance to play with the dumbbell themselves they did not imitate the unsuccessful action but modified the action to achieve that outcome (Meltzoff, 1995). Importantly, they did not do so if they only observed the failed action once and therefore could not tell whether it was accidental (Meltzoff et al., 1999). Relatedly, infants imitated rationally when an agent used an uncommon action to achieve her goal, for example, when an agent used her head to switch on the light. When she did so because her hands were occupied, children imitated this action by replacing it with the more common action: using their hands. In the other condition, the agent could have used her hands but still used her head. Here, children imitated the uncommon action and used their heads, too (Gergely et al., 2002). By the age of two to three years, children begin to refer to intentions verbally to explain action (Bretherton, 1991). Around that age, they can also identify intentional action in verbal tasks. They can do so for themselves as well as for other agents. For instance, 3-year-olds stated that they themselves or another agent did not mean to make mistakes when they tried to repeat a tongue twister (Shultz et al., 1980).

However, these more advanced but still early capacities have one signature limit: Children fail in their intentionality judgments when they cannot simply match the desired outcome and the achieved outcome. The tasks described so far can be solved via a simple matching strategy (Astington & Gopnik, 1991; Shultz & Wells, 1985): The child compares the desired outcome to the achieved outcome. If they match, the child labels the action as intentional; if they do not match, she labels it as unintentional. Evidence that children are indeed prone to such a strategy comes from the following study by Shultz and Wells (1985): In an electronic target shooting game, the shooter first chose a target she wanted to hit. Yet, the experimenter secretly manipulated the apparatus and controlled whether the desired target was hit. Children of age three to seven mainly based their answer on whether the stated outcome and the achieved outcome matched. While even older children appear to find a matching strategy compelling, they refrain from applying this strategy when it is not applicable. In contrast, children younger than five seem to rely on this strategy irrespective of its applicability. In another target-shooting game, children hit colored target-cans. Some of these cans contained prizes. Before each shot, children had to state which target they wanted to hit. Accordingly, there were two outcomes in this game: the intended action outcome (the targeted can) and the

desired outcome (the prize). When children hit the intended can and realized that it did not contain the prize, children younger than five falsely applied the matching strategy. They stated that they had intended to hit the prize-containing target (Phillips et al., 1998). Children younger than five also appear to falsely apply the matching strategies in the case of deviant causal chains. An agent wanted a new doll and intended to buy it at the store. However, before she could perform the intended action her mother gave her the doll as a present. Children understood that her desire was fulfilled. But only by the age of five, children understood that her intention was not realized, because she did not do what she had planned to do (Schult, 2002).

It seems that as powerful as children's early reasoning about intentions is, it is also substantially limited. Even infants appear to be able to identify goal-directedness and intentionality of action. However, until much later, children appear to base intentionality only on the desired outcome and its relation to reality. Children do not appear to differentiate between desires and intentions (Astington, 1991; Perner, 1991). Only by the age of five, children develop a more sophisticated understanding of intentions that appreciates more complex, intention-specific properties: that intentions commit you to a certain action (Schult, 2002) and that intentions have to cause the intended action themselves (Phillips et al., 1998). Baird and Astington (2005, p. 256) describe this as a "shift from inferences based on the observable to those based on the unobservable". They argue that shift occurs because 5-year-olds, in contrast to younger children, have metarepresentational abilities (see also Astington, 2001; Perner, 1991). By the age of five, children have become able to represent an agent's intention. This allows them to understand how the agent represents the intention from her subjective perspective. But does older children's more sophisticated intention reasoning really depict a subjective conception? An agent states the goal "to obtain a doll" and that the action "go to the store and buy one" will bring about this goal. This is rational on an objective level and based on publicly available information. There is no need to relativize to the agent's subjective standpoint to understand that her intention will not be carried out if she receives the doll as a present. Hence, from these studies alone, we cannot tell whether 5-year-olds are really able to represent an agent's intention from her subjective perspective.

3.2.3.3 Testing for a Subjective Conception of Intentions.

One way in which intentions are clearly subjective is that they are aspectual. Under which description an action is intentional depends on how the agent represents the action: If an agent mis-represents a description or does not represent it at all, the action will not be intentional under this description. Oedipus' marriage is an example of the latter case of partial

representation. In this case, an agent represents some descriptions of her action, but she is ignorant of others. In consequence, her action will only be intentional under the descriptions she represents. Oedipus, for example, does not know that Yocasta is also *his mother*. Thus, from his subjective standpoint the action of *marrying the beautiful woman* is intentional. Yet, his action is not intentional under the description of *marrying his mother*.

Another way in which aspectuality of intentions comes about is on the basis of mis-representations. In the case of partial representation, an object is both X and Y, but you represent only X and you have *no representation* regarding Y. In the case of mis-representation, an object is both X and Y, but you believe that it is X and you *believe that it is not Y*. Consider this rather extreme example. You know you have a mole in your company. But you do not know who it is. You are, however, quite sure it is not your longstanding employee Peter, because you trust him. Unfortunately, this is a mis-representation. *Your longstanding employee Peter is also the mole*. You assign Peter the task to investigate who is giving away secret information. Did you intentionally ask the *mole* to find the mole? Your action is intentional under your mis-represented description *ask the longstanding employee who is not the mole to find the mole*. But, crucially, it is not intentional under the reality-congruent description *ask the mole to find the mole*¹.

To appreciate that actions are only intentional under the descriptions the agent represents, the ascriber has to take the subjective perspective of the agent. This makes the aspectuality a perfect test case for a subjective conception of intentions. First evidence for children's capacity to appreciate aspectuality of intentions based on partial representation comes from a study by Kamawar and Olson (2011): In this study, the agent gives keys to the policeman (X). The policeman is also Cathy's dad (Y), but the agent does not know that (she partially represents only X). Only by the age of eight, children appreciated the aspectuality of intentions. Younger children said that the agent intentionally gave the keys to Cathy's dad (Y). However, a more recent study indicates that children's difficulties in this task reflect a

¹ Aspectuality of intentions can also come about in the form of side effects which are foreseen but themselves not intended. In this case you are aware of the action descriptions, but you do not care about one description or value it less strongly than another description. For instance, a physician administers a treatment to a patient. That treatment will save her life, but it will also make her loose her hair. In some crucial sense, the physician did not intend to make the patient bald. She intended to save her life and accepted that this comes along with the hair loss. While these cases require a subjective notion of intention, they provide less clear-cut test cases than mis-representation and partial representation.

performance limitation rather than a genuine competence deficit. In a simplified and more engaging context, children appreciated the aspectuality of intentions already by the age of five (Proft et al., 2019). In contrast to Kamawar and Olson's task, this task was not an abstract vignette. Instead, children learned about all information themselves in an interactive game context. This decreased the cognitive load, especially in terms of working memory. Children in this study played a game together with a puppet. In the first phase of this game, the child and the puppet chose animals for their farm. Only in the second phase, it turned out that all these animals also had a second identity. They were either also Blickets or Zickets, for example the cow was also a Blicket. To induce further relevance to the second identity, Blickets caused the child and the puppet to lose stickers. As both were ignorant of the second identity in the beginning, choosing the animal was intentional under the description of choosing the cow but not choosing the Blicket. In this simplified and relevant format, children showed an appreciation of the aspectuality of intentions already by the age of five.

Thus, children seem to develop a fairly sophisticated notion of intentions by the age of five. However, it is unclear whether this notion also allows them to represent the agent's intention subjectively. There is initial evidence that children have some subjective conception of intentions by the age of five. Yet, it is unclear how comprehensive this conception is. Children of this age have not only developed a concept of partial representation but also of mis-representations. Thus, they can ascribe mis-representations and appreciate aspectuality of intentions that is based on partial representations. By theory, this should allow children to also consider the aspectuality of intentions based on mis-representations. However, to date, at least to my knowledge, this has not been addressed empirically.

4 Conclusion and Open Questions

Part of a fully-fledged concept of cognitive and conative mental states is a subjective conception. For cognitive mental states, the literature provides extensive insights about the developmental trajectory of such a concept. Already from early on children can reason on an objective level about what an agent knows because of her perceptual access. By the age of four, children develop a subjective conception of cognitive states. At this age, they can take the agent's subjective perspective to represent how she represents or mis-represents reality. Based on this extensive empirical work, much theoretical work discusses how children's concept of cognitive states develops.

In contrast, much less empirical work exists on children's subjective understanding of conative mental states. This is surprising considering the critical role conative states play in everyday life and their high explanatory value. One reason for the asymmetry of empirical evidence on subjective reasoning might stem from the opposite direction of fit of cognitive states. The fact that conative states cannot be false makes it difficult to determine whether the ascriber relativizes to a subjective standpoint. Also, theoretical work is much less differentiated and spelt out for conative states than cognitive states. This might be a consequence of the lack of empirical basis. Nevertheless, it is possible to capture subjective conception of conative mental states. Such empirical designs build on the following properties: Desires are subjective in that they can be incompatible with other desires of norms and values. Intentions are subjective in that they are aspectual.

The scarce existing empirical work that has built on these properties points in the following direction: We cannot simply presume that children's capacities as prevalently measured also taps a subjective conception. Existing evidence on desires is inconclusive. Some research suggests that a subjective conception of desires develops rather late in tandem with a subjective conception of beliefs (for an overview see Steglich-Petersen & Michael, 2015). Evidence on the subjective conception of intentions is scarce. It suggests a subjective conception that develops even later. However, this assumption is based on the results of two studies (Kamawar & Olson, 2011; Proft et al., 2019). This lack of clarity regarding children's subjective conception of conative states leaves a fundamental gap in our understanding of theory of mind and its ontogenetic development. To fill this gap, we need research that systematically tests for a subjective conception of conative states and when it develops in relation to cognitive states.

But how could children's subjective conception of conative and cognitive mental states develop? Theoretically, different trajectories seem plausible. As described before, Perner and Roessler (2010) apply their teleological account not only to beliefs but also to children's subjective conception of desires. From early on, children can reason about an agent's states, but this reasoning is limited to an objective teleological form. Only later do children develop the capacity to take the agent's perspective. This allows them to reason subjectively not only about beliefs but also desires. Apperly and Butterfill's two-system account (2009) has not been elaborated regarding conative states. However, it seems plausible that from early on, children can operate on a system 1 not only to ascribe cognitive but also conative states. This system 1 is, as argued by Apperly and Butterfill, limited to a level-I-perspective taking process of

registration. This could not only allow children to reason about what an agent has access to, it might also allow them to consider an objective form of desirability and to make sense of goal-directed actions. However, as it cannot account for how agents represent reality this notion of conative states should be restricted to such proxies and an objective level. Only the later developing system 2 would then allow them to reason about cognitive and conative mental states as propositional attitudes that represent reality from the agent's subjective perspective.

Building on these accounts, it seems plausible that the following two-system structure might apply for cognitive and conative mental states in general: One early developing system is limited to an objective form of reasoning that functions on some form of proxy-states. A later developing system can reason about proper propositional attitudes and relativize to the agent's subjective perspective. Such a trajectory would be reflected by a pattern of results as follows: First, children come to reason on a subjective level about cognitive and conative states at the same age. Second, their subjective reasoning capacity for beliefs, desires and intentions is strongly related.

Another plausible trajectory builds on the fact that beliefs and desires are essential components of intentions. Moses (2001) argues that for this reason a subjective conception of intentions should develop in a protracted manner after desires and beliefs. In other words, a subjective conception of desires and beliefs should develop first. Only on this basis, should children become able to develop a subjective conception of intentions. Empirically, such a trajectory would be supported if a subjective reasoning capacity for intentions is found only in older children who can reason subjectively about desires and beliefs.

Of course, there are other possible trajectories. One could imagine that conative states and cognitive states develop in more or less complete, modular-like separation from each other. Yet, this appears unlikely considering their strong conceptual relation and how strong performance in ascribing these states is related. Like Baron-Cohen (1995) or Premack and Woodruff (1990), one could also imagine an innate or very early developing full-blown subjective concept for conative states. This full-blown concept of conative states could then be the basis for a further development of a cognitive mental state concept. However, this seems unlikely as children's development of intentions is still undergoing fundamental developmental changes in preschool years and possibly even later.

The present dissertation tests when children develop a subjective conception of desires and intentions. To this end, it builds on the described approaches and addresses the according

limitations. It thereby goes beyond the traditional somewhat limited focus on cognitive states. In this way, this dissertation will build an essential foundation for future more comprehensive empirical and theoretical work on ontogenetic theory of mind-development.

5 Aim of Dissertation

Empirical and theoretical work on theory of mind has neglected the subjective conception of mental states. This dissertation aims to address this fundamental gap in theory of mind research. To this end, this dissertation identified what determines that conative states are subjective. Building on this and existing empirical work, it proposes study designs that explicitly test for children's subjective conception of desires and intention. Based on these study designs, this dissertation systematically tested when children develop a subjective conception of desires (Project 1) and intentions (Project 2a and b). Based on these findings, it aims to propose a comprehensive view on the developmental trajectory of cognitive and conative states.

The rationale of Project 1 was to test when children develop a genuine subjective understanding of desires. Early work has established the prevailing assumption that children by the age of two to three years have developed a nuanced concept of desires (e.g., Bartsch & Wellman, 1995; Steglich-Petersen & Michael, 2015; Wellman & Wolley, 1990). However, more recent empirical and theoretical work sheds substantial doubt on this assumption (e.g., Perner et al., 2018; Perner & Roessler, 2010; Priewasser et al., 2017). This work explains existing evidence by an objective teleological reasoning strategy. A genuine subjective understanding can only be captured via strongly subjective desires. Desires are strongly subjective in that they can be incompatible with other desires or values and norms. Existing evidence on interpersonally incompatible desires is mixed and suffers from a substantial lack of clarity, conceptually and methodologically (e.g., Fizke et al., 2014; Lichtermann, 1991; Moore et al., 1995; Priewasser et al., 2013; Proft et al., 2021; Rakoczy et al., 2007). The test case of value-incompatible desires is clear on a conceptual level. Yet, the evidence so far is restricted to emotion ascription, which is a rather indirect measure and is potentially easily impacted by moral valence (Yuill, 1984; Yuill et al., 1996). Project 1 relies on the conceptually clear case of value-incompatible desires to test for a subjective conception of desires. But in contrast to existing studies, it refrains from emotion ascriptions. Instead, it measures children's understanding of strongly subjective desires via their memory for the propositional content of these desires. Such memory-for-complements tasks have been found to tap children's understanding of beliefs (de Villiers & Pyers, 2002) and desires (Perner et al., 2003).

The aim of Project 2 was to test when children develop a genuine subjective understanding of intentions. Intentions are subjective in one fundamental way: They are aspectual. Aspectuality can come about in different ways. The scarce existing research has focused on aspectuality of intentions based on partial representation (Kamawar & Olson, 2011;

Proft et al., 2019). An agent represents an action under description X but not under Y. Therefore, her action is only intentional under description X. In an engaging, relevant, and simplified format, children appreciated that actions are only intentional under the descriptions the agent subjectively represents by the age of five (Proft et al., 2019). Project 2 investigated how comprehensive this subjective conception of intentions is. To this end, Project 2a tested for children's appreciation of aspectuality of intentions based on mis-representations. Project 2b extends on 2a in the following important way: Project 2a did not conclusively ensure that children understood they had to answer the test questions from the agent's subjective perspective. For this reason, Project 2b adapted the study design. The aim of this adaptation was to emphasize the relevance of the agent's subjective perspective. The task was transferred to a morally relevant context. Children were asked to evaluate the agent's actions that either intentionally or unintentionally brought about negative consequences. The responsibility to punish or reward the agent should motivate children to carefully consider under which description the agent represented her intention.

6 Summary of Empirical Findings

In the following section I will summarize the main findings of the studies of Project 1 and 2 which I conducted in the course of this dissertation: (1) Schünemann, Schidelko, Proft, & Rakoczy (2021) – Children understand subjective (undesirable) desires before they understand subjective (false) beliefs; (2a) Schünemann, Proft, & Rakoczy (2021) – Children's developing understanding of the subjectivity of intentions – a case of “advanced Theory of Mind”; (2b) Schünemann, Bleijlevens, Proft, & Rakoczy (2021) – Children's meta-representational notion of intentions – Understanding the subjectivity of intentions. I will describe the experimental design and the main results. For further details regarding subjects, design, procedure, analysis, and results, please refer to the original manuscripts (Appendices A, B, and C).

6.1 Project 1: Schünemann, Schidelko, Proft, & Rakoczy (2021)

Previous work has provided inconclusive evidence on children's subjective understanding of desires. This study systematically tested whether a subjective conception develops earlier than beliefs or later, in tandem with subjective belief reasoning. To this end, I relied on value-incompatible desires as a conceptually clear test case for subjective desire understanding, and memory-for-complements tasks as a straightforward measure. Moreover, I compared children's desire reasoning directly to their subjective belief reasoning capacity.

Value-incompatible desires, as wicked desires, are strongly subjective because they are good from the desirer's subjective perspective but bad on an objective level. Accordingly, the child has to relativize to the agent's perspective to ascribe a wicked desire. Previous work measured children's understanding of such value-incompatible desires via emotion ascription (Yuill, 1984; Yuill et al., 1996). However, emotion ascription is rather indirect and possibly confounded by the moral significance of wicked desires. To avoid such methodological limitations, I capitalized on memory-for-complements tasks as an indicator of mental state ascription. Originally, these were applied in research on cognitive states. The logic is the following: First, children are told about A's belief about what B is doing in the format "A thinks *that X*" (e.g., A thinks that B is reading a book). This is shown on a picture (A looks at B who is sitting behind a pile of books). Then, children learn that this belief is actually false, "But Y is the case" (But B is playing cards). Again, this is depicted on an according picture. The children are then asked to reproduce the sentential complement of A's initial belief: "What did A think?" – "*that X*". As simple as this task seems, children's success to reproduce the complement appears to depend on children's ability to ascribe false beliefs. Children fail this task until age four and the task correlates with the standard false belief task (de Villiers & Pyers, 2002). This is not only the case for beliefs but also for desires. In contrast to English, the German language also allows to form desires as a that-complementation. For this reason, the memory-for-complements task can also be used for desires in German. In a very elegant adaptation, Perner and colleagues (2003) found that children do not only reproduce belief complements at the same age they solve belief tasks. They also reproduced the complements of neutral (thus compatible) desires around the same age as they succeed in other desire tasks. However, this study focused on linguistic aspects and, therefore, stuck to neutral desires that do not require subjective reasoning. Extending this, memory-for-complements tasks provide the ideal measure to tap and compare children's understanding of value-incompatible and compatible desires. To compare children's performance to their subjective belief reasoning capacity, I also conducted a standard change-of-location false belief task. This allowed me to directly compare children's reasoning about value-incompatible desires to their subjective belief reasoning capacity.

Sixty-one 2.5- to 4-year-olds were tested. They observed a puppet A who stated desires. In all these desires she wanted another puppet B to perform a certain action. These desires were either neutral (e.g., "B should put the necklace on the stand") or wicked (e.g., "B should destroy the drawing"). B always performed some alternative action (put the drawing in a shelf).

Children were then asked to reproduce A's desire complement ("What did A want B to do?; German phrasing: "What did A want that B does with the painting?"). In addition, children received a standard change-of-location false belief task.

In line with the asymmetric account, results showed that young children were better at reasoning about both forms of desires than about beliefs. Results were not in line with the symmetric account. Children did not show more difficulties to reason about subjective desires than about neutral desires. Rather, I found the opposite effect. Young children were even better at reasoning about subjective wicked desires than neutral desires. Moreover, subjective belief reasoning capacity was not related to children's reasoning about subjective desires. These results provide evidence that a subjective conception of desires and beliefs develops asymmetrically: Children first become able to reason about subjective desires before they can reason about false beliefs.

6.2 Project 2a: Schünemann, Proft, & Rakoczy (2021)

This study targeted children's subjective conception of intentions. To this end, it tested when children appreciate the aspectuality of intentions based on mis-representations in three experiments. In Experiment 1, I tested sixty-six 3- to 6-year-olds. Children learned about an agent who acted intentionally under one description X, but because of her mis-representation not under another description Y. The scenarios built on the classic change-of-location false belief test vignette (Wimmer & Perner, 1983) and a more recent variation of this vignette (Rakoczy et al., 2015). The basic structure was as follows: The agent observed that an object A (e.g., a pen) was put in box 1 and an object B (a ball) was put in box 2. The agent left and during her absence the object A was moved to box 2. Consequently, unbeknownst to the agent, box 2 now contained not only B but also A. She returned and intentionally grasped box 2 in order to obtain B. Children were then asked the test question "Did the agent also intentionally take A (the pen)?"¹. In half of the trials, children were asked for the agent's belief instead of her intention. In these belief trials, the agent did not grasp one of the boxes but stated the aim to play with A (which she mis-represented to still be in box 1). Children then had to predict

¹ This is a case of aspectuality of intentions based on mis-representation for the following reason: The agent knows that there is an A and a B. As a result, there are two possible box-descriptions related to A: "contains A" and "does not contain A" (note, this is a different case than not representing A at all). The agent mis-represents box 2: She falsely applies the description "does not contain A". In consequence, her action of grasping box 2 is not intentional under the description "contains A".

where the agent would look for A. I found that none of the included age groups (4- to 6-year-olds) appreciated the aspectuality of the agent's intentions. Interestingly, I found an increase of performance on the intention question with age for children who succeeded on the belief trials but not for children who failed the belief trials. Thus, it seems that understanding mis-representations constitutes a necessary basis to then develop an appreciation for the aspectuality of intentions. Yet, something else has to develop or mature before children come to appreciate the aspectuality of intentions. Thus, a mis-representations seems to be necessary, but not sufficient, for appreciating the according aspectuality of intentions.

To follow up on this, Experiment 2 extended Experiment 1 in three ways. First, I broadened the age window to 3- to 9-year-olds, to be able to determine the age of onset when children come to appreciate the aspectuality of intentions. Second, I further explored the "necessary but not sufficient"-developmental trajectory: I tested what other cognitive and linguistic capacities might be necessary for children to apply their understanding of mis-representations when ascribing intentions. Third, I tested for a relation of my intention task and second-order false belief tasks (A believes that B believes that X). Both tasks follow a similar recursive structure: The ascriber first has to ascribe a mis-representation and then based on that mis-representation she has to draw conclusions concerning the second-order belief or respectively the intention. One hundred and nineteen 3- to 9-year-olds received the same intention and belief tasks as in Experiment 1. In addition, I measured children's working memory, inhibitory control and verbal intelligence, and their understanding of second order false beliefs. Results showed that children appreciated the aspectuality of the agent's intention only by the age of eight. Again, understanding mis-representations was necessary but not sufficient for this capacity. Yet, none of the included cognitive and verbal capacities accounted for this "necessary but not sufficient"-developmental trajectory. Likewise, performance on the intention task was not related to second order belief reasoning.

Thus, based on the first two experiments, results suggest a rather late development of an understanding of aspectuality of intentions. However, this intention task poses high inferential demands in its structure: Children have to first ascribe the mis-representation and then, based on this first reasoning step, ascribe the according intention. Research on other cognitive capacities that require such two-step inferential processes finds an equally late age of onset– but only at first glance. Children fail spontaneous task formats that require them to do all inferential steps at once. However, they succeed in every step on its own. For instance, young children fail to extrapolate from past behavior to future behavior based on personality

traits. This would require the following inferential chain: behavior - personality trait - behavior. Yet, they succeed in each step separately: They can infer a personality trait from behavior as well as a behavior from a trait (Liu et al., 2007; also see Proft & Rakoczy, 2018 for similar findings on moral reasoning). In a similar manner, it is plausible that children conceptually understand the aspectuality of intentions based on mis-representations earlier. However, they fail the spontaneous task format of Experiments 1 and 2 because of its complex inferential demands. For this reason, I conducted a third experiment.

In Experiment 3, children were scaffolded through the reasoning chain. Seventy-nine 4- to 7-year-olds first saw the same scenarios as in Experiments 1 and 2. However, before children had to evaluate the agent's intention, they were reminded of the agent's mis-representation: They were asked where the agent believes A to be. When children answered this question incorrectly, they were corrected. Only, when it was made clear that the agent falsely believed A to be in the initial location children were asked the test question, whether the agent took A intentionally. In this scaffolded format, even 6-year-olds correctly stated that the action was unintentional under the mis-represented description.

Thus, spontaneously children appreciated the aspectuality of intentions based on mis-representations only by the age of eight. In a scaffolded simplified design, they succeeded earlier, at the age of six. However, it is not clear whether children's difficulties really depict a lack of competence. Experiment 3 avoided certain performance factors that masked children's competence. Possibly, other performance factors masked even younger children's competence. One such potential linguistic performance factor results from the form of reading children apply to the intention test questions: Two possible readings can be applied to intention reports, the so-called *de dicto* and the *de re* reading (see Jacob, 2019; Nelson, 2019; Quine, 1956). Our test question asked the child to report the agent's intention. Hence, it establishes an intensional (with an s) context. Intensional contexts require a *de dicto* (about what is said) reading: The intention report is true when the agent herself would also report her intention in that particular way. Accordingly, if we substitute coreferential terms, this can affect the truth-value of the intention report. Oedipus himself would have reported his intentions as "I intend to marry *the beautiful woman*" but not as "I intend to marry *my mother*". Thus, on a *de dicto* level the intention report "Oedipus intends to marry *his mother*" would be false. However, one can also apply a *de re* (about the thing) reading to our intention test question: "The action (in itself) of marrying Yocasta is such that Oedipus intends it." This reading allows to substitute Yocasta by coreferential terms without changing the truth value. We assumed that the more intuitive

reading of our test question is a *de dicto* reading: Would the agent say that she took A intentionally? And indeed, a validation study with adults showed that for them, this is clearly the obvious reading. The fact that this was not the case for children could have two reasons: First, a competence deficit: Young children cannot yet reason on a *de dicto* level about intentions; or second, a performance limitation: Young children are conceptually able to reason on a *de dicto* level. However, they did not see the necessity to do so here. The example of Oedipus and my study have one fundamental difference: Whether or not Oedipus intended to marry his mother is quite relevant. Whether or not the agent intended to take a pen is much less relevant. Possibly, the *de dicto* reading just was not relevant, salient or obvious enough for young children.

In conclusion, this study clearly showed that aspectuality of intentions based on mis-representations is not self-evident for young children. However, the results does not conclusively show the absence of a grasp of the aspectuality of intentions. Children low performance might simply reflect a higher threshold of recognizing when to relativize to an agent's subjective standpoint.

6.3 Project 2b: Schünemann, Bleijlevens, Proft, & Rakoczy (2021)

Project 2b addressed exactly this issue: It tested for children's appreciation of aspectuality of intentions based on mis-representations in a context that clearly required the ascriber to relativize to the agent's subjective standpoint (apply a *de dicto* reading). What determines such a context? Consider the following. If I ask you "Did Oedipus intend to marry his mother?", why does it become so apparent that I ask you to relativize to his perspective? If you were to falsely evaluate this action as intentional, you would blame Oedipus for breaking the norm of incest (a norm violation he would even be prosecuted for). Thus, it is fairly relevant to consider his perspective when you judge his intentions. This relevance makes it very salient that I actually expect you to relativize to Oedipus' standpoint. This is exactly what this study built on. The basic structure was very similar to Project 2a for two exceptions: First, the agent's intentions were embedded in a morally relevant context. And second, children were not asked directly for intentions. Instead, I tapped their understanding of intentions via action evaluation. The logic behind this was the following: If children understand under which description the agent acts intentionally, they should punish an agent who intentionally harms someone. Yet, they should refrain from punishing an agent who does so unintentionally.

Seventy-five 4- to 6-year-olds learned about two agents, A and B. B asked A to give her a desirable object (e.g., a pen). The pen and an undesirable object (e.g., a banana peel) were put in two boxes. A left to get a stool to be able to reach the pen. In the experimental condition, the objects were then switched while A was still absent. In consequence, when A returned, she mis-represented the boxes' contents. She gave B the box which she mis-represented to contain the pen, but which in reality contained the banana peel. Thus, because of her mis-representation she unintentionally gave B the undesirable object. The control condition was similar with one crucial difference: A returned before the objects were switched. Thus, she correctly represented the description under which she acted and gave B the undesirable object intentionally. Children were asked to assign marbles to A to evaluate her behavior: If A was good, give a marble; if A was bad, take one away; if A was neither good nor bad, do nothing. Only 5- and 6-year-olds punished the agent more when she intentionally gave the undesirable object than when she did so unintentionally. In contrast, 4-year-olds did not consider the agent's intentions in their evaluations. Accordingly, it seems that this highly salient context made it easier for children to realize that they have to relativize to the agent's perspective. By the age of five, children already appreciated that the agent did not act intentionally under the mis-represented description. Thus, for the cases of mis-representations we find a similar pattern of results as for partial representations: When extraneous task demands are low and the necessity to apply a *de dicto* reading is highly salient children can appreciate the aspectuality of intentions by the age of five (compare Proft et al., 2019).

In conclusion, Project 2a and b show the aspectuality of intentions is not self-evident for children. First, it poses quite high demands in terms of extraneous performance factors. And second, children only take the extra reasoning step to relativize to the agent's subjective perspective in contexts that are sufficiently relevant. Nevertheless, together with previous work on partial representations, these results suggest that children develop a comprehensive understanding that intentions are aspectual by the age of five. Accordingly, at least by the age of five, children seem to have developed a subjective conception of intentions.

7 General Discussion

The present dissertation aimed to investigate children's subjective understanding of conative states. To this end, this dissertation presented ways in which conative states are subjective and how it is possible to tap a subjective understanding of conative states empirically. These possibilities were then implemented in three studies. In the following, I will discuss this dissertation's empirical findings on desires and intentions with regards to the development of subjective understanding. Further, I will relate my findings to existing empirical and theoretical work. I will then explain how my findings contribute to a comprehensive understanding of the developmental trajectory of theory of mind. I will introduce a speculative alternative interpretation of my findings. Finally, I will give an outlook on future research to extend this dissertation and address open questions.

7.1 Subjective Conception of Desires

7.1.1 Summary of Project 1

It is predominantly assumed that children can reason about desires in a sophisticated manner much earlier than about beliefs (for an overview, see Steglich-Petersen & Michael, 2015). Yet, such empirical and theoretical work did not resolve whether this early reasoning capacity includes an important feature: Desires are subjective mental states. For this reason, a full-blown concept of desires needs to incorporate a subjective conception. Perner and Roessler (2010, 2012; see also Perner et al., 2018) proposed that even more sophisticated forms of early desire reasoning only reflect objective teleological reasoning. A subjective conception of desires would require the same capacity to take subjective perspective as belief ascription and therefore develop in tandem. Empirical work on children's subjective understanding of desires is based on incompatibility. Desires are subjective in that agents can hold desires that are mutually exclusive or incompatible with general norms and values. Previous work remains inconclusive about when children develop such a subjective understanding of desires (e.g., Moore et al., 1995; Priewasser et al., 2013; Proft et al., 2021; Rakoczy et al., 2007; Yuill, 1984; Yuill et al., 1996).

This dissertation's first study (Schünemann, Schidelko, et al., 2021) tested for a subjective conception of desires in a more systematic way than earlier work. It relied on the conceptually clear test-case of value-incompatible desires and used memory-for-complements as a direct measure. I found that children can reason about value-incompatible and therefore strongly subjective desires earlier than about beliefs. Also, young children were even better at

reasoning about value-incompatible desires than about neutral desires. Moreover, subjective desire reasoning was not related to subjective belief reasoning. Based on these findings, the following appears to be the case: Children develop a subjective conception of desires significantly earlier than of beliefs and the two are not related.

7.1.2 Implications

What does this tell us with regard to existing theoretical work? These findings clearly contradict the assumption that early desire reasoning is limited to an early objective teleological form of reasoning. They are in line with previous work, which assumes that a genuine understanding of desires develops asymmetrically before an understanding of beliefs. The current project extends this work by showing that this early capacity also comprises a subjective conception. Yet, do these findings also contribute to the discussion why children might come to understand desires earlier?

One plausible reason that has been suggested lies in the higher explanatory value of desires. In most situations, it suffices to refer to desires to explain rational action. This asymmetry in explanatory value appears to be mirrored in explanatory practice. We use this elliptical explanation form more often than an explanation that refers to beliefs. This asymmetry in practice might then be mirrored in children's development: The more frequently applied desire reasoning reaches a sophisticated level earlier than the less frequently applied belief reasoning (Rakoczy et al., 2007; Steglich-Petersen & Michael, 2015). My findings appear to conflict with this explanation. Consider the following: In how many situations do we really have to take the desirer's subjective perspective? It appears that quite often, we can ascribe desires based on objectively available facts (she runs towards the bus – she wants to catch that bus). Accordingly, if it was only a question of practical training, a subjective desire reasoning capacity should also develop later than an objective one.

Another plausible explanation for the asymmetric desires-before-beliefs development lies in the directions of fit these mental states have. Beliefs aim at representing the world accurately. Desires, in contrast, aim to change the world according to their propositional content (Searle, 1983). For this reason, a belief has to be evaluated against the normative standard of truth. Thus, the ascriber does not only have to ascribe the belief but also she has to coordinate whether the belief is true or not. If the belief does not comply with the default option (being true), the ascriber also has to inhibit this default option. In consequence, ascribing beliefs poses more demands in terms of inhibition and executive functions than desires (Rakoczy et

al., 2007). The world-to-mind direction of fit of desires is unaffected by the necessity of subjective perspective taking. So, in this aspect, this explanation seems to be compatible with this dissertation's finding that subjective and objective desire reasoning develops prior to belief reasoning. Yet, one might argue that ascribing strongly subjective desires also requires the inhibition of a certain default of objective desirability (she desires to hurt him, but one should not desire such thing) or first-person desirability (I would not want to hurt him). And indeed, in a study using objective desires, desire ascription was only weakly related to executive function (Moses et al., 2009; cited in Moses & Tahiroglu, 2010). In contrast, ascribing subjective (inter-personally incompatible) desires was related substantially with executive function (Rakoczy, 2010). As Project 1 did not measure children's executive functions, the findings cannot address this objection directly. However, young children were even better at reasoning about value-incompatible desires than about neutral desires. This makes it unlikely but not impossible that the first had higher demands in terms of executive function than the latter. Yet, more research is needed that systematically relates subjective desire reasoning to executive function.

7.1.3 Limitations

My findings clearly support an early developing subjective conception of desires. The design has the substantial advantage that it relies on value-incompatible desires which are clearly subjective. Furthermore, memory-for-complement tasks provide a more direct and purer measure than emotion ascription. Nevertheless, one might be skeptical that reproducing sentential complements is anything more than simple echolalia. Children in this study might have only parroted what they have heard. Wicked desires were even more salient and therefore made it even easier for children to parrot their complements. However, this appears unlikely for several reasons. First, previous work has found close relations between reproduction of complements and other measures of mental state ascription (e.g., de Villiers et al., 2014; de Villiers & Pyers, 2002; Perner et al., 2003). Second, we also validated this study's adaptation of the memory-for-complements task. In a pre-study, children received a belief version of this adaptation similar to the desire version of the main study. Their reproduction of belief complements was closely related to their performance on a standard false belief task.

Nevertheless, this study adds to an inconclusive empirical situation regarding subjective desire understanding, in which different approaches reveal different results. Future research needs to compare different approaches in a systematically and a priori planned way and identify why different approaches have revealed different developmental patterns.

7.1.4 Conclusion

With regard to children's subjective conception of conative states, Project 1 indicates that a subjective conception of desires develops earlier than an understanding of beliefs and is not related to belief understanding. This supports the asymmetry view that children come to understand desires earlier than beliefs.

7.2 Subjective Conception of Intentions

7.2.1 Summary of Project 2

The subject of intentions is paradox. The mental state is close to the observable action, closer than beliefs or desires. However, at the same time, intentions are multifaceted and complex in their relation to reality, actions, beliefs and desires, and intentions themselves. This paradoxon is mirrored in the ontogenetic development of a concept of intentions. While a very early notion of intention is present in infancy already this notion is far from a full-blown understanding of intentions (Astington, 2001). A fairly sophisticated notion of intentions appears to develop around the age of five (for an overview, see Baird & Astington, 2005). However, it is unclear whether this sophisticated notion of intentions considers that intentions are subjective. One way in which intentions are subjective is that they are aspectual. First evidence indicates that children appreciate the aspectuality in irrelevant contexts only by the age of eight (Kamawar & Olson, 2011). However, in relevant contexts and simplified task designs, children begin to appreciate the aspectuality of intentions around the age of five (Proft et al., 2019). At that age, they also consider more sophisticated properties, such as causal self-referentiality (e.g., Schult, 2002). The scarce evidence on aspectuality of intentions is limited to the case of partial representation (an action is not intentional under a description the agent does *not* represent; Kamawar & Olson, 2011; Proft et al., 2019). This dissertation looked at how comprehensive this subjective conception of intentions is. To this end, Project 2 tested children's understanding of the aspectuality of intentions based on mis-representations (an action is not intentional under a description the agent *mis-represents*).

Project 2b showed that, in a simplified and relevant context, children appreciated the aspectuality of intentions by the age of five (Schünemann et al., 2021). In comparison with the earlier conducted Project 2a, this reveals two factors that can mask this competence: One such factor are high inferential demands. Children in Project 2a faced substantial difficulties when the inferential reasoning chain was not made evident to them. Another factor is the saliency of the subjectivity. Children younger than six failed to appreciate the aspectuality of intentions

when the relevance of the agent's subjective perspective was not sufficiently salient (Schünemann, Proft, et al., 2021). Thus, it seems that children have developed a subjective conception of intentions that represents different forms of how the agent represents an action by age five. Nevertheless, this competence is still easily masked and more likely to be applied in obvious contexts.

7.2.2 Implications

Interestingly, the onset of this capacity at age five concurs with children's appreciation of aspectuality of intentions based on partial representations (Proft et al., 2019) and children's grasp of commitment to action and causal self-referentiality (Baird & Astington, 2005). The first concurrence speaks in favor of a comprehensive understanding of the aspectuality of intentions at age five. Remarkably, children's appreciation of aspectuality also appears to be equally impeded by irrelevance and extraneous task demands. For both, the case of partial representation and mis-representation, children failed irrelevant and demanding tasks until rather late, around age eight (Kamawar & Olson, 2011; Schünemann, Proft, et al., 2021).

It is also remarkable that this age of onset complies with the findings that 5-year-olds have developed a sophisticated notion of intentions. This fits with the assumption that this sophisticated notion is caused by their meta-representational abilities (e.g., Astington, 2001; Baird & Astington, 2005; Perner, 1991). I have argued that considering intention's causal self-referentiality and commitment to action does not necessarily require a subjective standpoint. Instead, it might be the other way round. Grasping these properties of intentions might be necessary to be able to appreciate the aspectuality of intentions. It seems plausible that aspectuality builds on these properties in the following way: The agent intends to act under description X (marry the beautiful woman). This commits her to perform this action. Moreover, it is the intention to perform the action under X that causes her to perform this action. This action also has the description Y (marrying his mother), but there was no intention to perform Y that caused this action or committed to this action.

So, do children have to understand that intentions are causally self-referential and commit to action in order to grasp the subjectivity of intentions in general? Or is this only necessary to appreciate this specific subjective feature aspectuality? Clear evidence for the latter would come from studies that tap children's subjective conception at an earlier age by referring to other subjective features. Strictly speaking, intentions are subjective in the same

way as desires: They can be incompatible between two agents¹. To tap children's subjective understanding of intentions in this way, one could measure children's understanding that agents can engage in similar actions to carry out different intentions. A runs towards the steering wheel, because she intends to steer the boat to the right. B performs the same action but because she intends to steer to the left. The child would be required to understand that A's and B's intentions cause them to carry out their intentions, and that their intentions commit them to certain actions which happen to be the same. Indeed, children fail an even simpler version of this task before they understand causal self-referentially and commitment to action. Until the age of five, children fail to ascribe intentions that are not mutually exclusive but only diverge to two agents performing similar actions: A runs to be home for dinner in time and B runs to exercise (Baird & Moses, 2001).

7.2.3 Conclusion

Project 2b shows that children develop a subjective conception of intentions by the age of five, together with a generally sophisticated notion of intentions. Project 2a indicates a high threshold of relevance for children to apply this subjective conception and difficulties to meet the extraneous demands such contexts often come with.

7.3 Developmental Trajectory

To explain other agents' actions, we ascribe cognitive and conative states. This dissertation seeks to contribute to a comprehensive understanding of how theory of mind develops. To this end, it goes beyond the rather narrow focus in theory of mind research on cognitive states and addresses the subjective conception of conative states. Drawing on existing evidence and this dissertation's findings, the following picture emerges: At a very early stage, infants begin to reason on a very effective but limited notion of pro-attitudes: goal ascription. They understand that agents act towards certain goals (e.g., Csibra et al., 2003; Woodward, 1998). Moreover, from early on infants have a notion of what agents have perceptual access to and can infer to a certain degree what an agent knows (e.g., Liszkowski et al., 2008; Pratt & Bryant, 1990). These first very basic and coarse notions of cognitive states seem to be limited to an objective reasoning level (e.g., Astington, 2001; Baird & Astington, 2005; Burge, 2018; Perner et al.,

¹ According to Bratman (1987), intentions entail compliance to consistency. An agent cannot hold two inconsistent intentions that commit him to action. She has to drop one. Against this constraint, it is conceptually unclear whether one agent can hold Ulysses-conflict-like incompatible intentions.

2018). Their early competency indicates that they can already identify features of actions (such as goal-directedness and information seeking) that are relevant for conative and cognitive states. This makes these early notions a very powerful tool to reason about the world. And it might lay the foundation for the development of more sophisticated notions.

At a later developmental stage, children develop a sophisticated notion of desires (e.g., Wellman & Wolley, 1990). The findings of this dissertation indicate that children's early understanding of desires also comprises a subjective conception of desires (Schünemann, Schidelko, et al., 2021). Thus, even before they can represent what an agent (mis-)represents, they can represent to which state an agent aims the world to change. A subjective understanding of beliefs then develops around the age of four (Wellman et al., 2001). This dissertation shows that a subjective conception of intentions develops only about a year later when children also grasp more sophisticated properties of intentions (Schünemann et al., 2021).

I have proposed two plausible trajectories how the human theory of mind might develop. First, I proposed a two-system approach. Inspired by two-system-accounts, especially by Perner and Roessler's teleological account (2010) and Apperly and Butterfill's two-system-account (2009), I have proposed an early developing objective reasoning capacity. This would function on a proxy form of genuine propositional attitudes, such as a notion of goals, desirability, and knowledge. Only later, a second system would develop that would function on proper propositional attitudes. I have proposed that this second system comes with the ability to reason subjectively about cognitive as well as conative states. While the first system would be restricted to an objective form of reasoning, the second system would allow children to view all mental states from the agent's perspective. The findings of this dissertation clearly contradict such a developmental trajectory. A subjective conception of desires appears to be in place earlier than a subjective conception of beliefs. Moreover, reasoning about subjective beliefs was not related to reasoning about subjective desires. A subjective notion of intentions develops even later. Thus, my findings make it unlikely that there is one system 2 that enables a general subjective understanding of all mental states more or less at once.

Second, I proposed a developmental trajectory in which children develop a subjective conception of mental states separately. Earlier developed subjective conceptions can then provide a basis for later developing subjective conceptions. Beliefs and desires are essential components of intentions. Thus, as long as children cannot reason subjectively about desires and beliefs, they should also be unable to reason subjectively about intentions. Only once they have developed a subjective conception of desires and beliefs, the way would be paved for a

subjective conception of intentions. This proposed trajectory is in line with my findings. A subjective conception of desires develops relatively early and possibly independently of a subjective notion of beliefs. A subjective notion of intentions develops even later than beliefs. Strict belief-desire theory suggests that beliefs and desires entail intentional action (Davidson, 1963). In contrast, other theories assume that there is more to intentional action than beliefs and desires (e.g., Bratman, 1987; Searle, 1983). This dissertation's findings point slightly in the latter direction. There seems to be a developmental gap. Children cannot reason subjectively about intentions at the same age they can reason subjectively about desires and beliefs. In line with Bratman (1987) and Searle (1983), I have described intention-specific features like causal self-referentiality and commitment to action as plausible candidates. However, this gap might also be caused by other factors, as executive function. Future research needs to systematically test what exactly causes this gap.

In sum, based on previous work and this dissertation's findings on subjective reasoning about conative states, the following developmental trajectory seems plausible: Already in infancy, children can engage in some form of reasoning about goals and knowledge. Children then develop a subjective conception of desires and later of beliefs. Based on that, they then develop a subjective conception of intentions.

7.4 Subjectivity as a Signature Characteristic for Full-Blown Concepts

7.4.1 Is Subjectivity a Signature for Full-Blown Concepts?

This dissertation aimed to identify when children develop a subjective notion of conative states. I found that young children were even better at reasoning about strongly subjective desires than about neutral desires (Schünemann, Schidelko, et al., 2021). Also, children appreciated the aspectuality of intentions by the same age they are able to consider more complex properties of intentions (Schünemann et al., 2021). This indicates that once children have developed a sophisticated notion of desires and intentions, this notion also comprises a subjective conception. Thus, when children understand desires as a state that is directed towards a certain content (e.g., she wants to find that rabbit and not just some animal), they also understand that desires are subjective representations. Likewise, when children understand that an intention commits to one certain action and has to be caused by the intention itself, they also appreciate the subjectivity of intentions.

I have argued that these results are compatible with a developmental trajectory in which children develop subjective conceptions of different mental states at different ages. Earlier

developed subjective conceptions might provide a basis for the development of subjective conceptions of other mental states. However, these results are also compatible with another, more speculative proposal for the role of children's subjective understanding of mental states: The ability to relativize to the agent's subjective standpoint might not be a signature of a full-blown or very mature concept of a mental state. Instead, it might be much more basic. This dissertation indicates that when children understand mental states on a sophisticated level, they are also able to relativize to the agent's subjective standpoint. But what if a subjective conception does not develop together with a sophisticated notion but is already present even in rather basic notions of mental states? Possibly children can already take the agent's perspective when they predict actions based on perceptual access and goals, or differentiate accidental from intentional behavior.

This does not suggest that children develop a full-blown meta-representational concept of mental states from early on, and that their competence is only masked by some extraneous factors (as suggested by underestimation accounts; e.g., Baillargeon et al., 2010; Carruthers, 2013; Leslie, 2005). It proposes that it is not the capacity to relativize to an agent's subjective standpoint that is lacking in earlier notions. Instead, what is lacking is a grasp of other features of these mental states, such as causal-self referentiality (Astington, 2001).

The idea that a subjective notion of mental states is present from early on was also proposed by Southgate (2020). Her theory builds on children's early proficiency in mastering conflicting perspectives and on their susceptibility to be impacted by other's perspectives. She proposes that human cognition actually starts off with a bias for others' attention while a self-representation develops only later. She argues that this has substantial advantages on a developmental level. Children's life is characterized by the constant challenge to learn and understand the world they live in. Fundamental support and guidance for this challenge comes from their social environment. This makes attending to others' perspectives and discovering the world through their perspectives a very reasonable and efficient approach.

Also, Phillips and Norby (2019) propose that, at least for cognitive states, children are sensitive to agents' perspectives from early on. It just depends on the nature of the state. They distinguish between factive and non-factive theory of mind: From early on, children can track what an agent represents as separate from their own representations. This allows them to track factive states. Factive states (like seeing or knowing) are directly tied to the reality (as you perceive it). You cannot represent Julia as knowing that there are cookies in the jar if you know that Paul ate all of them. Phillips and Norby argue that it is sufficient that children can keep the

agent's representations separate from their own to show that they relativize to the agent's perspective. This makes factive theory of mind a genuine form of theory of mind. Later, children also develop a non-factive theory of mind. At that stage they can also process that what other agents represent can be inconsistent with their own representation of reality. This allows them to represent non-factive states (like beliefs or guesses).

7.4.2 Approaches to Test This

How could we identify the specific role of children's subjective conception? The first step would be to test whether children's sophisticated notions of mental state reasoning comprise a subjective conception. If this is the case, one can look at more basic notions of mental states and how these appreciate the agent's perspective. Let us first turn to subjective conception of sophisticated notions. Why does research on subjective conceptions of desires and intentions but not beliefs contribute to this question? Ascribing false beliefs, requires the ascriber to relativize to the agent's subjective standpoint (at least most cases). In contrast to false beliefs, we can reason about desires and intentions (at least to some degree) without an obvious need to relativize to the agent's subjective standpoint (if someone says she is hungry, she objectively will desire food; if someone intends to buy the puppet, she objectively has to perform this action to carry out her intention). For this reason, one could imagine for desires and intentions something like the following: Children have an early capacity to reason about desires and intentions, but this does not appreciate that these states are subjective. A subjective conception develops only later. Thus, in contrast to false beliefs, desires and intentions allow us to test two options against each other: Option one, there is a notion of this state that is quite sophisticated, but it lacks subjectivity. Alternatively, option two, subjectivity is so basic that it will be considered by sophisticated notion and possibly in even early forms of state reasoning¹. This dissertation contributes to testing these two options against each other in the following way: For desires, I found that the subjectivity of desires did not pose any additional challenge. Children were even better in reasoning about strongly subjective desires than about neutral desires. For intentions, I found that children considered subjectivity of intentions at the same age as they engage in rather sophisticated reasoning about intentions, about the age of five.

¹ One could probably assume something similar for true beliefs, a sophisticated reasoning about true beliefs which is not subjective. However, children's surprising difficulties to reason about true beliefs are a whole field of research on its own (e.g., Rakoczy & Oktay-Gür, 2020).

Thus, for both conative states, it seems that at least a sophisticated notion of this state already allows the child to relativize to the agent's subjective perspective. This clearly speaks in favor of the second option that subjectivity is rather basic.

However, if a subjective conception is indeed present from early on, we should also find a subjective conception in children's early reasoning about others' mental states. One way to test whether young children relativize to an agent's subjective standpoint when they reason about goals or knowledge could build on aspectuality of mental states. However, because of the young age at which these early capacities are present, such studies would be limited to implicit measures. What could that look like? For goals, one option could be to adapt the paradigm by Woodward (1998). In this paradigm, infants expect an agent to keep grasping for the same goal even if the path of that reaching movement changes. To test if they also understand that the agent only reaches for this goal under a certain description one could use an adaptation of the following logic: An agent always chooses a die over a ball. The child but not the agent learns that the die is also a rattle, and the ball is also a bell. The agent then has to choose between a hidden object that makes a rattling sound and one that rings. If children understand that the agent aims for the die only under the description die but not rattle, they should expect random actions. In a related manner, one could test for children's appreciation of the aspectuality of knowledge. The child and the agent observe that A is put in an empty box. Unbeknownst to the agent, A is also B. Children who appreciate the aspectuality of knowledge should expect the agent to look for a B at another location than box 1.

The aspectuality of mental states does also allow to test for a subjective conception of earlier more basic notions of each mental state. Consider the children's early notion of intentions, that allows them to differentiate between intentional and accidental behavior (Behne et al., 2005). An aspectual version of this task could be implemented as follows: Someone withholds an object from the child under description X, which the child dislikes. Yet, unbeknownst to the agent this object is also Y, which the child likes. Do children consider under which description the agent withholds the object intentionally and under which she does so unintentionally?

Such a test a battery that systematically taps children's subjective conception of different mental states could reveal the actual role of children's subjective understanding in their theory of mind development. If children can indeed reason about an agent's subjective perspective from early on and for different notions this would support that a subjective conception develops early and is not a signature characteristic only of full-blown concepts.

However, if a subjective conception is only part of highly sophisticated notions, this would indicate that a subjective conception is indeed a signature limit that distinguishes a proper meta-representational concept from earlier notions.

7.5 Summary of Outlook

How should this project be continued? I have pointed out two ways in which future work needs to extend this dissertation. First, future research needs to validate the results of this dissertation. In Project 1, this dissertation found that children can reproduce the sentential complements of strongly subjective value-incompatible desires before they can solve standard false belief tasks. Young children are also better at reproducing the complements of value-incompatible desires than those of neutral desires. This indicates an early developing subjective conception of desires. Yet, these results remain inconclusive to a certain degree because they conflict with existing evidence (e.g., Yuill et al., 1996). Future research needs to test systematically why different methods yield different developmental patterns. Project 2 found that children appreciated that an agent acts unintentionally under a description she mis-represents by the age of five. However, children's performance was substantially impeded by inferential demands and in contexts in which the necessity to take the agent's perspective was less salient. Future research needs to investigate if children's genuine competence in Project 2b is still masked by extraneous factors.

Second, future research needs to address the significance of subjectivity. Does children's mental state reasoning at each state already comprise a subjective notion? Or is subjective reasoning tied to a proper understanding of mental states as propositional attitudes? I have proposed the aspectuality of mental states as one approach to test for children's subjective conception of mental states. A very systematic way to implicate this would be to adapt existing measures that tap children's state reasoning capacities in a way that requires children to consider the descriptions under which agents hold these states. Although this sounds like a very straightforward strategy, there are at least three caveats we have to consider. First, to appreciate aspectuality, the ascriber needs to represent that an agent represents some but not all descriptions. In other words, the ascriber has to process, at least, some form of perceptual access. Thus, if early in infancy children fail to appreciate the aspectuality of, for example, goals, this might only reflect insufficient understanding of perceptual access. Thus, there might be a lower limit in terms of age for which aspectuality tasks are appropriate. Possibly though, for such young age groups, subjective conception could be tapped by other approaches, for instance, as in Project 1 via incompatibility.

A second caveat concerns task demands. Project 2a of this dissertation clearly shows how prone aspectuality tasks are to suffer from extraneous task demands. The same can be seen for aspectuality of intentions based on mis-representations (compare Kamawar & Olson, 2011 to Proft et al., 2019) and the aspectuality of beliefs (compare, for example, Apperly & Robinson, 2003 to Rakoczy et al., 2015). Hence, aspectuality tasks have to be designed very carefully. Extraneous task demands have to be reduced as far as possible and the necessity to take the agent's subjective perspective has to be rather salient. Also, regarding the interpretation of children's performance, the impact of these competence-masking factors has to be considered very carefully.

A third caveat is that aspectuality requires children to link two different conceptual descriptions with each other (e.g., rattle and pen). Yet, theoretical and empirical work suggests that children fail to engage in such linking processes even from a first-person perspective. Children in these tasks learned that object A has property X (e.g., the green key opens the lion's cage). They then learn that A is also B (the green key is also the yellow key). Counterintuitively, children fail to acknowledge that X applies not only to A but also to B until the age of four (Perner et al., 2011)¹. If children younger than four are indeed unable to link identities, we should not find them to succeed on any aspectuality task before this age. Accordingly, to address this caveat, a first study should test whether children develop an understanding of aspectuality of desires even before age four. This would be a good indicator that children are not hindered to solve aspectuality tasks because they fail to link two identities.

Bearing all these caveats in mind, testing for children's aspectuality of mental states has the potential to reveal the exact developmental trajectory of children's subjective conception of other people's mind. This would allow to understand the exact role subjectivity plays in theory of mind development.

¹ The authors find that children's ability to link different identities is related to their ability to solve false belief tasks. They assume that children construe mental file cards for each identity. To understand that A=B they have to link the two files. Similarly, children construe copies of their file cards for other agents. To reason about the agent's representation children have to link her files as well. They assume that this form of linking requires the similar capacities as the first-person linking (for more details on this mental files theory, see Huemer et al., 2018).

7.6 Conclusion

The present dissertation is an important first step to investigate children's subjective conception of conative states. Existing work has neglected this substantial property of conative states. This is striking as conative states have a high explanatory value and might be ascribed in everyday life more often than beliefs. To a certain degree, this might be the case because testing for such a subjective conception is less straightforward for conative mental states than for cognitive states. This dissertation has described possibilities that allow us to test for children's subjective conception of conative states. In applying such designs, this dissertation shows that a subjective conception of desires develops much earlier than a subjective conception of intentions. However, in both cases, a subjective conception seems to be present when children develop a sufficiently sophisticated notion to reason competently about these states. This gives first evidence that a subjective conception of mental states does not develop at one time for all states. It rather supports a trajectory in which children develop a subjective conception of desires before they develop a subjective conception of beliefs. Subsequently, this might allow children to develop a subjective conception of intentions. However, my results are also compatible with a view that assumes that a subjective conception of a mental state is not a signature characteristic of an adult-like concept of this state but is present even in earlier notions. In conclusion, this dissertation provides a fundamental basis to achieve a comprehensive understanding of theory of mind as the capacity to reason on a subjective level about other agents' mental states.

8 References

- Anscombe, G. E. M. (1957). *Intention*. Basil Blackwell.
- Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, *116*(4), 953.
<https://doi.org/10.1037/a0016923>
- Apperly, I. A., & Robinson, E. J. (2003). When can children handle referential opacity? Evidence for systematic variation in 5- and 6-year-old children's reasoning about beliefs and belief reports. *Journal of Experimental Child Psychology*, *85*(4), 297–311.
[https://doi.org/10.1016/S0022-0965\(03\)00099-7](https://doi.org/10.1016/S0022-0965(03)00099-7)
- Apperly, I. A. (2010). *Mindreaders: The Cognitive Basis of "Theory of Mind"* (1st ed.). Psychology Press. <https://doi.org/10.4324/9780203833926>
- Astington, J. W. (1991). Intention in the child's theory of mind. In D. Frye & C. Moore (Eds.), *Children's theories of mind: Mental states and social understanding* (pp. 157–172). Psychology Press.
- Astington, J. W. (2001). The paradox of intention: Assessing children's metarepresentational understanding. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (Vol. 4, pp. 85–103). The MIT Press.
- Astington, J. W., & Gopnik, A. (1991). Theoretical explanations of children's understanding of the mind. *British Journal of Developmental Psychology*, *9*(1), 7–31.
<https://doi.org/10.1111/j.2044-835X.1991.tb00859.x>
- 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>
- Baird, J. A., & Astington, J. W. (2005). The development of the intention concept: From the observable world to the unobservable mind. In R. R. Hassin, J. S. Uleman, & J. A. Bargh (Eds.), *The new unconscious*. Oxford University Press.
- Baird, J. A., & Baldwin, D. A. (2001). Making sense of human behavior: Action parsing and intentional inference. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 193–206). MIT Press.
- Baird, J. A., & Moses, L. J. (2001). Do Preschoolers Appreciate That Identical Actions May Be Motivated by Different Intentions? *Journal of Cognition and Development*, *2*(4), 413–448. https://doi.org/10.1207/S15327647JCD0204_4

- Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. MIT press.
- Bartsch, K., & Wellman, H. M. (1995). *Children talk about the mind*. Oxford university press.
- Behne, T., Carpenter, M., Call, J., & Tomasello, M. (2005). Unwilling Versus Unable: Infants' Understanding of Intentional Action. *Developmental Psychology*, *41*(2), 328–337. <https://doi.org/10.1037/0012-1649.41.2.328>
- Bennett, J. (1978). Some remarks about concepts. *Behavioral and Brain Sciences*, *1*(4), 557–560. <https://doi.org/10.1017/S0140525X00076573>
- Bermúdez, J. L. (2020). *Frame It Again*. *Cambridge Books*.
- Brandl, J. (1996). Intentionality. In L. Albertazzi, M. Libardi, & R. Poli (Eds.), *The School of Franz Brentano* (pp. 261–284). Springer.
- Bratman, M. (1987). *Intention, plans, and practical reason* (Vol. 10). Harvard University Press Cambridge, MA.
- Bretherton, I. (1991). Intentional communication and the development of an understanding of mind. In D. Frye & C. Moore (Eds.), *Children's theories of mind: Mental states and social understanding* (pp. 49–75). Psychology Press.
- Burge, T. (2018). Do infants and nonhuman animals attribute mental states? *Psychological Review*, *125*(3), 409–434. <https://doi.org/10.1037/rev0000091>
- 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>
- Butterfill, S. A. (2020). *The developing mind: A philosophical introduction*. Routledge.
- Butterfill, Stephen A., & Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind & Language*, *28*(5), 606–637. <https://doi.org/10.1111/mila.12036>
- Callaghan, T., Rochat, P., Lillard, A., Claux, M. L., Odden, H., Itakura, S., Tapanya, S., & Singh, S. (2005). Synchrony in the onset of mental-state reasoning: Evidence from five cultures. *Psychological Science*, *16*(5), 378–384. <https://doi.org/10.1111/j.0956-7976.2005.01544.x>
- Cannon, E. N., & Woodward, A. L. (2012). Infants generate goal-based action predictions: Infants generate goal-based predictions. *Developmental Science*, *15*(2), 292–298. <https://doi.org/10.1111/j.1467-7687.2011.01127.x>

- Carey, S. (2009). Where our number concepts come from. *The Journal of Philosophy*, *106*(4), 220–254. <https://doi.org/10.5840/jphil2009106418>
- 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](https://doi.org/10.1016/S0163-6383(98)90009-1)
- Carruthers, P. (2013). Mindreading in infancy. *Mind & Language*, *28*(2), 141–172. <https://doi.org/10.1111/mila.12014>
- Charles, D. (2012). Teleological causation. In C. Shields (Ed.), *The Oxford Handbook of Aristotle* (pp. 227–266). OUP USA.
- Choe, K. S., Keil, F. C., & Bloom, P. (2005). Children's understanding of the Ulysses conflict. *Developmental Science*, *8*(5), 387–392. <https://doi.org/10.1111/j.1467-7687.2005.00426.x>
- Christensen, W., & Michael, J. (2016). From two systems to a multi-systems architecture for mindreading. *New Ideas in Psychology*, *40*, 48–64. <https://doi.org/10.1016/j.newideapsych.2015.01.003>
- 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](https://doi.org/10.1016/0885-2014(94)90012-4)
- Csibra, G., Bíró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science*, *27*(1), 111–133. https://doi.org/10.1207/s15516709cog2701_4
- Csibra, G., Gergely, G., Bíró, S., Koos, O., & Brockbank, M. (1999). Goal attribution without agency cues: The perception of 'pure reason' in infancy. *Cognition*, *72*(3), 237–267. [https://doi.org/10.1016/S0010-0277\(99\)00039-6](https://doi.org/10.1016/S0010-0277(99)00039-6)
- Davidson, D. (1963). Actions, Reasons, and Causes. *The Journal of Philosophy*, *60*(23), 685–700. <https://doi.org/10.2307/2023177>
- De Villiers, J. G., & De Villiers, P. A. (2000). Linguistic determinism and the understanding of false beliefs. In K. J. Riggs & P. Mitchell (Eds.), *Children's reasoning and the mind* (Vol. 189, pp. 191–226). Psychology Press.
- 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](https://doi.org/10.1016/S0885-2014(02)00073-4)

- De Villiers, J., Hobbs, K., & Hollebrandse, B. (2014). Recursive Complements and Propositional Attitudes. In T. Roeper & M. Speas (Eds.), *Recursion: Complexity in Cognition* (Vol. 43, pp. 221–242). Springer International Publishing.
https://doi.org/10.1007/978-3-319-05086-7_10
- Dennett, D. C. (1978). Beliefs about beliefs [P&W, SR&B]. *Behavioral and Brain Sciences*, *1*(4), 568–570.
- Doherty, M. J. (2006). The development of mentalistic gaze understanding. *Infant and Child Development*, *15*(2), 179–186. <https://doi.org/10.1002/icd.434>
- 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>
- Fabricius, W. V., Boyer, T. W., Weimer, A. A., & Carroll, K. (2010). True or false: Do 5-year-olds understand belief? *Developmental Psychology*, *46*(6), 1402–1416.
<https://doi.org/10.1037/a0017648>
- Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314. <https://doi.org/10.1016/j.tics.2004.05.002>
- Fizke, E., Barthel, D., Peters, T., & Rakoczy, H. (2014). Executive function plays a role in coordinating different perspectives, particularly when one's own perspective is involved. *Cognition*, *130*(3), 315–334. <https://doi.org/10.1016/j.cognition.2013.11.017>
- Fizke, E., Butterfill, S., van de Loo, L., Reindl, E., & Rakoczy, H. (2017). Are there signature limits in early theory of mind? *Journal of Experimental Child Psychology*, *162*, 209–224. <https://doi.org/10.1016/j.jecp.2017.05.005>
- 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. <https://doi.org/10.1037/0012-1649.17.1.99>
- Flavell, J. H., Green, F. L., Flavell, E. R., Watson, M. W., & Campione, J. C. (1986). Development of knowledge about the appearance-reality distinction. *Monographs of the Society for Research in Child Development*, i–87. <https://doi.org/10.2307/1165866>
- Flavell, J. H., Shipstead, S. G., & Croft, K. (1978). Young children's knowledge about visual perception: Hiding objects from others. *Child Development*, 1208–1211.
<https://doi.org/10.2307/1128761>

- Fodor, J. A. (1992). A theory of the child's theory of mind. *Cognition*, *44*(3), 283–296.
[https://doi.org/10.1016/0010-0277\(92\)90004-2](https://doi.org/10.1016/0010-0277(92)90004-2)
- Ganglmayer, K., Attig, M., Daum, M. M., & Paulus, M. (2019). Infants' perception of goal-directed actions: A multi-lab replication reveals that infants anticipate paths and not goals. *Infant Behavior and Development*, *57*, 101340.
<https://doi.org/10.1016/j.infbeh.2019.101340>
- Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature*, *415*(6873), 755–755. <https://doi.org/10.1038/415755a>
- Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences*, *7*(7), 287–292.
[https://doi.org/10.1016/S1364-6613\(03\)00128-1](https://doi.org/10.1016/S1364-6613(03)00128-1)
- Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*(2), 165–193. [https://doi.org/10.1016/0010-0277\(95\)00661-H](https://doi.org/10.1016/0010-0277(95)00661-H)
- Goldman, A. I. (1970). *Theory of human action*. Princeton University Press.
- Goldman, A. I. (1993). The psychology of folk psychology. *Behavioral and Brain Sciences*, *16*(1), 15–28. <https://doi:10.1017/S0140525X00028648>
- 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*, *26*–37. <https://doi.org/10.2307/1130386>
- Gopnik, A., & Wellman, H. M. (1994). The theory theory. In A. Lawrence & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 257–293).
- Gordon, R. M. (1986). Folk psychology as simulation. *Mind & Language*, *1*(2), 158–171.
<https://doi.org/10.1111/j.1468-0017.1986.tb00324.x>
- 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>
- Hadwin, J., & Perner, J. (1991). Pleased and surprised: Children's cognitive theory of emotion. *British Journal of Developmental Psychology*, *9*(2), 215–234.
<https://doi.org/10.1111/j.2044-835X.1991.tb00872.x>
- Harman, G. (1976). Practical reasoning. *The Review of Metaphysics*, *3*(29), 431–463.

- Harris, P. L., Johnson, C. N., Hutton, D., Andrews, G., & Cooke, T. (1989). Young Children's Theory of Mind and Emotion. *Cognition & Emotion*, 3(4), 379–400. <https://doi.org/10.1080/02699938908412713>
- Heyes, C. (2014). Submentalizing: I Am Not Really Reading Your Mind. *Perspectives on Psychological Science*, 9(2), 131–143. <https://doi.org/10.1177/1745691613518076>
- Hogrefe, G.-J., Wimmer, H., & Perner, J. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child Development*, 567–582. <https://doi.org/10.2307/1130337>
- Huemer, M., Perner, J., & Leahy, B. (2018). Mental files theory of mind: When do children consider agents acquainted with different object identities? *Cognition*, 171, 122–129. <https://doi.org/10.1016/j.cognition.2017.10.011>
- Jacob, P. (2019). Intentionality. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2019). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/spr2017/entries/intentionality/>
- Kamawar, D., & Olson, D. R. (2011). Thinking about representations: The case of opaque contexts. *Journal of Experimental Child Psychology*, 108(4), 734–746. <https://doi.org/10.1016/j.jecp.2010.10.005>
- Kampis, D., Karman, P., Csibra, G., Southgate, V., & Hernik, M. (2020). A two-lab direct replication attempt of Southgate, Senju, & Csibra (2007) [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/gzy26>
- Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, 89(1), 25–41. [https://doi.org/10.1016/S0010-0277\(03\)00064-7](https://doi.org/10.1016/S0010-0277(03)00064-7)
- Knudsen, B., & Liszkowski, U. (2012). 18-month-olds predict specific action mistakes through attribution of false belief, not ignorance, and intervene accordingly. *Infancy*, 17(6), 672–691. <https://doi.org/10.1111/j.1532-7078.2011.00105.x>
- Kulke, L., & Hinrichs, M. A. B. (2021). Implicit Theory of Mind under realistic social circumstances measured with mobile eye-tracking. *Scientific Reports*, 11(1), 1–13. <https://doi.org/10.1038/s41598-020-80614-5>
- Kulke, L., Reiß, M., Krist, H., & Rakoczy, H. (2017). How robust are anticipatory looking measures of Theory of Mind? Replication attempts across the life span. *Cognitive Development*. <https://doi.org/10.1016/j.cogdev.2017.09.001>

- 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., Wübker, M., & Rakoczy, H. (2019). Is implicit Theory of Mind real but hard to detect? Testing adults with different stimulus materials. *Royal Society Open Science, 6*(7), 190068. <https://doi.org/10.1098/rsos.190068>
- Lalonde, C. E., & Chandler, M. J. (2002). Children's understanding of interpretation. *New Ideas in Psychology, 20*(2–3), 163–198. [https://doi.org/10.1016/S0732-118X\(02\)00007-7](https://doi.org/10.1016/S0732-118X(02)00007-7)
- Leahy, B., Rafetseder, E., & Perner, J. (2014). Basic conditional reasoning: How children mimic counterfactual reasoning. *Studia Logica, 102*(4), 793–810. <https://doi.org/10.1007/s11225-013-9510-7>
- Leslie, A. M. (2005). Developmental parallels in understanding minds and bodies. *Trends in Cognitive Sciences, 9*(10), 459–462. <https://doi.org/10.1016/j.tics.2005.08.002>
- Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in 'theory of mind.' *Trends in Cognitive Sciences, 8*(12), 528–533. <https://doi.org/10.1016/j.tics.2004.10.001>
- Lichtermann, L. (1991). Young children's understanding of desires. Third year project report. [Unpublished manuscript].
- Liszkowski, U., Carpenter, M., Striano, T., & Tomasello, M. (2006). 12- and 18-month-olds point to provide information for others. *Journal of Cognition and Development, 7*(2), 173–187. https://doi.org/10.1207/s15327647jcd0702_2
- Liszkowski, U., Carpenter, M., & Tomasello, M. (2008). Twelve-month-olds communicate helpfully and appropriately for knowledgeable and ignorant partners. *Cognition, 108*(3), 732–739. <https://doi.org/10.1016/j.cognition.2008.06.013>
- Liu, D., Gelman, S. A., & Wellman, H. M. (2007). Components of Young Children's Trait Understanding: Behavior-to-Trait Inferences and Trait-to-Behavior Predictions. *Child Development, 78*(5), 1543–1558. <https://doi.org/10.1111/j.1467-8624.2007.01082.x>
- 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>

- 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>
- Luo, Y., & Johnson, S. C. (2009). Recognizing the role of perception in action at 6 months. *Developmental Science*, *12*(1), 142–149. <https://doi.org/10.1111/j.1467-7687.2008.00741.x>
- Malle, B. F., & Knobe, J. (1997). The Folk Concept of Intentionality. *Journal of Experimental Social Psychology*, *33*(2), 101–121. <https://doi.org/10.1006/jesp.1996.1314>
- Marsh, H. L., Stavropoulos, J., Nienhuis, T., & Legerstee, M. (2010). Six- and 9-Month-Old Infants Discriminate Between Goals Despite Similar Action Patterns: DISCRIMINATION OF UNWILLING VERSUS UNABLE PARTNERS. *Infancy*, *15*(1), 94–106. <https://doi.org/10.1111/j.1532-7078.2009.00002.x>
- Meltzoff, A. N. (1995). Understanding the Intentions of Others: Re-Enactment of Intended Acts by 18-Month-Old Children. *Developmental Psychology*, *31*(5), 838–850. <https://doi.org/10.1037/0012-1649.31.5.838>
- Meltzoff, A. N., Gopnik, A., & Repacholi, B. M. (1999). Toddlers' understanding of intentions, desires and emotions: Explorations of the dark ages. In *Developing theories of intention: Social understanding and self-control* (pp. 17–41). Lawrence Erlbaum Associates Publishers.
- Moore, C., Jarrold, C., Russell, J., Lumb, A., Sapp, F., & MacCallum, F. (1995). Conflicting desire and the child's theory of mind. *Cognitive Development*, *10*(4), 467–482. [https://doi.org/10.1016/0885-2014\(95\)90023-3](https://doi.org/10.1016/0885-2014(95)90023-3)
- Moses, L., Carlson, S. M., Claxton, L. J., & Stieglitz, S. (2009). Inhibitory control and children's theories of mind: The sequence and generality problems. [Unpublished manuscript].
- Moses, L. J. (2001). Some Thoughts on Ascribing Complex Intentional Concepts to Young Children. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and Intentionality: Foundations of Social Cognition*. The MIT Press. <https://doi.org/10.7551/mitpress/3838.001.0001>
- Moses, L., & Tahiroglu, D. (2010). Clarifying the Relation between Executive Function and Children's Theories of Mind. *Self and Social Regulation: Social Interaction and the*

- Development of Social Understanding and Executive Functions*, 218–233.
<https://doi.org/10.1093/acprof:oso/9780195327694.003.0009>
- Nelson, M. (2019). Propositional Attitude Reports. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2019). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/spr2019/entries/prop-attitude-reports/>
- Newen, A., & Wolf, J. (2020). The Situational Mental File Account of the False Belief Tasks: A New Solution of the Paradox of False Belief Understanding. *Review of Philosophy and Psychology*, 11(4), 717–744. <https://doi.org/10.1007/s13164-020-00466-w>
- Oktaç-Gür, N., Schulz, A., & Rakoczy, H. (2018). Children exhibit different performance patterns in explicit and implicit theory of mind tasks. *Cognition*, 173, 60–74. <https://doi.org/10.1016/j.cognition.2018.01.001>
- Olineck, K. M., & Poulin-Dubois, D. (2005). Infants' Ability to Distinguish Between Intentional and Accidental Actions and Its Relation to Internal State Language. *Infancy*, 8(1), 91–100. https://doi.org/10.1207/s15327078in0801_6
- 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>
- Paul, S. (2020). *Philosophy of Action: A Contemporary Introduction*. Routledge.
- Perner, J. (1991). *Understanding the representational mind*. The MIT Press.
- Perner, J., Mauer, M. C., & Hildenbrand, M. (2011). Identity: Key to Children's Understanding of Belief. *Science*, 333(6041), 474–477. <https://doi.org/10.1126/science.1201216>
- 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. (1988). Developing semantics for theories of mind: From propositional attitudes to mental representation. In J. W. Astington, P. L. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (pp. 141–172). CUP Archive.
- Perner, J., & Esken, F. (2015). Evolution of human cooperation in *Homo heidelbergensis*: Teleology versus mentalism. *Developmental Review*, 38, 69–88. <https://doi.org/10.1016/j.dr.2015.07.005>

- Perner, J., Huemer, M., & Leahy, B. (2015). Mental files and belief: A cognitive theory of how children represent belief and its intensionality. *Cognition*, *145*, 77–88. <https://doi.org/10.1016/j.cognition.2015.08.006>
- Perner, J., Priewasser, B., & Roessler, J. (2018). The practical other: Teleology and its development. *Interdisciplinary Science Reviews*, *43*(2), 99–114. <https://doi.org/10.1080/03080188.2018.1453246>
- Perner, J., & Roessler, J. (2010). Teleology and causal understanding in childrens' theory of mind. In J. H. Aguilar & A. Buckareff (Eds.), *Causing human action: New perspectives on the causal theory of action* (pp. 199–228). Bradford.
- 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., 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., Wellman, H. M., & Slaughter, V. (2012). The mind behind the message: Advancing theory-of-mind scales for typically developing children, and those with deafness, autism, or Asperger syndrome. *Child Development*, *83*(2), 469–485. <https://doi.org/10.1111/j.1467-8624.2011.01728.x>
- Phillips, J., & Norby, A. (2019). Factive theory of mind. *Mind & Language*, *36*(1), 3–26. <https://doi.org/10.1111/mila.12267>
- Phillips, W., Baron-Cohen, S., & Rutter, M. (1998). Understanding intention in normal development and in autism. *British Journal of Developmental Psychology*, *16*(3), 337–348. <https://doi.org/10.1111/j.2044-835X.1998.tb00756.x>
- 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](https://doi.org/10.1016/0022-0965(89)90066-0)
- 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>

- Povinelli, D. J. (2001). On the possibilities of detecting intentions prior to understanding them. *Intentions and Intentionality: Foundations of Social Cognition*, 83(2), 225–248. <https://doi.org/10.1111/j.1467-8624.2011.01728.x>
- 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. (1990). The infant's theory of self-propelled objects. *Cognition*, 36(1), 1–16. [https://doi.org/10.1016/0010-0277\(90\)90051-K](https://doi.org/10.1016/0010-0277(90)90051-K)
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(4), 515–526. <https://doi.org/10.1017/S0140525X00076512>
- Priewasser, B., Rafetseder, E., Gargitter, C., & Perner, J. (2017). 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>
- Priewasser, B., Roessler, J., & Perner, J. (2013). Competition as rational action: Why young children cannot appreciate competitive games. *Journal of Experimental Child Psychology*, 116(2), 545–559. <https://doi.org/10.1016/j.jecp.2012.10.008>
- Proft, M., Hoss, C., Paredes, K. M., & Rakoczy, H. (2021). Do children understand desires before they understand beliefs? A comparison of 3-year-olds' grasp of incompatible desires, competitive games and false beliefs. *Cognitive Development*, 57, 101009. <https://doi.org/10.1016/j.cogdev.2021.101009>
- Proft, M., & Rakoczy, H. (2018). The ontogeny of intent-based normative judgments. *Developmental Science*, e12728. <https://doi.org/10.1111/desc.12728>
- Proft, M., Schünemann, B., & Rakoczy, H. (2019). Children's understanding of the aspectuality of intentions. *Journal of Experimental Child Psychology*, 181, 17–33. <https://doi.org/10.1016/j.jecp.2018.12.001>
- Quine, W. V. (1956). Quantifiers and propositional attitudes. *The Journal of Philosophy*, 53(5), 177–187.

- Rafetseder, E., O'Brien, C., Leahy, B., & Perner, J. (2021). Extended difficulties with counterfactuals persist in reasoning with false beliefs: Evidence for teleology-in-perspective. *Journal of Experimental Child Psychology*, *204*, 105058. <https://doi.org/10.1016/j.jecp.2020.105058>
- Rafetseder, E., & Perner, J. (2018). Belief and Counterfactuality: A Teleological Theory of Belief Attribution. *Zeitschrift Für Psychologie*, *226*(2), 110–121. <https://doi.org/10.1027/2151-2604/a000327>
- Rakoczy, H. (2010). Executive function and the development of belief-desire psychology: Executive function and belief-desire psychology. *Developmental Science*, *13*(4), 648–661. <https://doi.org/10.1111/j.1467-7687.2009.00922.x>
- Rakoczy, H. (2017a). Theory of mind. *The Cambridge Encyclopedia of Child Development*, 505–512.
- Rakoczy, H. (2017b). In defense of a developmental dogma: Children acquire propositional attitude folk psychology around age 4. *Synthese*, *194*(3), 689–707. <https://doi.org/10.1007/s11229-015-0860-8>
- Rakoczy, H., Bergfeld, D., Schwarz, I., & Fizke, E. (2015). Explicit Theory of Mind Is Even More Unified Than Previously Assumed: Belief Ascription and Understanding Aspectuality Emerge Together in Development. *Child Development*, *86*(2), 486–502. <https://doi.org/10.1111/cdev.12311>
- Rakoczy, H., & Oktay-Gür, N. (2020). Why Do Young Children Look so Smart and Older Children Look so Dumb on True Belief Control Tasks? An Investigation of Pragmatic Performance Factors. *Journal of Cognition and Development*, *21*(2), 213–239. <https://doi.org/10.1080/15248372.2019.1709467>
- Rakoczy, H., Warneken, F., & Tomasello, M. (2007). “This way!”, “No! That way!”—3-year olds know that two people can have mutually incompatible desires. *Cognitive Development*, *22*(1), 47–68. <https://doi.org/10.1016/j.cogdev.2006.08.002>
- Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology*, *33*(1), 12–21. <https://doi.org/10.1037/0012-1649.33.1.12>
- Rowling, J. K. (1997). *Harry Potter and the Philosopher's Stone*. Bloomsbury.

- Ruffman, T., Aitken, J., Wilson, A., Puri, A., & Taumoepeau, M. (2018). A re-examination of the broccoli task: Implications for children's understanding of subjective desire. *Cognitive Development, 46*, 79–85. <https://doi.org/10.1016/j.cogdev.2017.08.001>
- Russell, J. (1987). "Can we say ...?" Children's understanding of intensionality. *Cognition, 25*(3), 289–308. [https://doi.org/10.1016/S0010-0277\(87\)80007-0](https://doi.org/10.1016/S0010-0277(87)80007-0)
- Schueler, G. F. (1991). Pro-Attitudes and Direction of Fit. *Mind, 100*(2), 277–281.
- Schult, C. A. (2002). Children's Understanding of the Distinction between Intentions and Desires. *Child Development, 73*(6), 1727–1747. <https://doi.org/10.1111/1467-8624.t01-1-00502>
- Schünemann, B., Bleijlevens, N., Proft, M., & Rakoczy, H. (2021). *Children's meta-representational notion of intentions – Understanding the subjectivity of intentions* [Unpublished manuscript]. Department of Developmental Psychology. University of Göttingen.
- Schünemann, B., Proft, M., & Rakoczy, H. (2021). *Children's developing understanding of the subjectivity of intentions – a case of "advanced Theory of Mind"* [Manuscript submitted for publication]. Department of Developmental Psychology. University of Göttingen.
- Schünemann, B., Schidelko, L., Proft, M., & Rakoczy, H. (2021). *Children understand subjective (undesirable) desires before they understand subjective (false) beliefs* [Manuscript submitted for publication]. Department of Developmental Psychology. University of Göttingen.
- Schuwerk, T., Kampis, D., Baillargeon, R., Biro, S., Bohn, M., Byers-Heinlein, K., Csibra, G., Dörrenberg, S., Fisher, C., & Franchin, L. (2021). *Action anticipation based on an agent's epistemic state in toddlers and adults*.
- 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>
- Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge university press.
- Shahaeian, A., Peterson, C. C., Slaughter, V., & Wellman, H. M. (2011). Culture and the sequence of steps in theory of mind development. *Developmental Psychology, 47*(5), 1239–1247. <https://doi.org/10.1037/a0023899>

- Shultz, T. R., & Wells, D. (1985). Judging the intentionality of action-outcomes. *Developmental Psychology, 21*(1), 83–89.
- Shultz, T. R., Wells, D., & Sarda, M. (1980). Development of the ability to distinguish intended actions from mistakes, reflexes, and passive movements. *British Journal of Social and Clinical Psychology, 19*(4), 301–310. <https://doi.org/10.1111/j.2044-8260.1980.tb00357.x>
- Sirois, S., & Jackson, I. (2007). Social cognition in infancy: A critical review of research on higher order abilities. *European Journal of Developmental Psychology, 4*(1), 46–64. <https://doi.org/10.1080/17405620601047053>
- Smith, M. (1987). The Humean theory of motivation. *Mind, 96*(381), 36–61.
- Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition, 96*(1), B1–B11. <https://doi.org/10.1016/j.cognition.2004.07.004>
- Southgate, V. (2020). Are infants altercentric? The other and the self in early social cognition. *Psychological Review, 127*(4), 505. <http://dx.doi.org/10.1037/rev0000182>
- Southgate, V., Chevallier, C., & Csibra, G. (2010). Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Developmental Science, 13*(6), 907–912. <https://doi.org/10.1111/j.1467-7687.2009.00946.x>
- 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>
- Sprung, M., Perner, J., & Mitchell, P. (2007). Opacity and Discourse Referents: Object Identity and Object Properties. *Mind & Language, 22*(3), 215–245. <https://doi.org/10.1111/j.1468-0017.2007.00307.x>
- Steglich-Petersen, A., & Michael, J. (2015). Why Desire Reasoning is Developmentally Prior to Belief Reasoning: Why Desire Reasoning is Developmentally Prior to Belief Reasoning. *Mind & Language, 30*(5), 526–549. <https://doi.org/10.1111/mila.12089>
- 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>

- 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>
- Surtees, A. D. R., Butterfill, S. A., & Apperly, I. A. (2012). Direct and indirect measures of Level-2 perspective-taking in children and adults. *British Journal of Developmental Psychology*, *30*(1), 75–86. <https://doi.org/10.1111/j.2044-835X.2011.02063.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>
- Vouloumanos, A., Martin, A., & Onishi, K. H. (2014). Do 6-month-olds understand that speech can communicate? *Developmental Science*, *17*(6), 872–879. <https://doi.org/10.1111/desc.12170>
- Wellman, H., 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., & Banerjee, M. (1991). Mind and emotion: Children's understanding of the emotional consequences of beliefs and desires. *British Journal of Developmental Psychology*, *9*(2), 191–214. <https://doi.org/10.1111/j.2044-835X.1991.tb00871.x>
- Wellman, H. M., & Bartsch, K. (1988). Young children's reasoning about beliefs. *Cognition*, *30*(3), 239–277. [https://doi.org/10.1016/0010-0277\(88\)90021-2](https://doi.org/10.1016/0010-0277(88)90021-2)
- Wellman, H. M., & Brandone, A. C. (2009). Early intention understandings that are common to primates predict children's later theory of mind. *Current Opinion in Neurobiology*, *19*(1), 57–62. <https://doi.org/10.1016/j.conb.2009.02.004>
- 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., & Phillips, A. (Eds.). (2001). Developing Intentional Understandings. In *Intentions and Intentionality: Foundations of Social Cognition*. The MIT Press. <https://doi.org/10.7551/mitpress/3838.001.0001>
- Wellman, H., & Wolley, J. (1990). From simple desires to ordinary beliefs: The early development of everyday psychology. *Cognition*, *35*(3), 245–275.

- Wenzel, L., Dörrenberg, S., Proft, M., Liszkowski, U., & Rakoczy, H. (2020). Actions do not speak louder than words in an interactive false belief task. *Royal Society Open Science*, 7(10), 191998. <https://doi.org/10.1098/rsos.191998>
- Wiesmann, C. G., 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>
- 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](https://doi.org/10.1016/0010-0277(83)90004-5)
- Woodward, A. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, 69(1), 1–34. [https://doi.org/10.1016/S0010-0277\(98\)00058-4](https://doi.org/10.1016/S0010-0277(98)00058-4)
- Woodward, Amanda, Sommerville, J. A., & Guajardo, J. (Eds.). (2001). How Infants Make Sense of Intentional Action. In *Intentions and Intentionality: Foundations of Social Cognition*. The MIT Press. <https://doi.org/10.7551/mitpress/3838.001.0001>
- Yuill, N. (1984). Young children's coordination of motive and outcome in judgements of satisfaction and morality. *British Journal of Developmental Psychology*, 2(1), 73–81. <https://doi.org/10.1111/j.2044-835X.1984.tb00536.x>
- Yuill, N., Perner, J., Pearson, A., Peerbhoy, D., & Ende, J. (1996). Children's changing understanding of wicked desires: From objective to subjective and moral. *British Journal of Developmental Psychology*, 14(4), 457–475. <https://doi.org/10.1111/j.2044-835X.1996.tb00718.x>

Appendix A: Schünemann, Schidelko, Proft, & Rakoczy (2021)

Schünemann, B.*, Schidelko, L.*, Proft, M., & Rakoczy, H. (2021). *Children understand subjective (undesirable) desires before they understand subjective (false) beliefs* [Manuscript submitted for publication]. Department of Developmental Psychology. University of Göttingen. *equal contribution as first authors

This article has been submitted for publication in Journal of Experimental Child Psychology, published by Elsevier.

Children understand subjective (undesirable) desires before they understand subjective (false) beliefs

Word count: 9866

Abstract

Our folk psychology is built around the ascription of beliefs (and related cognitive states) and desires (and related conative states). How and when children develop a concept of these different kinds of propositional attitudes has been the subject of a long-standing debate. Asymmetry accounts assume that children develop a conception of desires earlier than a concept of beliefs. In contrast, the symmetry account assumes that conceptions of both kinds of attitudes are based on the same underlying capacity to ascribe subjective perspectives. Accordingly, a genuine subjective understanding of desires develops in tandem with subjective belief understanding. So far, existing evidence that tested these two accounts remains inconclusive, with inconsistent findings from diverging methods. Therefore, the present study tested between the two accounts in a more systematic way: First, we used a particularly clear test case – value-incompatible (wicked) desires. Such desires are strongly subjective because they are desirable only from the agent’s but not from an objective perspective. Second, we probed children’s ascription of such desires in the most direct and simplified ways. Third, we directly compared children’s desire understanding to their ascription of subjective beliefs. Results revealed that young children were better in reasoning about subjective desires than about beliefs. Desire reasoning was not correlated with subjective belief reasoning and children did not have more difficulties to reason about strongly subjective wicked desires than neutral desires. All in all, these findings are not line with the predictions of the symmetry account but speak in favor of the asymmetry account.

Keywords: Theory of mind, beliefs, desires, memory-for-complements

Children understand subjective (undesirable) desires before they understand subjective (false) beliefs

Rational action is based on practical reasoning. Practical reasoning, from the first-person perspective, proceeds from evaluative (e.g., “More oxygen in the room would be good”) and factive premises (e.g., “opening the window will let in more oxygen”) to intentions (“I will open the window”) and finally to actions (opening the window). From the third-person perspective, explaining the rational action of another agent proceeds, in a reverse inference, by reconstructing the practical reasons underlying her action: Why did she open the window? Because she wanted to let in more oxygen and thought that opening the window was a means to do so. Actions are thus explained by ascribing two kinds of underlying subjective attitudes that jointly rationalize it: conative attitudes like desires (e.g., “...more oxygen”) and cognitive attitudes like beliefs (e.g., “opening the window will bring in more oxygen”). For this reason, our folk psychology (or Theory of Mind) is at its core a “belief-desire psychology” (Fodor, 1975; Wellman, 2011).

From an ontogenetic perspective, one of the central questions of Theory of Mind research is how belief-desire psychology develops. One particular question concerns the acquisition of concepts of desires and concepts of beliefs relative to each other. There are two broad competing views: The asymmetry view assumes that an understanding of subjective desires develops before an understanding of beliefs. In contrast, the symmetry view assumes that the understanding of subjective desires and subjective beliefs develop in close tandem because they are based on the very same underlying cognitive capacity (Perner, Priewasser, & Roessler, 2018; Perner & Roessler, 2010).

So, what is the reasoning behind these bold claims? From the perspective of the asymmetry account, there are two plausible, not necessarily mutually exclusive, reasons why an understanding of subjective desires might develop before an understanding of subjective beliefs. First, the desire-belief asymmetry may be due to asymmetries in explanatory practice: If I were asked to explicitly state a psychological explanation for your opening the window I would most likely just answer: “Because she wanted to let in more oxygen”. Strictly speaking, such forms of rational action explanation are elliptical because they omit the premise regarding the agent’s belief. But since in most cases we can easily take the content of the other agent’s beliefs (opening windows let in oxygen) as factual common ground, this elliptical form is usually perfectly fine. Desires thus have a certain primacy in terms of explanatory power (Steglich-Petersen & Michael, 2015): In most cases, we can make sense of

actions by only ascribing desires to the agent. Accordingly, reference to desires is way more frequent than reference to beliefs. Ontogenetically, this primacy of desire ascription in terms of explanatory power and frequency may lead to a primacy in acquisition such that children come to ascribe desires earlier and more readily relative to beliefs (Rakoczy, Warneken, & Tomasello, 2007).

Second, the desire-belief asymmetry may be due to differences in the logical structure of beliefs versus desires. Desires have what is often called a “world-to-mind direction of fit”: They aim at adjusting the world (lack of oxygen) according to the propositional content of the desire (more oxygen). In contrast, beliefs have a mind-to-world direction of fit. They aim at representing the world accurately, by bringing the representational content of beliefs in line with the world (Searle, 1983). This difference in logical structure comes with different normative implications. Beliefs, in contrast to desires, aim at truth and are thus evaluated against the normative standard of being accurate; and their default is being true (unless we ascribed mostly true beliefs, belief ascription could never get off the ground). Belief attribution thus requires the interpreter to coordinate the ascription of belief contents with the true states of affairs in a way not required for desire ascription, and, in the case of false beliefs, to inhibit the default ascription of true beliefs. This makes belief ascription more demanding in terms of inhibition and executive function more generally (Rakoczy, 2010).

In contrast, the symmetry view builds on the assumption that a genuine understanding of both desires and beliefs requires the very same conceptual capacity to understand subjective attitudes and perspectives (Perner & Roessler, 2012). Accordingly, a proper understanding of desires and beliefs will develop in tandem around the age of four. Earlier in development, children do engage in rational action explanation of sorts. But this early rational action explanation is restricted to a teleological form of reasoning, based on objective facts and values or goals (rather than on subjective beliefs and desires): “Why did she open the window?”, “Because more oxygen was needed in the room and opening the window let in more oxygen”. Young children thus understand that agents act towards certain ends (more oxygen in the room), and that certain means are appropriate to reach these ends (open windows → more oxygen) (Perner & Esken, 2015; Perner et al., 2018; Perner & Roessler, 2010). But nothing in this explanatory scheme requires the interpreter to take the agent’s subjective perspective (for related proposals, see Apperly & Butterfill, 2009; Gergely & Csibra, 2003; Gordon, 1986). Even though it does not involve any ascription of subjective perspectival attitudes, this teleological reasoning schema is quite powerful; it allows children (and arguably, adults in many circumstances) to predict and explain many kinds of actions.

On a theoretical level, both kinds of accounts thus provide reasons for why an understanding of desires should develop before (asymmetry) or in tandem with (symmetry) belief reasoning. But what about the empirical level? At first glance, testing these two accounts against each other appears to be an easy task. Just ask children to ascribe desires and beliefs and see what is achieved first. And indeed, a large body of empirical work has followed this approach. What they have found clearly *prima facie* seems to support asymmetry accounts. Long before children can reason about beliefs, they proficiently handle desires (Wellman & Liu, 2004). They predict that agents will cease to perform an action once their desire is fulfilled (Wellman & Wooley, 1990) and ascribe positive emotions to agents who have fulfilled their desires and negative emotions to agents who have not (Hadwin & Perner, 1991; Wellman & Banerjee, 1991). Yet on closer inspection, these results turn out to be perfectly compatible with the symmetric teleological account as well. The tasks do not require a subjective genuine understanding of desires but can be solved by the restricted teleological form of reasoning. Based on objective facts and goals, one can predict that someone will perform the action which is the mean to achieve a certain objectively desirable end and cease that action once this end is achieved. Likewise, objective reasoning suffices to make the connections “achieve desirable end \rightarrow positive emotions” and “not achieve desirable end \rightarrow negative emotions”. Teleological reasoning even allows a little teleologist to predict and understand actions based on person-relative goals such as the following one from a famous study (Repacholi & Gopnik, 1997): Children can predict that, faced with a choice of broccoli and crackers, different agents may make different choices (an adult may make the – from the child’s perspective- absurd choice of broccoli, whereas another child will choose the cracker). In order to do this, one need not revert to subjective desires; rather, an understanding that different things are desirable for different agents is sufficient: broccoli is good for adults whereas crackers are good for children (compare: rotten meat is good for hyenas, but not for humans).

How then can we empirically decide between asymmetry and symmetry accounts more stringently? Crucially, the accounts make competing predictions with regard to understanding desires that stringently require subjective reasoning. According to the symmetry account, these strongly subjective desires require the same subjective reasoning capacities as mis-representations of reality (false beliefs). Teleological reasoning cannot make sense of either of them. Accordingly, the teleological account predicts that an understanding of such subjective desires should develop in tandem with an understanding for false beliefs. In

contrast, the asymmetry account predicts that even strongly subjective desires should be understood and ascribed before beliefs.

Strongly subjective desires in this sense involve some incompatibility between the subjective content of the desire and something else. For example (this will not be focus here), several desires may be mutually incompatible within one agent (e.g., losing weight vs. eating this cookie; (Choe, Keil, & Bloom, 2005).

Another way in which desires can be incompatible is when desires of two agents are mutually exclusive. For instance, A wants to win the race (p) but B also wants to win the race (q). p and q are mutually incompatible. The moment one agent wins the race the other cannot win the race anymore. This cannot be framed in terms of person-relative desirability. Both A and B have a certain attitude towards p. These attitudes are in direct conflict. Thus, to reason about A and B's desires and predict their actions and reactions, one needs to relativize to A's and B's subjective standpoints: p is good from A's point of view but bad from B's point of view.

The same holds for desires that are incompatible with objective values and norms such as wicked desires. Suppose A wants to hurt B. Then the outcome "B is suffering" will be objectively bad, but good from A's wicked perspective. Like in the case of mutually incompatible desires, an interpreter can only make sense of such a situation by reverting and relativizing to A's subjective perspective: objectively bad, but subjectively good from A's standpoint.

Existing research investigating the ascription of strongly subjective desires in order to contrast asymmetry and symmetry accounts has mainly focused on interpersonally incompatible desires. Yet, the evidence from this line of research so far is mixed and hard to interpret. Some studies found that children can reason about such incompatible desires earlier than they can reason about beliefs (e.g., Fizke, Barthel, Peters, & Rakoczy, 2014; Proft, Hoss, Paredes, & Rakoczy, 2021; Rakoczy, 2010; Rakoczy et al., 2007). Others found that children develop this ability only later, when they can also reason about beliefs (e.g., Lichtermann, 1991; Moore et al., 1995; Priewasser, Roessler, & Perner, 2013). These inconsistencies in findings might stem from substantial underlying conceptual and methodological disagreements. Conceptually, there is disagreement regarding the question which types of situation require a genuine understanding of subjective perspectives. Some studies confront children with desires that are incompatible in the sense that if A's desire is fulfilled B's cannot be fulfilled anymore (e.g., Lichtermann, 1991; Rakoczy et al., 2007). Yet, it has been

argued that strictly, only those cases require subjective reasoning where fulfilling A's desire also entails obvious negative consequences for B (like in competitive game contexts; Perner & Roessler, 2010; Priewasser et al., 2013).

On a methodological level, it is not clear what types of tasks are appropriate implementations of such situations. One such implementation is to ask whether children ascribe the respective emotions to A and B. Initial studies showed poor performance before age four (Lichtermann, 1991), but once the questions were simplified and suitably introduced in subsequent research, 3-year-olds performed competently in most cases (Rakoczy et al., 2007). Other studies implement incompatible desires by engaging children in competitive games with other agents. Again, initial studies found young children to struggle to ascribe their opponent's incompatible desire (Moore et al., 1995) and to act in a way that would fulfill their own desire while hindering the opponent in fulfilling hers (Priewasser et al., 2013). But also, in this implementation, children had no difficulties in reasoning about their opponent's conflicting desires if the inferential complexity was reduced or children were not required to harm their opponent, but only had to report the incompatible desires (Proft et al., 2021; Rakoczy et al., 2007).

In contrast to the question what really counts as an interpersonally conflicting desire, more conceptual clarity and agreement hold in the case of desires that conflict with values and norms. In existing studies on children's understanding of wicked desires, subjects were asked to rate an agent's emotion after her wicked desire (e.g., pushing someone off the swing) was fulfilled. The typical developmental pattern found was that until the age of four, children ascribe negative emotions to this agent although her desire was fulfilled; from age four on, they then ascribe positive emotions ("she is happy because she succeeded in pushing the other child off"¹; Yuill, 1984; Yuill, Perner, Pearson, Peerbhoy, & Ende, 1996). On a conceptual level, this fits nicely with the prediction of the symmetry account: "Pushing someone off the swing for no reason" is objectively not desirable and is incompatible with (in this case, moral)

¹ Interestingly (but not in our focus here), the developmental pattern actually constitutes a U-shaped curve. Children younger than four ascribed negative emotions to the harming agent. Children older than four ascribed positive emotions to her. Yet, 10-year-olds again ascribed negative emotions to her. Based on children's explanations of these emotions, the authors interpret this as development from first objective to subjective reasoning and then later to enlightened moral reasoning ("although she may feel Schadenfreude, she cannot be really happy"; Yuill et al., 1996).

norms and values. Thus, ascribing a desire with such a content requires relativization to the agent's subjective standpoint (as such it is bad, but from her perspective it is good).

So, *prima facie*, the evidence from children's development of understanding wicked desires speaks for the symmetry account. However, taken by themselves, these findings are difficult to interpret for at least two reasons: First, ascribing attitude-dependent emotions adds an extra layer of complexity. Inferentially, it goes a crucial step beyond the mere ascription of the state in questions (here, desires): Ascribe the attitude in question, relate it to reality and infer the emotion on the basis of the (non-)fulfillment of the attitude (desire fulfilled → happy; unfulfilled → sad / belief fulfilled → not surprised; non fulfilled → surprised). Correspondingly, the ability to ascribe mental state-dependent emotions (happiness/surprise) has been found to develop with a delay relative to the ability to ascribe the mental state itself (desire/belief) (Hadwin & Perner, 1991; Harris, Johnson, Hutton, Andrews, & Cooke, 1989). Taking the ascription of an attitude-dependent emotion as an indicator of understanding the attitude in question may thus underestimate children's competence.

Second, the evidence so far does not show directly that ascribing desire-dependent emotions in the case of wicked desires is based on the same type of general subjective reasoning that also underlies subjective belief ascription. Existing studies found similar ages of onset in the two types of capacities (ascribing emotions based on wicked desires; and ascribing false beliefs). Yet, similar ages of onset are not sufficient to show shared underlying competencies¹. What is missing is more direct evidence that the two types of competencies, understanding wicked desires and subjective reasoning about beliefs, go together (i.e., correlate).

Against this background, the rationale of the present study was to test between the symmetry and the asymmetry accounts in systematic and novel ways. To this end, we used value-incompatible desires as a test case and capitalized on the memory-for-complements task as an indicator of mental state ascription, and thus avoids problems of previous studies that

¹An interesting case from Theory of Mind research is children's performance in the so-called False Photo tasks relative to their performance in standard false belief tasks. Children come to master both types of tasks around the same age (Zaitchik-Samet, 1989), and so it was originally thought that they tap a shared underlying capacity. But subsequent studies showed that the two tasks do not correlate in neurotypical development and actually dissociate in autism (where children have much less difficulty in the False Photo compared to the false belief task), and thus not tap the same capacity (Leekam, Perner, Healey, & Sewell, 2008).

have used emotion ascription. To our knowledge, the present study is the first to utilize this direct measure for value-incompatible desire and to draw direct comparisons to subjective belief reasoning capacities. Conceptually, these cases are clear: Value-incompatible desires are strongly subjective and cannot be grasped on a restricted teleological level.

Methodologically, we sought to implement more stringent investigations of children's understanding of such cases than previous studies and directly compared children's understanding of wicked desires to neutral desires and their ability to ascribe false beliefs. The logic here is the following: The symmetry account predicts that children who cannot ascribe false beliefs yet should also fail to ascribe subjective wicked desires. But they should have no difficulty ascribing neutral desires (which fall within the scope of simple objective teleological reasoning). In contrast, the asymmetric account predicts that children should be able to ascribe all desires, neutral as well as wicked ones, before they can ascribe false beliefs.

Against the background of the methodological limitations of existing studies, we employed a direct measure of children's desire understanding (rather than an indirect one based on desire-dependent emotion ascription) and compared it to children's false belief understanding directly (rather than relying on indirect comparisons of age of onset).

As a direct measure of children's understanding of desires, we capitalized on so-called memory-for-complements tasks. These were originally applied in the research of children's understanding of cognitive attitudes like beliefs. The logic of this task is as follows: Children see a picture of two agents (e.g., Protagonist A sees B doing something behind a pile of books). They are told about A's belief. In the format "She believes *that X*" (She believes that A reads a book). Then, children are told that this belief is false "But Y is the case" (But B is playing cards) and see a picture of the reality. The test question then asks children to reproduce the sentential complement (*that X*) of the agent's initial belief ("What did she believe?"). Children fail this seemingly easy task until age four, and performance in this task strongly correlates with traditional belief ascription tasks such as the false belief task (de Villiers & Pyers, 2002). This makes memory-for-complements tasks a convenient and direct measure for tapping children's understanding of the concept of those mental states that can be expressed via *that*-complements. In contrast to English, in German this is also the case for desires. Perner and colleagues (2003) already made use of this in a very elegant design to directly compare children's understanding of desires and beliefs. They found that children can reproduce *that*-complements of neutral desires around the same age as they can ascribe these in other tasks, and thus before they reproduce belief-*that* complements and ascribe beliefs more generally.

Yet, as this study was mainly interested in linguistic components, the ascription of desires in this study was restricted to neutral desires and thus, did not require subjective reasoning. In the present study, we build on this direct measure of desire understanding and extend it to study children's understanding of value-incompatible desires.

We implemented these approaches in the following way: We directly compared children's understanding of value-incompatible wicked desires to neutral desires that can be made sense of with an objective teleological notion of desirability, and related both to children's ascription of subjective beliefs in a standard FB task (Wimmer & Perner, 1983). In the memory-for-complements tasks, children observed a puppet who expressed desires regarding what a monkey should do. These could be either wicked (e.g., Monkey should destroy someone else's painting) or neutral (e.g., Monkey should hang the necklace on the stand). The monkey never fulfilled these desires but performed alternative actions (put the drawing in the shelf, put the necklace in a box). Children were then asked what the puppet wanted the monkey to do (in the German "want + that" complementation construction). These scenarios followed the logic of Perner and colleagues' second study (2003), where they used acted-out versions that made the scenarios more vivid, concrete and easier to follow¹. The asymmetry account would be supported by the following pattern: Children can reproduce neutral and wicked desires even before they solve the standard false belief task. In contrast, the symmetry account would be supported by this pattern: Children can reproduce wicked desires only when they also solve the false belief task. Neutral desires can be reproduced even earlier.

Method

Participants

Sixty-one 2½ - to 4-year-old (31-58 months, *mean age* = 41.8 months; 30 girls) monolingual German children were included in the final sample. Three additional children

¹ To validate our adaptation of the memory-for-complements task, we conducted a pilot validation-study. In this study, children received memory-for-complements tasks that were similar in design and structure to our main desire task. But instead of desire complements, children reproduced belief complements, and – as an external validation – we directly compared performance in this task to performance in a standard false belief task. Results revealed that the two types of measures were strongly correlated. The memory-for-belief-complements task thus taps into the same capacity as the standard false belief task. More generally, just like original memory-for-complements tasks, our adaption appears to be a valid measure of children's understanding of mental states (see Appendix A for a detailed description of method and results).

were tested but excluded from data analyses because they were uncooperative ($n = 2$) or did not fulfill the language requirements ($n = 1$). This age range was similar to the original studies by Perner and colleagues (2003, *Exp. 1*: 3,6 - 4,8 years, *Exp. 2*: 2,5 - 4,5 years) and allowed the inclusion of children who have and have not developed an explicit false belief understanding, yet. Participants in this and all subsequent studies were recruited from a databank of children whose parents had previously given consent to experimental participation. The testing was conducted in single sessions by two female experimenters (E) in the laboratory.

Design

In the desire task, each child was tested in two conditions: complementation of neutral desires and complementation of wicked desires. The conditions consisted of two trials each (two neutral desires and two wicked desires), resulting in four total trials for the desire task. The order of the trials was counterbalanced (24 orders of presentation). Before the four experimental trials of the desire task started, children received a two-step warm-up phase. Additionally, children received two trials of a standard false belief task. One half of the children received the false belief task before the desire task and the other half after the desire task. The whole test session lasted ~20-25 minutes.

Materials and Procedure

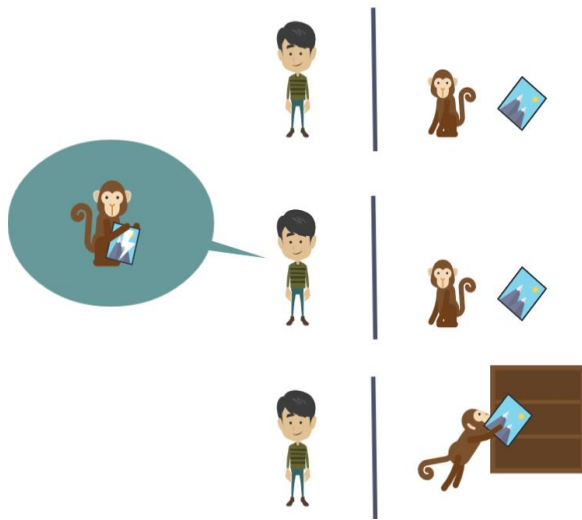
Desire Task. The overarching theme of the desire task was a hand puppet theater with the boy called Tom, Monkey and Frog, acted out by E2. Tom was introduced to the children in the beginning of the test session. Monkey and Frog joined the scene in the beginning of the experimental trials.

Warm-up phase. In the first step of the warm-up phase, children were acquainted to commenting on Tom's actions (e.g., putting a stone out of his trouser pockets). In the second step, children had to verbalize actions displayed by a woman in a short video (e.g., putting a toy drum in a box). This ensured that all participants were able to verbalize the relevant actions for the desire task (for details, see Appendix B).

Experimental trials. For the experimental trials, Tom and Monkey were placed on two sides of an occluder so that they could not see each other. The child sat in front of the scene so that she was able to see both Tom and Monkey. E1 made sure that the child understood that Tom and Monkey were not able to see each other because of the occluder. Frog joined the scene and gave E1 and the child a box containing some of his beloved, self-made belongings.

Frog then had to leave again and left the box with E1 and the child. Afterwards, the four experimental trials of the following scheme began:

E1 took one of the objects out of the Frog's box and in front of Tom and Monkey on the table (e.g., a painting or a self-made necklace). Tom then stated his desire to the child. The desire – neutral or wicked – was always about Monkey conducting an action with the object (e.g., wicked desire: “Monkey should tear the painting” or neutral desire: “Monkey should hang up the necklace on the stand”). Tom did not directly state what he wanted Monkey to do but only what he should do to make the task as parallel as possible to the desire condition of the complementation task by Perner and colleagues (2003). Consequently, children needed to infer the desire from Tom's utterance. All desires shared by the boy were spoken in a friendly voice. There was no difference in pitch and expression between the neutral and wicked desire trials. The objects necessary to conduct the desired action (e.g., the stand to hang the necklace on) were available in the scenery of each trial. However, Monkey always conducted an alternative neutral action with the object (e.g., Monkey put the painting in the shelf). As Tom was on the other side of the occluder, he could neither see Monkey performing the action nor see the outcome of the action. After Monkey performed the action, E1 asked the test question about Tom's desire (“What did Tom want Monkey to do with the painting?”) and the control question about Monkey's actual action (“What did Monkey actually do with the painting?”; see Figure 1 for example procedure).



1. Tom and Monkey are separated by an occluder. Child can see both Tom and Monkey. E1 puts the Frog's self-made painting on the table. Child, Tom and Monkey can see it.
2. Tom to the child: "The monkey should tear the painting."
3. Monkey puts the painting carefully on a shelf. Child, but not Tom can see Monkey's action.
4. Test question: "What did Tom want Monkey to do with the painting?"¹
5. Control question "What did Monkey actually do with the painting?"

Figure 1. Example procedure. Memory-for-complements trial with wicked desire.

The direct translation of the test question in German language includes a that-complementation ("What wanted Tom that Monkey does with the painting?"). If the child did not answer, E1 gave the child the first part of the sentence to answer ("Tom wanted that Monkey..."). To make the test and control question seem less academic, E1 turned her back on the scene while Tom uttered his desire and Monkey performed the action. After the control question, E1 took another object out of Frog's box and the next trial began.

FB Task

The children received two trials of the standard change-of-location task with different stimuli (Wimmer & Perner, 1983) acted out by E1 with little plastic figures: Protagonist A, for example, the boy and his object, for example, his ball were presented to the child. Before leaving the scene, the boy placed his ball in one of two boxes (box 1). In his absence, protagonist B, for example, the girl moved the ball to the other box (box 2) and the following control questions and the test question were asked:

¹ Direct translation from German language "What wanted Tom that Monkey does with the painting?"

- Control Question 1: In which box did the boy put his ball in the beginning? [correct answer: box 1]
- Control Question 2: Where is the ball now? [correct answer: box 2]
- Control Question 3: Who put it there? [correct answer: the girl]
- Test question: When the boy returns, where will he look for the ball first? [correct answer: box 1]

Children were corrected when giving false answers to the control questions.

The order and sides of the two trials in the FB task were counterbalanced.

Results

Coding

Two observers coded all sessions from the videotape. An additional independent coder who was blind to the hypotheses coded a random sample of 25% of all sessions for reliability. The consistency of the ratings was high for all test questions in the desire and belief task (all Cohen's $\kappa \geq .84$).

Desire Task. In each trial, children received the test question that asked them to reproduce the desire's complement and a control question that asked them for the agent's actual action. When children answered this control question incorrectly, we excluded this trial from analyses (11.48% of trials). We scored the answer to the test question as correct when children reproduced the correct complement. In cases, where children stated ignorance, the actual action or gave random answers, we coded the answer as incorrect. This resulted in a binary outcome.

False Belief Task. Answers to the test question were scored as correct when children predicted that the protagonist would look for the object in the initial location. Based on their score we categorized children in two categories of False Belief Understanding: Children who answered two of two administered false belief test questions correctly were assumed to have a reliable understanding of beliefs and were classified as Passers ($n = 28$). Children who failed to correctly answer one or both trials were classified as Non-Passers ($n = 33$).

Plan of analysis

The aim of this study is to test the asymmetry and the symmetry account against each other. To do so, we tested the accounts' explicit predictions by applying two different approaches of analysis:

First, we looked at children's absolute performance as a function of mental state (wicked desires, neutral desire and belief) and age. The asymmetric account makes a very clear prediction here: Younger children should be better in ascribing desires than in ascribing beliefs.

The symmetric account makes a different but equally clear prediction: Younger children should be better in ascribing neutral desires than in ascribing wicked desires.

Second, we took a relational approach. How is subjective reasoning (operationalized via belief reasoning) related to wicked and neutral desire reasoning? Here, the symmetric account makes a clear prediction: Performance in ascribing wicked desires should be strongly related with ascribing false beliefs.

Absolute Performance

We compared children's absolute performance as a function of mental state (neutral desire, wicked desire, false belief; see Fig. 2) and age to test the asymmetric prediction: younger children perform better in ascribing desires than beliefs, and the symmetric prediction: younger children perform better in ascribing desires than beliefs.

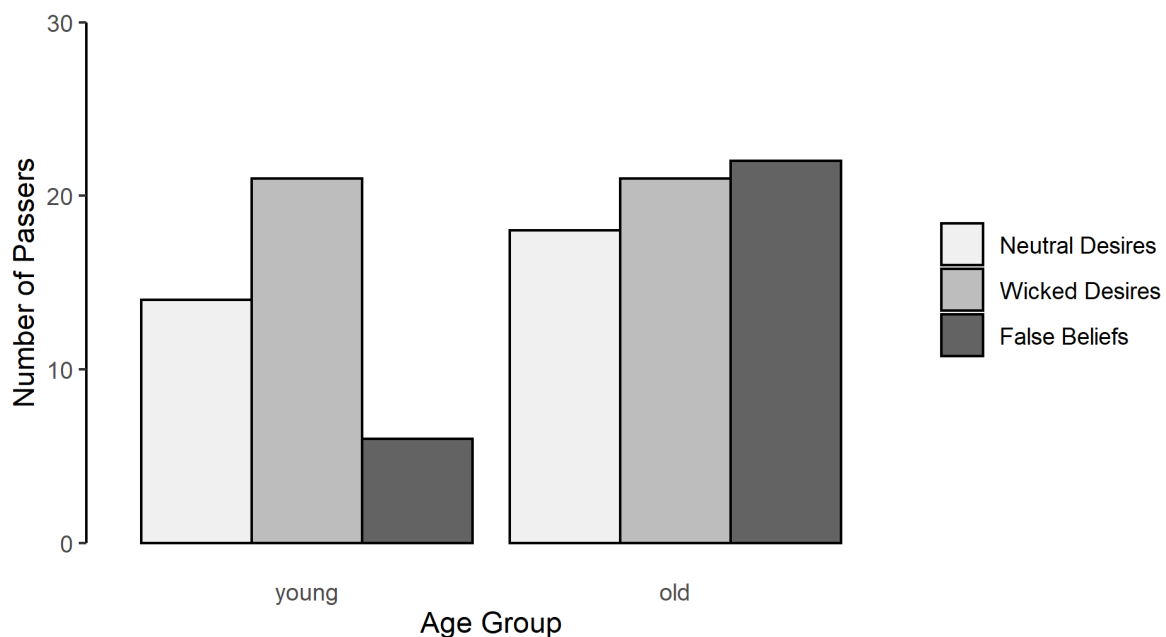


Figure 2. Number of children passing all available trials by mental state for children younger and older than median age.

To this end, we first divided our sample in younger and older children via median split ($Mdn = 1251$ days; 31 children \leq median; 30 children $>$ median). Performance in each task was operationalized as passing vs. non-passing. We did this for two reasons: First, we had

excluded trials from analyses in which children failed the control question. This makes sum scores incomparable. Second, we used different measures for desire reasoning and belief reasoning, which further reduces the informativeness of comparing sum or mean scores. For each mental state, passers were defined as children who succeeded in all available trials (i.e., non-excluded trials) for this state¹. To test the prediction of the asymmetric account, we compared the number of passers in desire and belief conditions. Table 1 shows how many children passed the desire tasks compared to the false belief tasks. In the younger age group, more children passed the desire tasks, neutral as well as wicked, than the false belief tasks. This was different for older children. There, more children passed the false belief tasks aligning performance in desire and belief tasks. This pattern was supported by separate one-sided McNemar's exact tests which revealed that, for younger children, significantly more children passed the desire tasks than the belief tasks (neutral desires: $OR = 3.667, p = .029$; wicked desires: $OR = 16, p < .001$). This was not the case for younger children (neutral desires: $OR = 0.333, p = .965$; wicked desires: $OR = 0.8, p = .746$). Thus, as predicted by the asymmetric account, we observed children to perform better in the desire tasks than in the belief tasks.

Table 1a

Contingencies of Passing: Neutral Desires x False Beliefs

			<i>Neutral Desires</i>	
			Non-Passers	Passers
Younger Children ($n = 30$)	<i>False Beliefs</i>	Non-Passers	13	11
		Passers	3	3
Older Children ($n = 29$)	<i>False Beliefs</i>	Non-Passers	5	2
		Passers	6	16

Note. One younger and one older child are not depicted here because they failed both control questions.

¹ For neutral desires, ten children failed to answer the control question for one trial correctly. These children were counted as passers if they succeeded on the remaining trial. Two children failed control questions for both trials and therefore could not be included in comparisons regarding neutral desires at all. For wicked desires, fourteen children failed on one control question and one on both. Note that for false beliefs, we followed the standard procedure. Children received control questions before test questions and false answer to control questions were corrected. Thus, no trials had to be excluded from analyses for this state.

Table 1b

Contingencies of Passing: Wicked Desires x False Beliefs

			<i>Wicked Desires</i>	
			Non-Passers	Passers
Younger Children (<i>n</i> = 31)	<i>False Beliefs</i>	Non-Passers	9	16
		Passers	1	5
Older Children (<i>n</i> = 29)	<i>False Beliefs</i>	Non-Passers	3	4
		Passers	5	17

Note. One older child is not depicted here because they failed both control questions.

To test the prediction of the symmetric account, we compared the number of passers in the neutral and the wicked desire condition. As can be seen in Table 2, results are opposite to the symmetric prediction: Younger children passed the wicked desire task more often than the neutral desire task. The same held for older children. Separate one-sided McNemar's exact tests support this pattern: For both age groups, we did not find that children pass the neutral desire task more often than the wicked desire task (younger: $OR = 1.25$, $p = .998$; older: $OR = 0$, $p = 1$). Rather, the opposite was the case. A one-sided McNemar's testing the reversed pattern found that significantly more of the younger children passed the wicked desire than the neutral desire task (younger: $OR = 8$, $p = .012$). For older children, McNemar's test in this direction could not be conducted as no older child failed the wicked desire task while passing the neutral desire task.

Table 2

Contingencies of Passing: Wicked Desires x Neutral Desires

			<i>Neutral Desires</i>	
			Non-Passers	Passers
Younger Children (<i>n</i> = 30)	<i>Wicked Desires</i>	Non-Passers	8	1
		Passers	8	13
Older Children (<i>n</i> = 28)	<i>Wicked Desires</i>	Non-Passers	8	0
		Passers	2	18

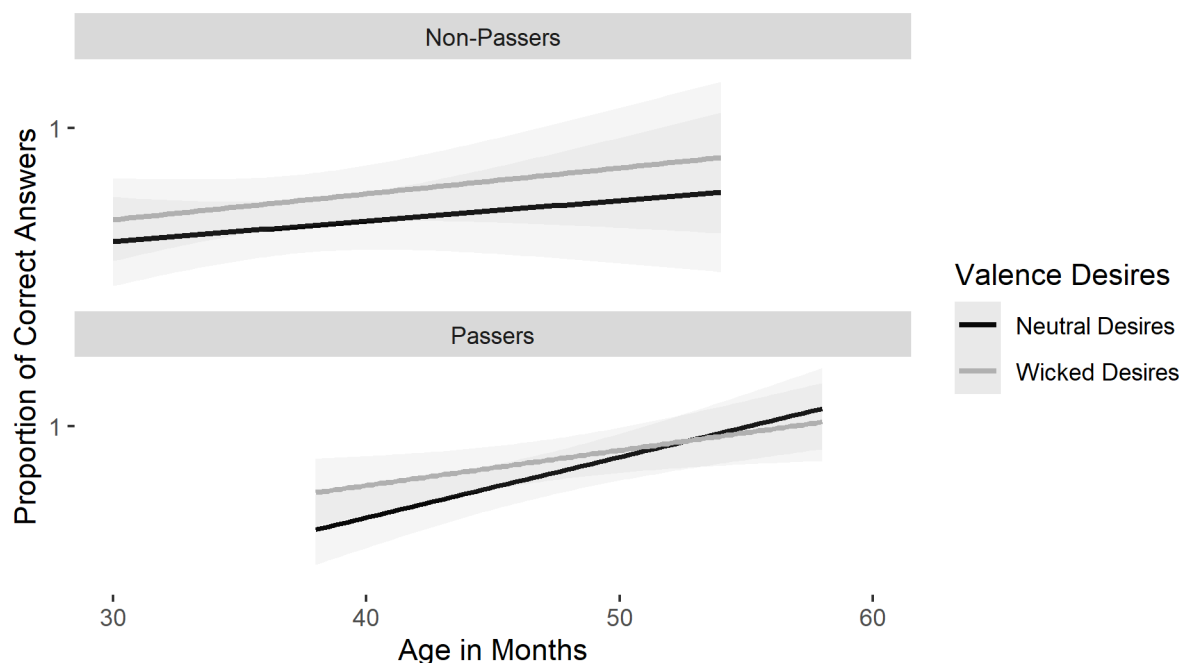
Note. One younger and two older children not depicted here because they failed both control questions.

Thus, we find the exact opposite of what the symmetric account predicted: Children performed actually better in the wicked than in the neutral desire task.

Relational Approach

We used a relational approach to test the prediction of the symmetric account that wicked desire and false belief reasoning share the same core competency (subjective reasoning) and accordingly should be related strongly. The symmetric account remains agnostic about the relation of false belief reasoning and neutral desire reasoning. It is possible but not necessary that there is a relation between these capacities caused by factors as for instance verbal demands. However, if any this relation should be less pronounced as it should not be caused by a shared underlying competency. Overall, such a pattern would be reflected by an interaction of children's belief reasoning capacity and the desire's valence in the following way: There is a relation between False Belief Understanding and children's performance on the wicked desires task, but no or a less strong relation between False Belief Understanding and the neutral desires task.

We set up a Generalized Linear Mixed Model with binomial error structure and a logit link function. We included the interaction of False Belief Understanding and Valence of Desires as a fixed effect. Age was included as a control variable and to account for repeated measures, we included children's ID as random intercept effect. We checked for the model stability (see Appendix) and multicollinearity (all $VIFs \leq 1.526$).



To test for an overall effect of False Belief Understanding and Valence of Desires, we compared this full model with a null model which only contained the control variable, age,

and the random intercept for children. This comparison revealed no significant effect (likelihood ratio test: $\chi^2 = 14.891$, $df = 3$, $p = 0.058$).

What does this tell us? Children's performance depends to a substantial degree on their age and not on their belief reasoning capacity or the desire's valence (see also Fig. 2). With regard to existing doubts to the strict interpretation of p -values and against the background that the obtained value exceeds the criterion of 0.05 only to a very slight degree, we took a closer look at the model. Yet, keep in mind that the overall model comparison did not reach significance. There was no interaction effect of False Belief Understanding and Valence of Desires ($b = 0.309$, $p = .802$) and also no main effect of False Belief Understanding ($b = 0.206$, $p = .908$). Only, the effect of Valence of Desires ($b = 1.403$, $p = .046$) and age ($b = 2.343$, $p = .008$) reached significance. Thus, even if the full-null model comparison was significant there would be no evidence that belief reasoning is related to desire reasoning, neither for wicked nor neutral desires¹. We only found children to become better over age and to be (possibly) more likely to reproduce wicked desires than neutral desires.

Discussion

The guiding question of the present study was how subjective conceptions of beliefs and desires develop in relation to each other. The symmetry account claims that a subjective understanding of desires develops before a subjective understanding of beliefs. The asymmetry account, in contrast, assumes that both concepts are acquired in tandem because they rely on a common core capacity to reason subjectively. Here, we tested these two accounts against each other in novel ways. First, we applied a more straightforward measure than earlier studies. While earlier studies asked children to ascribe desire-dependent emotions, we asked children to directly reproduce sentential complements of value-incompatible (wicked) and -compatible (neutral) desires. Second, we directly compared children's performance in desire reasoning to their subjective belief reasoning capacity.

¹Separate regression analyses on aggregated mean scores of children performance yielded similar results. Regression analyses revealed that belief reasoning was neither related to wicked desire reasoning ($b = 0.183$, $p = .129$) nor to neutral desire reasoning ($b = 0.215$, $p = .087$), especially when controlling for age (wicked desires: $b = 0.041$, $p = .788$; neutral desires: $b = -0.001$, $p = .948$). In contrast, there was a relation between wicked and neutral desire reasoning ($b = 0.761$, $p < .001$), even when controlling for age ($b = 0.708$, $p < .001$).

The main findings of the present study were the following: First, younger children were generally better in reasoning about desires than about beliefs. This was the case not only for neutral but also for wicked desires. Older children succeeded in both, belief and desire ascription. Second, younger children in our study were better in reasoning about wicked desires than about neutral desires. And third, we found no relation between wicked desires and belief reasoning.

Overall, thus, the present findings are compatible with the predictions of the asymmetry account. This account makes one clear prediction: Children develop an understanding of desires before beliefs. In line with this prediction, the present results indicated that younger children were more proficient in ascribing both value-compatible and value -incompatible desires than in ascribing beliefs while older children succeeded in both.

In contrast, the present results are not in line with the predictions of the symmetry account in several respects. First, the symmetric account predicts that young children should be able to reason about neutral desires since these can be grasped on the basis of a purely objective teleological reasoning scheme. In contrast, they should fail on tasks that require the ascription of value-incompatible desires since these cannot be understood on the basis of objective teleology but require truly subjective reasoning. However, the results of the present study showed that children did not only have no such difficulties, but were, in fact, even better at reasoning about value-incompatible desires. Second, the symmetric account predicts that genuine desire reasoning and belief reasoning, since they share an underlying core capacity (a grasp of subject perspective), should emerge in tandem. So far, evidence for this assumption came from indirect findings that children begin to reason about value-incompatible desires and beliefs about the same age. Here, we took a more direct approach than comparing ages of onset and investigated inter-task correlations. In contrast to the predictions of the symmetry account, we did not find any correlations between false belief and wicked desire ascription.

The present findings stand in contrast with the results of two previous lines of research that provided evidence for the symmetry account. First, as reviewed above, previous research on children's developing understanding of wicked desires used emotion ascription as its dependent measure. The central finding was that children ascribe subjective, wicked desire dependent emotions (such that an agent feels good after realizing her bad ends) only from age four to five (Yuill, 1984; Yuill et al., 1996). Why, then, is there this discrepancy between our findings and those previous results? Clearly, from the present results alone we cannot tell, and

currently we can thus only speculate. But the following seems plausible: Emotion ascription involves extraneous complexities that do not apply in the case of memory-for-complements tasks. In general, ascribing emotions based on mental states appears to be inferentially more demanding than ascribing the mental state itself (Harris et al., 1989). Even more so in the case of desires conflicting with moral norms. This case requires children to coordinate the representations of an outcome for the victim that causes negative emotions (e.g., being hurt) with positive emotions on the desirer's side (having succeeded in hurting someone). Compare this to the case of neutral desires, in which neutral/slightly positive outcomes and emotions on the "victim's" side have to be paired with positive emotions on the desirer's side. One can easily see how the first case adds further challenge to the task. More generally, it is not even completely clear what the normatively correct answer is in such wicked desire cases. Mature reasoners would at least ascribe some form of mixed emotion ("in some sense she is glad because of her 'success', but true happiness looks different"). Developmentally, indeed, older children, from around the age of ten, stop ascribing purely positive emotions to the agent. Instead, they add a negative notion to account for some form of remorse on side of the successful yet wicked desirer (Yuill et al., 1996).

Second, previous research on children's understanding of incompatible desires in the context of competition has suggested that children acquire a notion of the subjectivity of desires only from around age four to five (Priewasser et al., 2013). How can we explain the discrepancy between those and the present findings? Again, from the present study alone we cannot conclusively tell, but the following seems plausible: Competitive moves may have been difficult for young children not for cognitive reasons (understanding the incompatibility of one's own and the opponent player's subjective desires), but for broadly motivational or ethical ones: Children may have well understood the different subjective desires, but may not have translated this into competitive action since they found it difficult to overcome norms of politeness, or may have simply cared more about playing together than about winning. Indeed, when in a recent study, children were asked directly in such competitive scenarios, they did not show difficulties in ascribing their own and their opponent's mutually incompatible desires (Proft et al., 2021). At this stage of inquiry, thus, the empirical situation remains complex. Currently, we can only speculate in post-hoc ways why different approaches (asking about neutral or wicked desires directly, or about desire-dependent emotions; or recording competitive actions) reveal different developmental patterns. Future research needs to go beyond post-hoc attempts at making sense of such seemingly contradictory patterns of findings and investigate the development of understanding

subjective desires and beliefs in different forms and with different type of measures in more systematic and comprehensive, a priori planned, ways.

To summarize, the present study investigated children's developing understanding of beliefs and various forms of desires in order to test competing predictions of two broad classes of accounts against each other. In contrast to previous studies, arguably it used the purest measure for desire understanding, memory-for-complements, and directly compared subjective belief and desire reasoning. The results suggest that genuine subjective understanding of desires and understanding of beliefs do not emerge in tandem. Instead, it suggests an asymmetric development: A concept of desires that can be relativized to subjective standpoints is already present before children can reason about beliefs in comparably subjective ways. How sophisticated this early subjective conception of desires is, however, and what exactly its scope and limits are, needs to be explored more systematically in future research.

References

- Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, *116*(4), 953.
- Choe, K. S., Keil, F. C., & Bloom, P. (2005). Children's understanding of the Ulysses conflict. *Developmental Science*, *8*(5), 387–392. doi: 10.1111/j.1467-7687.2005.00426.x
- 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. doi: 10.1016/S0885-2014(02)00073-4
- Fizke, E., Barthel, D., Peters, T., & Rakoczy, H. (2014). Executive function plays a role in coordinating different perspectives, particularly when one's own perspective is involved. *Cognition*, *130*(3), 315–334. doi: 10.1016/j.cognition.2013.11.017
- Fodor, J. A. (1975). *The language of thought* (Vol. 5). Harvard university press.
- Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences*, *7*(7), 287–292. doi: 10.1016/S1364-6613(03)00128-1
- Gordon, R. M. (1986). Folk psychology as simulation. *Mind & Language*, *1*(2), 158–171.
- Hadwin, J., & Perner, J. (1991). Pleased and surprised: Children's cognitive theory of emotion. *British Journal of Developmental Psychology*, *9*(2), 215–234. doi: 10.1111/j.2044-835X.1991.tb00872.x
- Harris, P. L., Johnson, C. N., Hutton, D., Andrews, G., & Cooke, T. (1989). Young Children's Theory of Mind and Emotion. *Cognition & Emotion*, *3*(4), 379–400. doi: 10.1080/02699938908412713
- Hogrefe, G.-J., Wimmer, H., & Perner, J. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child Development*, *57*, 567–582.
- Leekam, Susan., Perner, J., Healey, L., & Sewell, C. (2008). False signs and the non-specificity of theory of mind: Evidence that preschoolers have general difficulties in understanding representations. *British Journal of Developmental Psychology*, *26*(4), 485–497. doi: 10.1348/026151007X260154
- Lichteremann, L. (1991). Young children's understanding of desires. Third year project report. *Unpublished Manuscript*.
- Moore, C., Jarrold, C., Russell, J., Lumb, A., Sapp, F., & MacCallum, F. (1995). Conflicting desire and the child's theory of mind. *Cognitive Development*, *10*(4), 467–482. doi: 10.1016/0885-2014(95)90023-3

- Perner, J., & Esken, F. (2015). Evolution of human cooperation in *Homo heidelbergensis*: Teleology versus mentalism. *Developmental Review*, 38, 69–88. doi: 10.1016/j.dr.2015.07.005
- Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, 5(2), 125–137.
- Perner, J., Priewasser, B., & Roessler, J. (2018). The practical other: Teleology and its development. *Interdisciplinary Science Reviews*, 43(2), 99–114. doi: 10.1080/03080188.2018.1453246
- Perner, J., & Roessler, J. (2010). Teleology and causal understanding in childrens' theory of mind. In *Causing human action: New perspectives on the causal theory of action* (pp. 199–228).
- Perner, J., & Roessler, J. (2012). From infants' to children's appreciation of belief. *Trends in Cognitive Sciences*, 16(10), 519–525. doi: 10.1016/j.tics.2012.08.004
- 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.
- Priewasser, B., Roessler, J., & Perner, J. (2013). Competition as rational action: Why young children cannot appreciate competitive games. *Journal of Experimental Child Psychology*, 116(2), 545–559. doi: 10.1016/j.jecp.2012.10.008
- Proft, M., Hoss, C., Paredes, K. M., & Rakoczy, H. (2021). Do children understand desires before they understand beliefs? A comparison of 3-year-olds' grasp of incompatible desires, competitive games and false beliefs. *Cognitive Development*, 57, 101009.
- Rakoczy, H. (2010). Executive function and the development of belief-desire psychology: Executive function and belief-desire psychology. *Developmental Science*, 13(4), 648–661. doi: 10.1111/j.1467-7687.2009.00922.x
- Rakoczy, H., Warneken, F., & Tomasello, M. (2007). “This way!”, “No! That way!”—3-year olds know that two people can have mutually incompatible desires. *Cognitive Development*, 22(1), 47–68. doi: 10.1016/j.cogdev.2006.08.002
- Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14-and 18-month-olds. *Developmental Psychology*, 33(1), 12.
- Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge university press.

- Sprung, M., Perner, J., & Mitchell, P. (2007). Opacity and Discourse Referents: Object Identity and Object Properties. *Mind & Language*, 22(3), 215–245. doi: 10.1111/j.1468-0017.2007.00307.x
- Wellman, H. M. (2011). Developing a theory of mind. In *The Wiley-Blackwell handbook of childhood cognitive development, 2nd ed* (pp. 258–284). Wiley-Blackwell.
- Wellman, H. M., & Banerjee, M. (1991). Mind and emotion: Children's understanding of the emotional consequences of beliefs and desires. *British Journal of Developmental Psychology*, 9(2), 191–214. doi: 10.1111/j.2044-835X.1991.tb00871.x
- Wellman, H. M., & Liu, D. (2004). Scaling of Theory-of-Mind Tasks. *Child Development*, 75(2), 523–541. doi: 10.1111/j.1467-8624.2004.00691.x
- 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.
- Yuill, N. (1984). Young children's coordination of motive and outcome in judgements of satisfaction and morality. *British Journal of Developmental Psychology*, 2(1), 73–81.
- Yuill, N., Perner, J., Pearson, A., Peerbhoy, D., & Ende, J. (1996). Children's changing understanding of wicked desires: From objective to subjective and moral. *British Journal of Developmental Psychology*, 14(4), 457–475. doi: 10.1111/j.2044-835X.1996.tb00718.x
- Zaitchik-Samet, D. (1989). *On the preschooler's problem with the false belief task*. Massachusetts Institute of Technology.

Appendix A

Validation study

An additional validation study was implemented to validate the new method of sentential complements with other, more established, tasks. We tested children's memory-for-complements for cognitive mental states (beliefs) and compared their performance to their belief ascription ability in the standard change-of-location false belief task. We tested ten 3- to 4-year-olds (37-59 months; *mean age*= 49.5 months, 6 girls). Children were recruited from the same databank as in the main study.

Design

In the belief memory-for-complements task, children had to reproduce sentential complementation of false beliefs similar to the "unexpected content task" (the "Smarties" test; (Hogrefe, Wimmer, & Perner, 1986; Perner, Leekam, & Wimmer, 1987). This task asked children to complement beliefs about the content of a given container. Children received scenarios which differed regarding the strength of the default expectation what the container might contain. In one scenario the container was neutral and provided no hint about its content, for example, a plain red. In the other scenario, the outer appearance of the container showed clearly what it usually contains, a Smarties container. This container produced a high expectation containing chocolate beans. Each condition consisted of two trials. The order of the trials was counterbalanced. After the memory-for-complements task, children received two trials of the standard change-of-location false belief task (Wimmer & Perner, 1983). The whole test session lasted ~20 minutes.

Material and Procedure

Memory-for-complements task with beliefs

In this validation study, we used the same set-up of a hand puppet theater with the boy Tom and Monkey, again separated by an occluder, and Frog. Frog again brought a big box containing four smaller containers into the scene. In each trial, Tom uttered his belief about the content of one of the four smaller containers (e.g., plain red box or Smarties container). In order to match the belief utterance to the format of the desire utterance in the main study (e.g., "Monkey should hang the necklace on the stand"), the belief about the container's content was framed as belief about Monkey's action (e.g., "Monkey certainly takes a hat out of the

red box/ Smarties out of the Smarties box”¹). After Tom’s utterance, Monkey opened the container and showed the child (but not Tom) its actual content. E1 then asked the test question about Tom’s belief (“What did *Tom* think that Monkey take out of the red box/ Smarties box?”) and control question about Monkey’s actual action (“What did he actually take out of the red Box/ Smarties box?”)

Change-of-location false belief task

Children received the same Change-of-location false belief task as in the main study.

Coding

Memory-for-complements task

Trials in which children answered the test question with the sentential complement about Tom’s respective belief verbally (e.g., “hat”) were coded as correct (1). When children failed to refer to the stated belief, trials were coded as incorrect (0). Control questions about the container’s actual content were coded as correct (1) when children referred it verbally (e.g., teddy bear) or used similar descriptions (e.g., puppet). All other answers and missing answers were coded as incorrect (0).

Change-of-location False Belief task

See coding for change-of-location false belief task in main study.

Results and Discussion

As in the main study we excluded trials of the memory-for-complements from analyses when children failed the respective control question (12.5% of trials). Table A1 shows how many children passed (succeed on all available trials) the memory-for-complements and the standard false belief task.

¹ Original phrasing in German language: “Der Affe holt bestimmt einen Hut aus der roten Box” or “Der Affe holt bestimmt Smarties aus der Smartiesbox“.

Table A1

Contingencies of Passing: Memory-for-Complements Belief x Standard False Belief

		<i>Standard False Belief</i>	
		Non-Passers	Passers
<i>Memory-for-Complements</i>	Non-Passers	4	0
<i>Belief</i>	Passers	2	4

A two-sided McNemar's exact test revealed no differences between children passing the memory-for-complements and the standard false beliefs ($OR=0, p=.5$). Mean performance in both tasks was highly correlated ($r_s=.888, p<.001$). These results support that our adaptation of the memory-for-complements tasks is a valid measure for children's concept of mental states.

Appendix B*Detailed procedure of the main study*

- E1 and child look together at animal photos
- Tom (E2) joins the scene, child gets introduced to Tom, look all together at animal photos

Warm-up phase step 1

- E1 and child watch together a video of a woman on a laptop computer.
- E1: "What is the woman doing?"
- The woman performs the following actions
 - o Put something (hedgehog) on grass
 - o Put something (hedgehog) in a basket
 - o Tear something (paper)
 - o Put something (paper) on a shelf
 - o Put something (cup) in a box
 - o Break something (cup)
 - o Hang something (watch) up on a stand
 - o Put something (watch) in a tin
- If child cannot name an action, E1 names the action and asks the child to repeat it

Warm-up phase step 2

- Tom joins the scene again
- E1 turns her back on the child
- Tom tells the child a fact about him (2 trials) “My favorite color is green” , “My favorite animal is the dog”
- E1 turns to child again and asks surprised “What did Tom say?”
- ... child answers
- E1 again turns her back on the child
- Tom performs an action (2 trials): puts stone on the table, takes sticker out of his trouser pocket
- E1 turns to child again and asks surprised “What did Tom do?”
- ... child answers
- Frog (E2) joins the scene and brings a box with some of his beloved, self-made belongings with him
- Frog leaves box with child and E1 and leaves the scene again

Experimental trials

- E1 takes one of the belongings out of the box and puts it on the table
- Tom and Monkey join the scene
- Monkey and Tom are separated by an occluder
- E1: “Is Tom able to see what the Monkey is doing?”
- ... child answers (is corrected if wrong)
- E1 turns back on child
- Tom: “Monkey should tear the painting.”
- Monkey performs alternative neutral action: Monkey puts painting on the shelf
- E1 turns to child again: “What did Tom want Monkey to do?”
- Child answers...
- E1: “What did Monkey actually do?”
- Child answers...
- Next trial begins...
- Other scenarios:
 - Desire “Monkey should break the play drum”, actual action: Monkey puts the play drum in a box
 - Desire “Monkey should put the necklace on the stand”, actual action: Monkey puts the necklace in a case

- Desire “Monkey should put the playdough dog on the grass”, actual action:
Monkey puts playdough dog in the basket

Change-of-location False Belief Task

Look, this is the boy with his ball “Hello!”. The boy loves to play with his ball (boy plays ball). Now, the boy wants to go to the playground. He puts his ball in the blue box and goes to the playground. The boy is on the playground now. He is very far away. He cannot see and cannot hear us anymore. Now, the girl comes. She takes the ball out of the blue box and puts it in the red box. Then the girl leaves again.

- Control question 1: “Where did the boy put the ball first”?
- Control question 2: “Where is the ball now?”
- Control question 3: “Who put it there?”
- Test question: “When the boy returns, where will he look for the ball first?”

Appendix B: Schünemann, Proft, & Rakoczy (2021)

Schünemann, B., Proft, M., & Rakoczy, H. (2021). *Children's developing understanding of the subjectivity of intentions – a case of “advanced Theory of Mind”* [Manuscript submitted for publication]. Department of Developmental Psychology. University of Göttingen.

This article has been submitted for publication in Journal of Cognition and Development, published by Taylor & Francis.

**Children’s developing understanding of the subjectivity of intentions – a case of
“advanced Theory of Mind”**

Word count: 10508

Abstract

When and how do children develop an understanding of the subjectivity of intentions? Intentions are subjective mental states in many ways. One way concerns their aspectuality: Whether or not a given behavior constitutes an intentional action depends on how, under which aspect, the agent represents it. Oedipus, for example, intended to marry Yocasta, but did not intend to marry his mother (even though in fact, but unbeknownst to him, Yocasta was his mother). In the present study, we investigated the trajectories and determinants of children's developing understanding of (less dramatic forms of) the aspectuality of intentions. In two studies, children aged 3-9 observed an agent who acted intentionally but based on some mis-representation regarding the target of her action. The agent grasped a box that contained A and B while believing that it only contained A but not B. Children were asked about the aspectuality of the agent's intention (in particular, whether she intended to grasp B). When asked to do so spontaneously, children younger than 8 failed (falsely claiming that the agent intended to grasp B). In contrast, in a simplified format in which children were scaffolded through the required inferential chains, children from age 6 succeeded. Children's general capacity for meta-representation appeared to be necessary but not sufficient by itself for understanding the aspectuality of intentions. The present findings suggest that the appreciation of the aspectuality of intentions is part of an advanced theory of mind that develops in much more protracted ways than basic theory of mind.

Keywords: Theory of Mind, Intentions, Aspectuality, Representations, Conative Mental States

In the series *Breaking Bad*, the Drug Enforcement Agency (DEA) detective Hank repeatedly talks to his inconspicuous, seemingly innocent brother-in-law Walt about the DEA's strategy in trying to hunt down the infamous Meth cook known by his nickname "Heisenberg". Hank, as almost everyone, believes Heisenberg to be some kind of monster living in a parallel mafia world somewhere out there. As it happens, though, the inconspicuous Walt is in fact Heisenberg. So, Hank in effect reports secret information about the DEA's strategy to Heisenberg himself. Did he do that intentionally? Well, in some broad sense yes, but in a crucial, narrower sense, he did not. While he intended to tell Walt about the DEA's strategy, he clearly did not intend to tell Heisenberg. How is this possible?

This is possible due to the so-called aspectuality of intentions (Anscombe, 1957; Searle, 1983): There are always many different descriptions that potentially apply to a given intentional action (such as moving your hand, moving oxygen molecules in the air, waving to a man, greeting your neighbor). But the action is typically performed intentionally only under some specific descriptions or aspects, and not under others. Whether or not a given intentional action description applies, depends on the way the agent represented the action. Intentionality is thus relative to the subjective standpoint of the agent. Such subjective aspectuality is a crucial feature of intentional action. It may come about in various ways: One way is that you actually know about the relevant aspects, but simply do not care about one of them or value another one more strongly. Side effects which are foreseen but not themselves intended by the agent are such a case much discussed in moral philosophy and psychology. Imagine a doctor who gives you a treatment that is the only means to save your life but which will also make you lose your hair. Did she intend to make you bald? In some crucial sense, she did not. She intended to save your life, realized that the only way to do so is to give you the treatment and foresaw (but did not intend to bring about) the side effect of making you bald.

Much more basically, however, the aspectuality of intentional action can be determined by what the agent knows or believes about the relevant situations and her actions. The agent may only have a partial representation of the relevant situation and may be simply unaware of some crucial information. Imagine, for example, A visits a friend B. Sitting in her kitchen while B is in the bathroom, A sees a cookie on the table and eats it. B enters the room and exclaims in shocked voice: "No, tell me you didn't eat my only birthday present!" As it turns out, the cookie was the only birthday present B got this year. So, A intended to eat the cookie but did not intend to eat B's only birthday present – simply because she was unaware of (and would never have been able to think of) this aspect of the cookie.

Another case comes about when the agent does not act on the basis of such partial representations (where an object is both X and Y, but the agent only represents it as X and has no representation regarding the Y-ness), but on the basis of mis-representation (where an object is both X and Y, but the agent believes it is X and believes it is not Y). Our Breaking Bad example above constitutes such a case: Hank represents his interlocutor as being Walt and as not being Heisenberg, and thus intentionally reveals DEA secrets towards him under the description “telling Walt” but not under the description “telling Heisenberg”.

From the point of view of cognitive development, the question is when and how children develop a conceptual framework that allows them to understand these complex but foundational features of intentions. Ontogenetically, understanding intentional action develops in degrees and stages over protracted time courses from infancy to school age. Even so much so that some researchers have wondered about a “paradox of intentions”: While they are “the simplest and most obvious mental state” whose understanding is indispensable for making sense of any interaction, they are at the same time the state “the most difficult to understand completely” (Astington, 2001, p. 88). Infants from around the end of the first year develop a rudimentary grasp of intentional actions in understanding an agent’s goals and distinguishing intentional from unintentional actions (Behne et al., 2005; Gergely et al., 2002; Meltzoff, 1995; Woodward, 1998). Much later, children from around age four to five begin to develop a more nuanced understanding of intentions that involves, for example, an appreciation of commitment (intentions commit you to performing an action in ways in which a mere desire does not; Shultz, 2002) or the causal self-referentiality of intentions (desires get fulfilled whenever their content is satisfied, but intentions only are fulfilled if they play the right kind of role in bringing it about that their content is fulfilled; Shultz & Wells, 1985).

So, when and how do children acquire a grasp of the subjective aspectuality of intentions? Generally, so far there has been very little research on this question. Initial studies focused on aspectuality based on the agent’s *partial representation* of the situation (e.g. Anne intentionally gives the keys to Kathy’s dad. Kathy’s dad happens to be the policeman, unbeknownst to Anne. So, did Anne intend to give the keys to the policeman?). Results revealed strikingly late competence: Children answered incorrectly (“yes”, in the above example) until age 8 or even later (Kamawar & Olson, 2011). Subsequent studies, however, suggest that these results revealed performance rather than competence limitations and significantly under-estimated younger children’s competence. Once a simpler and more engaging method with more relevant content was administered, even 5-year-olds performed competently (Proft et al., 2019).

But what about the more complex case of subjective aspectuality of intentions based on the agent's *mis-representation* of the relevant situations? When and how do children develop an appreciation of this form of aspectuality? This is the topic of the present paper. Inspired by classical Theory of Mind vignettes (Wimmer & Perner, 1983) and more recent variations (Rakoczy et al., 2015), we investigate children's representation of scenarios of the following structure: Object A is put in box 1 and object B is put in box 2 in the presence of an agent. Unbeknownst to the agent, A is then transferred to box 2, so that really, both A and B are in box 2 whereas the agent mistakenly thinks that box 2 contains only B and not A. The agent then intentionally grasps box 2 (that, factually, contains both A and B) and the target question is: did she intend to grasp A? In our scenario, the agent knows about the existence of A and B in the scene. This results in two possible A-related descriptions applicable to the boxes, "contains A"/"does not contain A". The agent then mis-represents box 2 to which she applies "does not contain A" (rather than simply being ignorant regarding A), and her action is performed under that mis-represented description. Accordingly, her action of grasping box 2 is not intentional under the description "contains A". This case where an agent actively thinks that a given description ("take the box that contains A") does not apply to her action is thus different from and more complex than cases of partial representation where an agent is simply ignorant about a potential description under which she acts.

On the basis of this paradigm, we address the following three interrelated questions: First, what is the age of onset of children's appreciation of the aspectuality of intentions based on an agent's mis-representation? Generally, in social cognitive development, an understanding of *what* other agents represent (i.e., whether they represent something) develops before an understanding of *how* they represent something. Regarding visual perspective taking, for example, children from their second year understand whether another agent can or cannot see all the things in a scene that they themselves can see (called "level I perspective-taking"; Flavell et al., 1978). But only from around age 4 do they understand that different agents can see the same thing (e.g., a digit) but see it in different ways depending on their viewpoint (e.g., as a "6" vs. "9") (called "level II perspective-taking"; Flavell et al., 1981). 2-year-olds can thus understand that another agent only has a partial representation of a scene (see objects X and Y, while the child herself sees X, Y and Z). But only 4-year-olds can explicitly understand that another agent may represent a given scene differently from how they themselves see it, and differently from how it really is and thus mis-represent it (e.g., thinking that a misleading object, actually a rock, was a sponge; Flavell et al., 1986).

Against this background, it is plausible to assume that children's understanding of different forms of the aspectuality of intentions – based on partial representation or mis-representation — follows an analogous trajectory. An understanding of mis-representation-based aspectuality would thus be expected to develop after an understanding of aspectuality based on partial representation. Since previous work has revealed an understanding of this arguably more basic type (aspectuality based on partial representation) in children at age 5, we here tested children from preschool to school age and expected competence in the arguably more complex understanding (aspectual intentions based on mis-representation) not before age 5 or later.

The second question pertains to the more fine-grained courses of development: how does children's understanding of mis-representation in general, and their understanding of the aspectuality of intentions based on an agent's mis-representation relate to each other? Clearly, the former is (conceptually) necessary for the latter. But is it sufficient as well? Theoretically, this is possible: Once children have acquired a general notion of mis-representation, they may flexibly put it to work in different areas, including the understanding of the aspectuality of intentions. Alternatively, a general grasp of mis-representation by itself may not be sufficient for understanding mis-representation-based aspectuality. Developmental trajectories of the latter kind (grasping mis-representation is necessary but not sufficient for some more complex understanding) have been found in related areas: concerning the understanding of complex emotions, it takes children some time after they have acquired a concept of "belief" before they can put this concept to use in ascribing belief-based emotions such as surprise (Hadwin & Perner, 1991; Harris et al., 1989). Similarly, regarding the development of ascribing higher-order beliefs: it takes children some time (up to two years) after they have acquired a concept of "belief" before they can put this concept to use in recursive ways in the ascription of higher-order beliefs ("She believes that he believes that p"; Perner & Wimmer, 1985; Sullivan et al., 1994).

In order to address the second question, we compared children's understanding of the agent's mis-representation (that box 2 contains only B and not A while in fact it contains both A and B) directly to their understanding of the aspectuality of intentions (that the agent intended to grasp the box with B but did not intend to grasp the box with A even though in real fact the box intentionally grasped contained both, A and B).

Third, if understanding mis-representations is indeed necessary but not sufficient for appreciating the aspectuality of intentions, what additional capacities are crucial, and under

which conditions can children show such an appreciation? The most obvious candidates are linguistic and domain-general cognitive capacities (working memory and inhibitory control, in particular). Why are these the most obvious candidates? The target tasks clearly pose considerable memory, executive and linguistic demands since children have to remember the agent's subjective viewpoint throughout the narration as well as the different descriptions that apply to the objects in question, and to inhibit their own representations and the urge to answer without considering the particular description. Children's difficulties might also stem from the task's recursive complexity more generally (see Halford et al., 1998). These demands may be particularly prominent in tasks with complex inferential structure. For example, in the present target tasks children have to engage in the following inferential chain: "She has not seen the transfer of A. Therefore, she believes that A is still in box 1 and box 2 contains only B. Therefore, although she has intentionally taken box 2, since she believed A to be somewhere else, she did not intend to take A". Research with inferentially similarly complex tasks in other areas of social-cognitive development (e.g., moral reasoning or trait ascription) has revealed developmental patterns of the following kind: when required to engage in similar inferential chains spontaneously, even 7-year-olds failed; but when guided through the inferential chain by relevant probes, even 4- to 5-year-olds succeeded (Liu et al., 2007; Proft & Rakoczy, 2019). In the context of the present research, it may thus be possible that children are able to put their understanding of mis-representation to use in coming to understand the aspectuality of intentions more easily under conditions of reduced demands on spontaneous inferences.

Accordingly, we applied two different approaches to address our three research questions. First, we measured children's understanding of aspectuality of intentions in a more spontaneous format in Study 1. As our starting point to investigate the development of this capacity, Study 1a addressed the first two research questions: What is the age of onset and how is its development related to understanding mis-representation? The results revealed that not even 6-year-olds grasped the aspectuality of intentions (although they had no trouble with understanding mis-representation). These findings thus suggest that understanding mis-representation is necessary but not sufficient for understanding the aspectuality of intentions based on mis-representation. Study 1b, therefore, addressed the first two questions in concert with the third research question and investigated the potential role of general cognitive, linguistic and recursive capacities in a correlational design. Study 1b replicated the basic results of Study 1a and found that children successfully appreciate the aspectuality of intentions in a spontaneous format by the age of 8. Yet, it failed to find evidence for a crucial role of linguistic,

recursive and domain-general cognitive capacities such as working memory and executive function.

Study 2 addressed the third research question with a different approach. By applying a scaffolded format instead of using a correlational design, we manipulated the task structure and thus experimentally reduced linguistic, memory and executive task demands: Children were guided through the two sequential inferential steps required to correctly answer the target question. Results revealed that in this guided format, even 6-year-olds performed competently and appreciated the aspectuality of intentions.

Study 1

Study 1 measured children's ability to spontaneously ascribe intentions to an agent who acted on the basis of a mis-representation as well as their ability to ascribe the underlying mis-representations themselves. Children received test questions on structurally similar scenarios that either addressed a protagonist's mis-representation (false belief about an object's location) or her intention based on such a mis-representation. In different conditions, the agent's mis-representation (concerning an object's location) was implemented in different ways: The Two-Objects scenario employed the classical change of location-vignette (Wimmer & Perner; 1983). The One-Object scenario followed a variation of this vignette in which the change of location is observed but happens under an unknown identity (Rakoczy et al., 2015).

Study 1a

In Study 1a, we administered this spontaneous format of our task to address the first two research questions: What is the age of onset of children's appreciation of the aspectuality of intentions based on an agent's mis-representation? And what pattern does the developmental trajectory follow?

Method

Participants

Seventeen 3-year-olds, sixteen 4-year-olds, seventeen 5-year-olds and sixteen 6-year-olds (39-83 months, $M=60.11$ months, $SD=13.30$; 35 male) were recruited from local childcare centers and from a databank of children whose parents had previously given consent to experimental participation. One 3-year-old participant was tested but had to be excluded from

analysis because she was uncooperative.

Design and Procedure

A 4 (Age group: 3-, 4-, 5- or 6-year-olds) x 2 (Scenario: Two-Objects or One-Object) x 2 (Test Question: Belief or Intention) mixed design was conducted, with Test Question and Scenario as within-subjects factors. Children received eight trials in counterbalanced order, two per combination of Test Question and Scenario. Furthermore, as a covariate we measured children's verbal intelligence at the beginning of each session.

Main Task

Scenarios. Following classical Theory of Mind vignettes (Wimmer & Perner; 1983) and more recent variations (Rakoczy et al., 2015), we included two different scenarios (see Fig. 1 for an overview of the procedure): In the Two-Objects condition (following Wimmer & Perner, 1983), there were two objects A and B. In the presence of the protagonist, A was put in box 1 and B was put in box 2. In the absence of the protagonist, object A was then transferred to box 2. The protagonist thus came to hold the false belief that A was still in box 1. In the One-Object condition (following Rakoczy et al., 2015) there was an object with two identities A and B, and the agent only knew about one identity (A) but not about the other one (B). The object was put into box 1 under its A-aspect, and then transferred under its B-aspect to box 2. The protagonist thus came to hold the false belief that A was still in box 1. To ensure that children had followed the scenario, we asked control questions about the protagonist's ignorance of the object's new location (Two-Object scenario)/second identity (One-Object scenario) and A's initial and current location. Incorrect answers to control questions were corrected.

Test Questions. In half of the test trials, the scenario was followed by a belief test question. In the other half, it was followed by an intention test question. In the belief test trials, the protagonist stated the desire to play with A and children were asked the belief test question.

Belief test question: Which box will the protagonist take now? (correct answer: "Box 1.")

In the intention test trials, the protagonist stated the desire to play with B (since B never changed its location, the protagonist knew it was in box 2). She then took box 2 and the experimenter asked the intention test question (see Fig. 1).

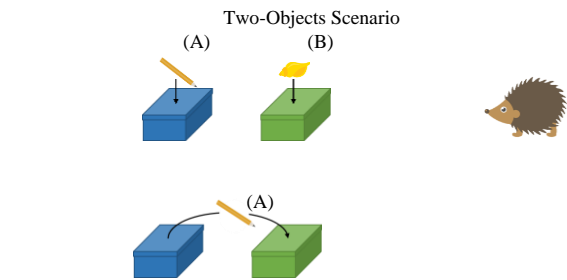
Intention test question: The protagonist intentionally¹ took the box containing B. The box also contains A. Did she also intentionally take A? (correct answer: “No.”)

To validate the task and ensure that our task analysis fits with mature folk psychology, we administered a paper pencil version of the task to 24 adults. In line with the task analysis, all adults replied “No” to the intention test question (see Appendix A for a detailed description of the method and results).

Verbal Intelligence

Furthermore, children’s Verbal Intelligence was assessed via the vocabulary subtest of the Kaufman Assessment Battery for Children II (KABC) (Melchers & Preuß, 2009). This test requires children to label increasingly difficult objects (see Appendix B for a more detailed description and psychometric principles).

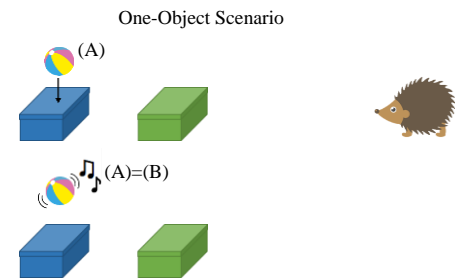
¹ In German, we employed the expression “absichtlich.” In contrast to the English expression “intentionally,” which has a rather stilted connotation, the German *absichtlich* belongs to common speech. A translation that lexically is slightly different but that depicts its acceptance more appropriately would be “on purpose.” Correspondingly, preschool-aged children in other studies handled the expression correctly when describing the intentionality of action (e.g., for the knee-jerk reflex; Lang & Perner, 2002).



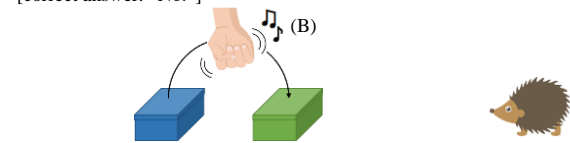
Ignorance CQ: Does the protagonist know that A is in box 2?
[correct answer: "No."]



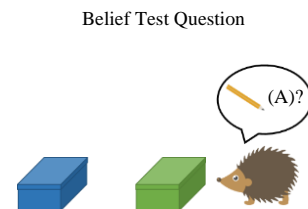
1st Location CQ: Where did I put A in the beginning?
[correct answer: "Box 1."]
2nd Location CQ: Where is A now?
[correct answer: "Box 2."]



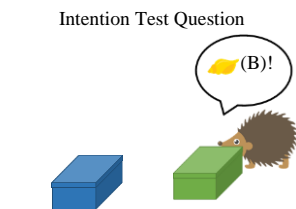
Ignorance CQ: Does the protagonist know that A is also B?
[correct answer: "No."]



1st Location CQ: Where did I put A in the beginning?
[correct answer: "Box 1."]
2nd Location CQ: Where is A now?
[correct answer: "Box 2."]



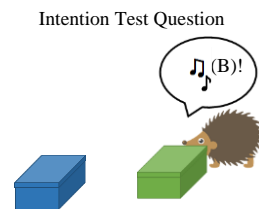
Belief TQ: Which box will the protagonist take now?
[correct answer: "Box 1."]



Intention TQ: ... Did she also intentionally take A?
[correct answer: "No."]



Belief TQ: Which box will the protagonist take now?
[correct answer: "Box 1."]



Intention TQ: ... Did she also intentionally take A?
[correct answer: "No."]

Figure 1. Procedure.

Results & Discussion

Scoring

Our study required 3- to 6-year-olds to ascribe beliefs and intentions to a protagonist who misrepresented the location of an object. To show an understanding of beliefs, children had to acknowledge that the agent acted on her mis-representation (false belief) and thus looked for object A at the wrong location (box 1). To show an understanding of the aspectuality of intentions, children had to acknowledge that the agent did *not* intend to take the object under the description “A”. Overall, children received four trials per test question. In each case, two trials referring to a Two-Objects scenario and two trials referring to a One-Object scenario. Correct answers were scored with 1 and incorrect with 0. Accordingly, for each test question children could receive a score between 0 and 4. The scoring of Study 1b and Study 2 followed the same principles.

The KABC allowed children to receive a verbal intelligence-score between 0 and 39, according to the number of objects they had labeled correctly (see Appendix B for a more detailed description and range of obtained scores).

Plan of analysis

Study 1a addressed the first two research questions of this paper. The first question concerns the age of onset of children’s appreciation of the aspectuality of intentions based on an agent’s mis-representation. To answer this question, we conducted an analysis of variance (ANOVA). In a second step, we compared each age group’s performance against chance performance. The second question is how children’s understanding of mis-representation and their understanding of the aspectuality of intentions relate to each other. Here, we compared the intention ascriptions of children who had shown a general understanding of mis-representation to those who had not. Next, we tested whether the relation of age and intention ascriptions were different in these two groups. In the following, we report the analyses as a function of the question they mainly speak to.

Age of onset

We conducted a 4 (Age) x 2 (Test Question) x 2 (Scenario) ANOVA with number of correct trials as dependent variable. We included all interactions and main effects in our model. Fig

2. depicts the mean number of trials in which children gave correct answers as a function of age and test question.

Because of the ordinal data level of the dependent variable (number of correct trials) and because Shapiro-Wilk normality test showed that data were not normally distributed ($W=0.69$, $p<.001$), we conducted the ANOVA on aligned rank transformed data using the ARTool (Kay & Wobbrock, 2015). This procedure allows robust non-parametric analyses of interaction and main effects (Wobbrock et al., 2011). Results revealed an interaction effect between age and test question ($F(3, 186)=3.93$, $p=.009$, $\eta^2=.04^1$) and a main effect for test question ($F(1, 186)=22.22$, $p<.001$, $\eta^2=.08$). There were no other significant effects (all F s <3.64 , all p s $>.058$; see Appendix C for details). Post-hoc Kruskal Wallis tests on the Age*Test Question interaction effect showed a main effect of age for the belief test question ($\chi^2(3)=8.57$, $p=.036$, $\varepsilon^2=0.13$) but not for the intention test question ($\chi^2(3)=6.02$, $p=.111$, $\varepsilon^2=0.09$).

Thus, in general children performed worse on the intention than on the belief test questions. Overall, children's age did not influence their performance in the intention test question. However, older children gave more correct answers to the belief test question than younger children. There was no such effect for the intention test question. The type of scenario, Two-Objects or One-Object, had no influence on children's performance.

¹ These are effect sizes based on non-transformed data. The ARTool package does not cover effect sizes, yet. However, effect sizes based on non-transformed data approximate the effect sizes based on aligned rank transformed data (Kay, 2020).

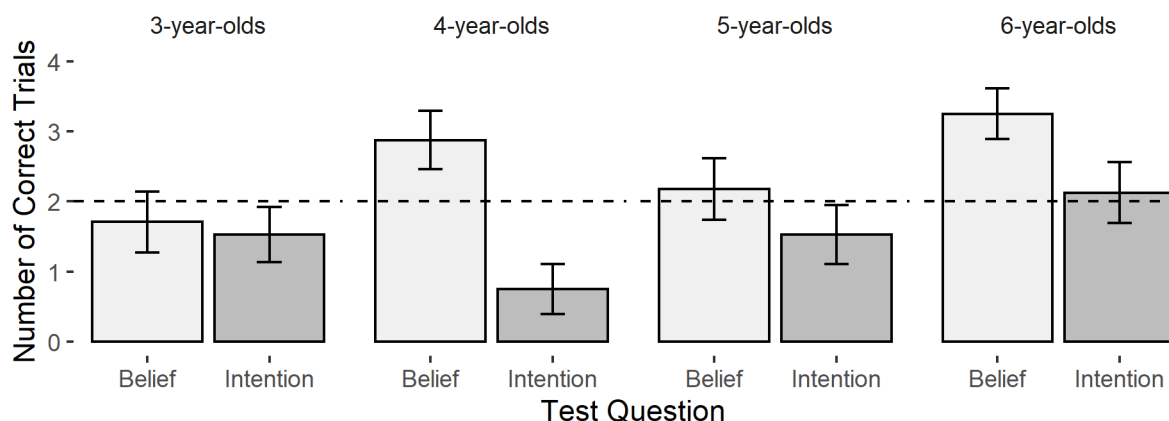


Fig. 2. Mean number of correctly answered trials out of four. The dashed line depicts chance performance. Error bars depict ± 1 Standard Error.

Children's performance on the intention test question did not improve with age. But was any age group able to ascribe the intentions correctly? We tested each age group's performance on the intention test question against chance performance (two out of four trials). In accordance with data level and distribution, we conducted one-sample Wilcoxon signed-rank tests. Results indicated that children of no age group performed significantly above chance. 4-year-olds' performance was even below chance (3-, 5- and 6-year-olds: all $V_s \geq 32$, all $p_s > .264$; 4-year-olds: $V=15$, $p=.008$). Thus, none of the included age groups showed proficient performance on the intention test questions. Performance on the belief test question was above chance for the 6-year-olds ($V=90$, $p=.008$) and at chance for younger age groups (all $V_s \geq 50$, all $p_s > .057$).

Thus, regarding our first question (age of onset of appreciating the aspectuality of intentions), Study 1a did not yield a clear answer. None of the age groups of 3- to 6-year-olds ascribed intentions proficiently in the spontaneous format. Furthermore, performance did not appear to progress within this age range.

Developmental Trajectory

We compared two plausible options: Understanding mis-representation is only necessary but not sufficient, or necessary and sufficient for understanding the aspectuality of intentions. For this reason, we compared the performance on the intention test question of children who showed a reliable understanding of mis-representation to those who did not. As criterion for the reliability of understanding mis-representation we used children's performance on the

belief test question. Children who had given correct answers to all four belief test questions were taken to have already obtained a reliable belief understanding (belief passers, $N=33$). In contrast, we attributed an unreliable belief understanding to children who failed on one or more belief test questions (belief non-passers, $N=33$). We then compared children's performance on the intention test question between belief passers and belief non-passers via a non-parametric Mann–Whitney U test. We found no differences in children's intention understanding depending on their belief understanding: Belief passers' ($M=1.36$, $SD=1.71$) and belief non-passers' ($M=1.61$, $SD=1.66$) performance on the intention test question did not differ significantly ($W=596$, $p=.483$). Thus, children's understanding of mis-representation alone did not determine whether they could appreciate the aspectuality of intentions. The former may thus be necessary but is not sufficient for the latter.

At first glance, it may appear surprising that children who are capable of ascribing beliefs in principle, have such a hard time applying this conceptual capacity in their evaluation of an action as (un-)intentional. However, such developmental trajectories are common in many areas: A given conceptual capacity (like understanding mis-representation) is of substantial importance for the acquisition of some related capacity (like intentionality judgments) but does not do the job alone. Other capacities are required in addition that develop only with progressing age. So, what would be evidence for such a necessary-but-not-sufficient pattern in the present case? One indicator could be different relations with age. The logic is the following: If the developmental trajectory follows indeed a necessary but not sufficient pattern, an appreciation of aspectuality of intentions should not come into play immediately after an understanding of mis-representation. Understanding mis-representations is the basis. But other additional factors have to come into place over time before children can consider mis-representations. Thus, only with increasing age should children (developing additional factors) become more and more able to use their insights about mis-representations for appreciating the aspectuality of intentions. Accordingly, only for those children who understand mis-representations (belief passers) we should observe that mastery of the intention question increases with age. In contrast, the progressing development of these additional factors should have no impact on the mastery of the aspectual intention question of children who do not yet understand mis-representations (belief non-passers), since they fail to fulfill the central necessary prerequisite. Thus, for non-passers there should be no relation between age and intention ascription.

Therefore, we looked at the correlation of age and performance on the intention test question separately for belief passers and belief non-passers. Again, we employed a non-parametric method and conducted Spearman's correlations on children's age in months and the aggregated score of their performance on the intention test question (see Fig 3.). The correlation did only reach significance for the belief passers ($r_s(33)=.46, p=.007$) but not for the belief non-passers ($r_s(33)=-.12, p=.500$). Thus, only for those children who already had obtained a reliable understanding of mis-representation, we found that the older they were the better they performed in response to the intention test question.

To gain further insight in the nature of this relation, we tested whether verbal intelligence had any influence on this relation. Stepwise multiple regressions showed that verbal intelligence did not moderate the relation of age and intention understanding for the "Belief Understanding" group ($\Delta R^2=0.05, \Delta F(1, 29)=1.96, p=.173., b=0.01, t(29)=1.40, p=.173$).

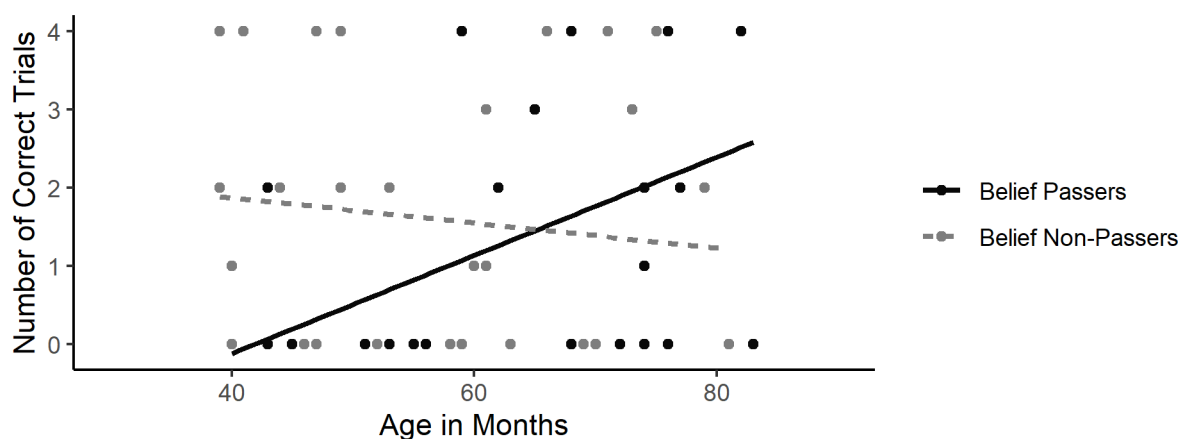


Fig. 3. Separate correlations between age and performance on the intention test question for belief non-passers and non-passers.

Thus, regarding our second question, i.e., the developmental trajectory of understanding mis-representation and understanding of the aspectuality of intentions, Study 1a revealed an interesting pattern. Understanding mis-representation alone did not determine children's ability to consider mis-representations when ascribing intentions. However, only for children that had obtained a reliable understanding of mis-representation, intention ascriptions improved with age. Thus, it seems that an understanding of mis-representation is the necessary first step in development. But understanding mis-representation alone is not

sufficient. Other factors need to come in place over time as indicated by the better performance of older children. And then, these other factors seem to enable children to apply their knowledge about an agent's mis-representation when they ascribe intentions to her. But what exactly are these factors? This brings on our third question that we address in Studies 1b and with a different approach in Study 2: If understanding mis-representations is indeed necessary but not sufficient for appreciating the aspectuality of intentions, what additional capacities are crucial, and under which conditions can children show such an appreciation?

Study 1b

Study 1b addressed the open questions of Study 1a. We administered the same spontaneous task as in Study 1a to test whether Study 1a's results supporting the necessary but not sufficient trajectory would replicate. To delineate the age of onset, we extended the age range to 3- to 9-year-olds. To address the third question concerning the crucial additional capacities, we employed a correlational design: Children received established tasks which measured their verbal intelligence, working memory and inhibitory control. To see whether children's difficulties relate to the task's recursive complexity we also assessed children ability to ascribe 2nd-order beliefs. The samples of 8- and 9-year-olds were tested at a later stage and due to the COVID-19 pandemic, data collection needed to be conducted online. This restricted us to the main task. Verbal intelligence, working memory and inhibitory control tasks could not be administered for these age groups.

Method

Participants

Seventeen 3-year-olds, seventeen 4-year-olds, seventeen 5-year-olds, seventeen 6-year-olds, seventeen 7-year-olds (36-95 months, $M=66.26$ months, $SD=16.96$; 35 male) were recruited from the same databank as in Study 1a. Children who had participated in Study 1a did not participate again. Four further children were tested but excluded from analysis because they were uncooperative ($N=3$) or due to experimental errors ($N=1$). In addition, seventeen 8-year-olds and seventeen 9-year-olds (96-118 months, $M=106.59$ months, $SD=7.03$; 15 male) participated in an online version of the main task. One further 9-year-old was tested but could not be included in the analyses because of technical issues.

Design and Procedure

A 5 (Age group: 3-, 4-, 5-, 6- or 7-year-olds) x 2 (Scenario: Two-Objects or One-Object) x 2 (Test Question: Belief or Intention) mixed design was conducted, with Test Question and Scenario as within-subjects factors. Again, children received eight trials in counterbalanced order, two per combination of Test Question and Scenario. The 2nd-order false belief test question was asked subsequent to the belief test question. As in Study 1a, the verbal intelligence task was always conducted at the beginning of each session. We counterbalanced whether the working memory and inhibitory control tasks were measured before or after the belief and intention tasks (see Appendix B for details).

In addition, we tested 8- and 9-year-olds¹. These data were only collected later as a consequence of the poor performance of the age groups originally included. Since not even 7-year-olds' performance on the intention test question exceeded chance performance, we extended the age window to 8- and 9-year-olds. Due to the COVID pandemic-related testing restrictions, these children could only be tested online. This allowed us to administer the main task and the 2nd-order belief question. The measures for linguistic and cognitive capacities could not be adapted to an online version. For our online version, we adapted the material of the adults' version of the main task (see Appendix A). Children saw videos of the main task, which used the same material and followed the same procedure as the acted-out version younger children received.

Main Task

Scenarios and Test Questions. Children observed the same scenarios as in Study 1a and received the same belief and intention test questions. Unlike in Study 1a, we also asked a 2nd-order false belief test question (following Perner & Howes, 1992):

2nd-order Belief TQ: If we ask the protagonist: Do you know where A is? What will she say? Will she say, “Yes, I know that“ or “No, I don't know that”?

Even though the protagonist mis-represents A's location, she believes she represents it accurately. Thus, the correct answer was “Yes”. Note, that we only administered the 2nd-order belief test question to children who had mastered the respective 1st-order question. This was

¹ We thank the anonymous reviewer for bringing up the idea of extending the age range even further.

possible as the 2nd-order test question was asked within the same trial as the 1st-order question. The reason behind this was to avoid false positives: If a child was unable to ascribe the belief to the agent she cannot ascribe a belief about this belief. Accordingly, a correct answer in such a case would display chance and not competence.

Linguistic and Cognitive Capacities

In addition to verbal intelligence, we tested for children's working memory and inhibitory control. A detailed description of each task and information regarding psychometric principles can be found in Appendix B.

Verbal Intelligence. To assess verbal intelligence we again administered the vocabulary subtest of the KABC.

Working Memory. To tap children's working memory, we conducted the Color Span Backwards task (Zoelch et al., 2005). This task requires children to remember a sequence of visually presented colors and to then reproduce these colors in the opposite order.

Inhibitory Control. As a measure of inhibitory control, we used the Head-Toes-Knees-Shoulders task (Ponitz et al., 2008), in which children are instructed to respond to the experimenter's requests in the opposite way. Hence, when the experimenter tells the child to touch her head, she has to inhibit abiding by this request and to touch her toes instead.

Results and Discussion

Scoring

The 2nd-order belief test question was scored as correct if children answered "Yes" (stating that the agent falsely believes her belief about the object's location to be true). Remember, we only administered the 2nd-order question when children had answered the respective 1st-order question correctly. All other 2nd-order belief trials were scored as missing values in the main analyses (though, see below for an alternative, more conservative analysis in which children who failed the 1st-order question were scored as failing the 2nd-order question as well).

For working memory children could score between 0 and 24 according to the number of correctly revised color sequences. For inhibitory control they could score between 0 and 60

according to the number of movements they transformed correctly (see Appendix B for a more detailed description of the scoring and range of obtained scores).

Plan of analysis

In Study 1b, we addressed all three questions. We first analyzed the data of the 3- to 7-year-olds who had received the acted-out versions of the task and the measures for linguistic and cognitive capacities. As in Study 1a, to test for the age of onset of an appreciation of aspectuality of intentions we conducted an ANOVA and compared each age group's performance against chance performance. To test for the developmental trajectory, like in Study 1a, we compared the performance on the intention test question and its relation to age between belief passers and non-passers. The overlap of task, age groups and analyses of Studies 1a and b allowed us to conduct these analyses on the merged sample of both studies' 3- to 6-year-olds. In addition, Study 1b addressed the question what other capacities are crucial in addition to understanding mis-representation. For this reason, we looked more closely at the relation between age and intention ascriptions for belief passers. We tested whether verbal intelligence, working memory or inhibitory control mediated this relation. Regarding children's 2nd-order belief reasoning, we asked whether the appreciation of the aspectuality of intentions is similarly impeded by processes as recursive complexity. For this reason, we tested for the correlation of the two tasks.

We collected supplementary data of 8- and 9-year-olds online in order to find out when children finally become able to consider the aspectuality of intentions in the spontaneous format of Study 1's task. Accordingly, we tested both age groups' performance against chance¹.

Age of Onset

As in Study 1a, we conducted a 5 (Age) x 2 (Test Question) x 2 (Scenario) ANOVA with

¹ We refrained from simply including these data in the analyses of 3- to 7-year-olds for two related reasons. First, the data were collected in very different ways (online testing rather than live act-out) and were thus not closely comparable to those of the 3- to 7-year-olds. Second, one task in particular (the one-object task) seems to have been somewhat compromised in its validity in the online testing format (see Appendix D for details).

number of correct trials as dependent variable. Again, we included all interaction and main effects. Fig 4. depicts the mean number of trials in which children gave correct answers as a function of age and test question.

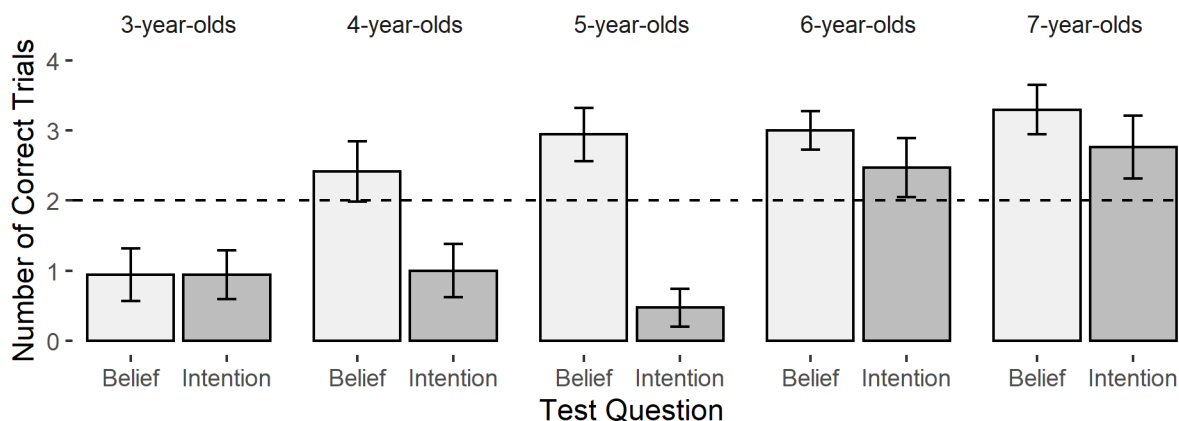


Fig. 4. Mean number of correctly answered trials out of four. The dashed line depicts chance performance. Error bars depict ± 1 Standard Error.

Again, our data were not normally distributed ($W=0.68, p<.001$). Accordingly, we conducted the ANOVA on aligned rank transformed data. This analysis yielded an interaction effect between age and test question ($F(4, 80)=9.36, p<.001, \eta^2=.00$) and a main effect for test question ($F(1, 80)=22.08, p<.001, \eta^2=.00$). In contrast to Study 1a, we also found a main effect for age ($F(4, 80)=9.85, p<.001, \eta^2=.02$). There were no other significant effects (all $F_s < 1.15$, all $p_s > .300$; see Appendix C for details). Post-hoc Kruskal Wallis tests on the Age*Test Question interaction effect showed a main effect of age for the belief test question ($\chi^2(3)=18.75, p<.001, \varepsilon^2=0.22$) and for the intention test question ($\chi^2(3)=20.98, p<.001, \varepsilon^2=0.25$).

Thus, as in Study 1a, children performed worse on the intention than on the belief test questions. However, in this study, we found that children's performance increased with age for the belief and the intention test question. Contingencies (see Table 1) show that this is most likely the direct result of including 7-year-olds. These were clearly divided in two groups with a majority of 65% 7-year-olds who solved all trials and a (still substantial) minority of 29% who failed all trials. The type of scenario, Two-Objects or One-Object, again had no influence on children's performance.

Table 1

Contingencies Age x Performance Intention Test Question

	0	1	2	3	4
3-year-olds	10	3	1	1	2
4-year-olds	11	1	2	0	3
5-year-olds	14	0	2	0	1
6-year-olds	4	2	1	2	8
7-year-olds	5	0	0	1	11

Note. Depicted is the number of correct intention test question trials.

In the next step, we again tested each age group's performance on the intention test question against chance via one-sample Wilcoxon signed rank tests. No age group performed significantly above chance. 3- to 5-year-olds' performance was even below chance (3-, 4- and 5-year-olds: all $Vs \leq 25.5$, all $ps < .033$; 6- and 7-year-olds: $V \geq 89$, $p > .134$). Thus, not even the 7-year-olds showed proficient performance on the intention test question. Performance on the belief test question was above chance for 5- to 7-year-olds (all $Vs \geq 81.5$, all $ps < .018$), at chance for 4-year-olds ($V = 65$, $p = .406$) and below chance for 3-year-olds ($V = 28.5$, $p = .028$).

Supplementary Online Data of 8- and 9-year-olds

The performance of the 8- and 9-year-olds in the complementary online study is depicted in Figure 5. The performance of both age groups on both the belief and intention test question exceeded chance performance (all $Vs \geq 75$, all $ps \leq .043$).

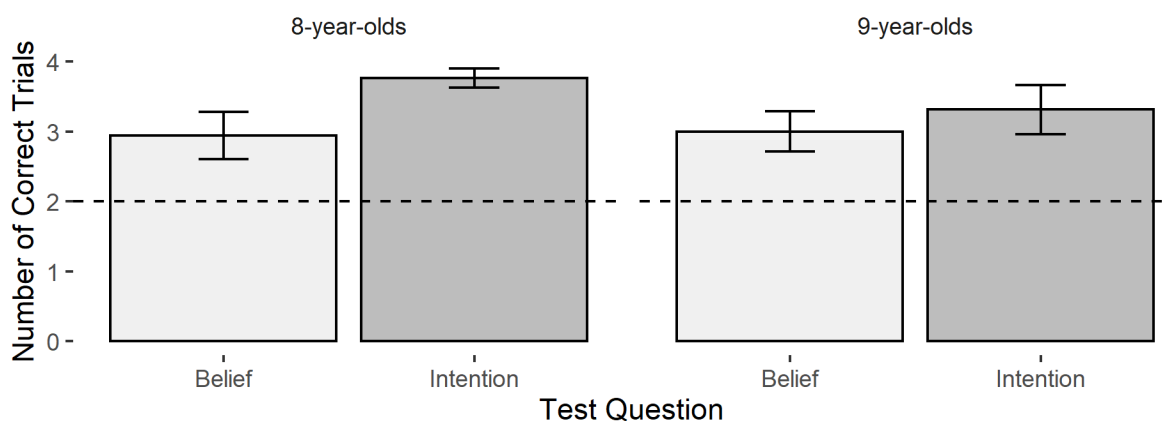


Fig 5. Mean number of correctly answered trials out of four. The dashed line depicts chance performance. Error bars depict ± 1 Standard Error.

Developmental Trajectory

As before, we split the 3- to 7-year-olds according to their understanding of misrepresentation in belief passers ($N=42$) and belief non-passers ($N=43$) and compared their performance on the intention test question via Wilcoxon rank sum tests. Belief passers ($M=1.95$, $SD=1.86$) performed significantly higher than non-passers ($M=1.12$, $SD=1.62$; $W=689.5$, $p=.040$, $r=0.22$). However, this difference appears to be the consequence of including older children than in Study 1b. As in Study 1a, there was no difference between belief passers' ($M=1.45$, $SD=1.05$) and non-passers' performance in the intentions test question ($M=1.70$, $SD=1.59$) for 3- to 6-year-olds ($W=689.5$, $p=.040$, $r=0.11$).

In the next step, we looked at the relation of age and intention ascriptions for belief passers and non-passers via Spearman's correlations on children's age in months and the aggregated score of their performance on the intention test question (see Fig 6.). Again, the correlation did only reach significance for the belief passers ($r_s(42)=.53$, $p<.001$) but not for the belief non-passers ($r_s(43)=.21$, $p=.171$).

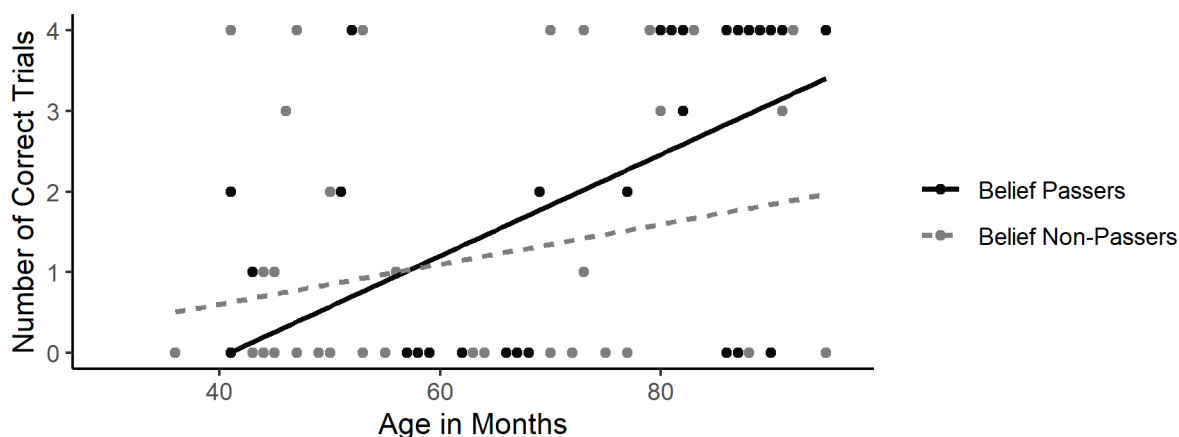


Fig. 6. Separate correlations between age and performance on the intention test question for belief non-passers and non-passers.

Analyses on merged sample of Study 1a and b

We administered the same basic task in Study 1a and b. The age windows of both studies overlap in that both samples include 3- to 6-year-olds. To increase test-power we merged both samples and conducted the analyses concerning age of onset and developmental trajectory on this merged sample. We found similar results. The merged sample consisted of thirty-four 3-, thirty-three 4-, thirty-four 5- and thirty-three 6-year-olds. The 4 (Age) x 2 (Test Question) x 2 (Scenario) ANOVA revealed a significant interaction between Age and Test Question ($F(3, 390)=8.97, p<.001$), a main effect for Test Question ($F(1, 390)=42.41, p<.001$) and age ($F(3, 130)=9.99, p<.001$). No other effect reached significance. One-sample Wilcoxon signed rank tests showed that also in our merged sample no age-group performed above chance for the intention test question. 6-year-olds' performance was at chance ($V=262.5, p<.305$). All younger age-groups' performance was even below chance (all $Vs \geq 95$, all $ps \leq .015$). For our merged sample of 3- to 6-year-olds we did not find a difference between belief passers' and non-passers' performance intention test question ($W=2276, p=.832$). The separate correlations between age and performance on the intention test question revealed such a correlation for belief passers ($r_s(62)=.44, p<.001$) but not for non-passers ($r_s(72)=0.03, p=.808$). Thus, not only did the results regarding our first two questions replicate. Also, the joined analyses of Study 1a and b's overlapping age groups and tasks found similar results. This indicates that results are also stable over different sample sizes and resulting test power.

Additional factors

To address the question, what other capacities are crucial, we first looked more closely at the relation between age and performance on the intention test question for those children who already showed a reliable understanding of mis-representation. We conducted Sobel tests to see whether children's scores for verbal intelligence, working memory or inhibitory control mediated the relation between age and intention ascription. None of these tests was significant (all $|z|s < 0.71$, all $ps \geq .481$). Thus, none of these capacities mediated this relation.

Next, we looked at children's performance on the intention test question and the 2nd-order belief test question. Both tasks require similar processing steps: First, ascribe the false belief and then, based on that insight, compute the 2nd-order belief or intention. If it was this similar recursive complexity that makes both tasks complicated we should find that they are correlated. We tested for the correlation of these two tasks via multiple regressions. This relation was significant ($b=1.12$, $t=2.39$, $p=.020$). However, as soon as age was entered as a second predictor the relation did not reach significance anymore ($b=0.51$, $t=1.03$, $p=.306$). The model including age also had a better fit ($F(1, 61)=13.77$, $p<.001$)¹. Also for the 8- and 9-year-olds tested online, intention and 2nd-order belief ascription was not related ($b=0.28$, $t=0.53$, $p=.600$).

Discussion

Regarding our first question, the age of onset of appreciating the aspectuality of intentions, Study 1 showed that only by the age of eight children were able to correctly ascribe intentions spontaneously. Study 1a showed that in the age window of three to six years children became

¹ Remember, if a child did not answer the 1st-order question correctly we did not administer the 2nd-order question, and such cases were coded as missing values. Alternatively, however, one could argue that answers in such cases can only be wrong (logically the 2nd-order question can only be correctly answered if the 1st-order question has been answered correctly), and that these cases should thus be coded as "incorrect" with regard to the 2nd-order question. An analysis based on this alternative coding scheme yielded a similar pattern of results. The relation of 2nd-order belief and intention ascriptions was significant ($b=1.56$, $t=3.68$, $p<.001$), but not if age was included as second predictor ($b=0.89$, $t=1.96$, $p=.053$). Again, the latter model explained significantly more variance ($F(1, 82)=10.31$, $p=.002$).

better in understanding mis-representations but not in considering these when ascribing intentions. When 7-year-olds were included, in Study 1b, we found children to become better in ascribing intentions over age. Still, a substantial subgroup of the 7-year-olds (five of seventeen children) ascribed intentions incorrectly in all trials. It appears that the spontaneous appreciation of the aspectuality of intentions develops in rather protracted ways.

Regarding our second question, the developmental trajectory of understanding mis-representation and appreciating the aspectuality of intentions, Study 1 clearly points towards a necessary but not sufficient pattern of development. Study 1a and b as well as the joint analyses found that only children who had already obtained an understanding of mis-representation, showed an increase over age in ascribing intentions correctly. No such relation was found for children who have not yet obtained an understanding of mis-representation. This indicates that once children have developed an understanding of mis-representation, some other capacity has to develop first, before children can consider the aspectuality of intentions.

Regarding our third question, what additional crucial factors are necessary for an appreciation of the aspectuality of intentions, we did not find conclusive evidence. Study 1b investigated the most obvious candidates, working memory, inhibitory control and verbal intelligence. None of them could explain what enables children to apply their understanding of mis-representation when ascribing intentions. Note, however, that we only had these data for 3- to 7-year-olds. None of these age groups performed proficiently in ascribing intentions. It is thus possible that these results would be different once sufficiently proficient performers were included that cause more variance in our data. Neither for 3- to 7-year-olds not for the 8- and 9-year-olds tested online, the demands of appreciating the aspectuality of intentions were related to the recursive nature of 2nd-order belief ascription. Thus, Study 1b did not find any evidence for any specific additional factors that were crucial for an appreciation of the aspectuality of intention.

But what then causes this “necessary but not sufficient”-trajectory? Another way to approach our third question is to look for conditions under which children show an appreciation of the aspectuality of intentions. One way to realize such an approach is by manipulating task demands, for example related to inferential complexity. The present intention ascription task has a very complex inferential structure. In the spontaneous format of Studies 1, children had to first infer the agent’s mis-representation (“She believes that A is still in box 1”). In a second step, they had to ascribe the intention on the basis of the agent’s mis-representation (“She intentionally took the box that in fact contained A. But since she believes that A is still in box

1, she did not intend to take the box with A.”). Previous empirical work in the domains of moral reasoning and trait ascriptions identified parallel inferential demands. Several studies found that guiding children through the inferential chain revealed much earlier competence (Liu et al., 2007; Proft & Rakoczy, 2018).

Study 2

The rationale of Study 2 was to test when children reveal competence in aspectual intention ascription tasks once the task demands have been radically reduced. Following similar work in other domains, we adapted the intention ascription task in Study 2 and scaffolded children through the requisite inferential chains. By leading the children in a step-wise manner through the components of this chain, working memory demands (regarding the amount of information children have to process simultaneously) were substantially reduced.

Method

Participants

Twenty-one 4-year-olds, eighteen 5-year-olds, twenty 6-year-olds and twenty 7-year-olds (48-96 months, $M=71.47$ months, $SD=14.14$; 37 male) were recruited from the same databank as in Study 1. Children who had participated in one of the earlier studies did not participate again. One further 4-year-old has been tested but was excluded from analysis because he was uncooperative.

Design and Procedure

A 4 (Age group: 4-, 5-, 6- or 7-year-olds) x 2 (Scenario: Two-Objects or One-Object) x 2 (Test Question: Belief or Intention) mixed design was conducted, with Test Question and Scenario as within-subjects factors. Children received four intention trials but only two belief trials, one per scenario. To be able to see the genuine impact of the scaffolding modification, belief trials were always administered after the intention trials. As in Study 1, the verbal intelligence task (KABC) was always conducted at the beginning of each session.

Main Task

Scenarios and Test Questions. Children observed the same scenarios as in Study 1. They also received the same test questions with one exception in the intention trials: Before we asked

the intention test question, we reminded them of the agent's mis-representation by asking them "Where does the protagonist believe that A is?" Incorrect answers to this question were corrected (and correct answers were confirmed in order to keep the amount of feedback constant irrespective of performance).

Results

Study 2 focused on the question under which conditions children display an appreciation of the aspectuality of intentions. For this reason, we conducted an ANOVA on their performance on the intention test questions and compared each age group's performance against chance performance. Next, we compared the intention ascriptions of belief passers and non-passers. To test for the impact of verbal intelligence, we conducted multiple regressions. Fig 7. depicts the mean number of trials in which children gave correct answers as a function of age and test question.

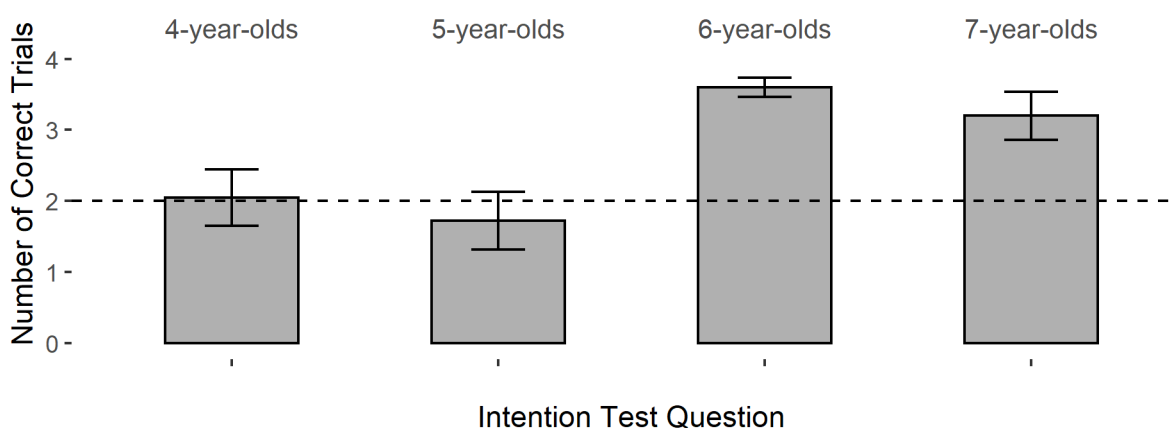


Fig. 7. Mean number of correctly answered intention trials out of four. The dashed line depicts chance performance. Error bars depict ± 1 Standard Error.

Again, our data were not normally distributed ($W=0.67$, $p<.001$). Accordingly, we conducted a 4 (Age) x 2 (Scenario) ANOVA with number of correct intention trials as dependent variable on aligned rank transformed data. As before, we included all interaction and main effects. This analysis yielded a main effect for age ($F(3, 75)=8.04$, $p<.001$, $\eta^2=.19$). There were no other significant effects (all $F_s < 1.44$, all $p_s > .237$; see Appendix C for details).

We tested each age group's performance on the intention test question against chance via one-sample Wilcoxon signed rank tests. 4- and 5-year-olds' performance did not exceed chance performance (all $V_s \leq 44$, all $p_s > .540$). In contrast, 6- and 7-year-olds performed

significantly above chance (all $V_s \geq 142.5$, all $p_s < .005$, all $r_s > .63$). Likewise, contingencies show that most 4- and 5-year-olds continuously failed all trials while most 6- and 7-year-olds solved all trials (see Table 2).

Table 2

Contingencies Age x Performance Intention Test Question

	0	1	2	3	4
4-year-olds	8	1	1	4	7
5-year-olds	8	0	3	3	4
6-year-olds	0	0	1	6	13
7-year-olds	3	0	2	0	15

Note. Depicted is the number of correct intention test question trials.

Next, we compared performance on the intention test question between belief passers ($N=41$) and belief non-passers ($N=38$). Belief passers ($M= 3.27$, $SD=1.32$) performed significantly more proficiently than non-passers ($M= 2.00$, $SD=1.72$; $W=1105$, $p < .001$)¹. Note however, that these findings may be somewhat difficult to interpret as the reminder in the scaffolded intention task was, in fact, a belief question (“Where does the protagonist believe that A is?”). Given that children received feedback on their answers to this question and that the false believe tasks were administered after the intention tasks (and thus after children had already received four trials of feedback on the reminder) their answers to the false belief task might be difficult to interpret due to potential learning effects.

Regarding the impact of verbal intelligence, multiple regression revealed that verbal intelligence and intention ascriptions were not related when controlling for age ($b=0.06$, $t=0.93$, $p=.354$).

¹ A similar pattern emerged when we determined children’s belief understanding based on their answers to the reminding question ($W=544$, $p=.015$).

Discussion

Study 2 focused on our third question (under what conditions do children display an appreciation of the aspectuality of intentions?) but also speaks to the other questions. In Study 2, we radically reduced tasks demands by scaffolding children through the required inferential chains. Now, even 6-year-olds performed proficiently in the intention ascription task.

This brings us back to the first question concerning the age of onset. This relatively early age of onset contrasts with children's low performance in the spontaneous format of Study 1. It seems that, indeed, the demands of the spontaneous format did obscure children's genuine competence; and that a basic form of appreciation of the aspectuality of intentions has emerged by the age of six. Regarding our second question, the results of Study 2 converge with those of Studies 1a and b: An understanding of mis-representations appears to be necessary but not sufficient for appreciating the aspectuality of intentions.

General Discussion

Summary of main findings

The present studies addressed three questions. First, what is the age of onset of children's appreciation of the aspectuality of intentions based on an agent's mis-representation? We found that only by the age of eight children spontaneously considered the agent's mis-representation in their intention ascription (Study 1). However, it appears that in these spontaneous ascriptions children's genuine competency was masked by extraneous performance factors. When we reduced these factors in Study 2, we found an age of onset around age six. The second question addressed the developmental trajectory. How is the development of the appreciation of aspectuality of intentions related to children's understanding of mis-representation in general? Study 1 found that while understanding mis-representation was necessary it was not sufficient. Following up on this, the third question was what additional capacities and conditions are necessary for children to show an appreciation of the aspectuality of intentions. Neither linguistic nor domain-general cognitive capacities explained the developmental trajectory (Study 1b). Only inferential complexity had a substantial impact on children's performance. When we guided children step-by-step through the inferential chain required to solve the test, 6-year-olds mastered the task proficiently.

Understanding aspectuality as a form of advanced Theory of Mind

These findings may seem surprising given children’s developing Theory of Mind competence around age four (or even earlier) (Wellman et al., 2001). But in fact, they do converge with results from many other studies in highlighting a more protracted development of complex forms of theory of mind than suggested by an overly narrow focus on false belief tasks. Similar patterns can be found in studies on ascribing complex emotions and higher-order mental states or appreciating the interpretive nature of representations (Harris et al., 1989; Perner & Wimmer, 1985; Liddle & Nettle, 2006; Osterhaus et al., 2016; Chandler & Lalonde, 1996). These studies, just like the present one, point towards a protracted developmental trajectory of advanced theory of mind that builds on, but goes beyond basic meta-representational theory of mind (such that the latter is necessary but not sufficient for the former).

But in what ways? What are the additional ingredients required to transform basic theory of mind into the advanced theory of mind of understanding the complex aspectuality of intentions? Currently, we do not know. Study 1b remained inconclusive and found no evidence for a role of linguistic or domain-general capacities – as measured here. Still, it is highly plausible that some form of complex domain-general and/or linguistic capacities are crucial and that Study 1b simply failed to tap them in the right kinds of ways¹. In broader theoretical perspective, plausible candidate capacities include the mastery of “relational complexity” (Halford et al., 1998), “cognitive complexity” (Zelazo et al., 1998) or general “recursive” capacities (Hauser et al., 2002). Future research needs to operationalize these rather abstractly described capacities and test for their role in the ontogenetic progression from basic to advanced theory of mind.

Competence versus performance

Still, before we interpret the present findings as indicating the relatively late emergence of a form of advanced theory of mind, one fundamental caveat needs to be discussed: Do the

¹ Remember, none of the age groups 3- to 7-year-olds ascribed intentions proficiently. Accordingly, it is possible that the necessary extent of linguistic and domain-general capacities exceeds the scope of our sample.

results of the present tasks really indicate a lack of competence (to understand the aspectuality of intentions) in children younger than age 6 or even later? Or alternatively, may children's actual competence have been masked by performance factors in these specific tasks? Indeed, Study 2 speaks in favor of the latter. Scaffolding children through the task enhanced performance of 6- and 7-year-olds. We thus cannot rule out that competence might already be present in even younger children. It might have simply been obscured by performance factors that go beyond those reduced in our guided format. One potential linguistic performance factor concerns the way children read the test question. Our test question requires the participant to apply a so-called *de dicto* (about what is said) reading (see Jacob, 2019; McKay & Nelson, 2014; Quine, 1956). Here, the truth of an intention report depends on whether the agent would report her intention in that particular way. This implies that the substitution of co-referential terms (*Walt/Heisenberg*) can affect the truth-value of the intention report. Take the intention report "Hank intends to convey secret information to Heisenberg". Hank would never have reported the intention in that way. Thus, on the *de dicto* level this report is plainly false. Yet, another way to approach the test question is to apply a *de re* (about the thing) reading. Here, the truth of the intention report does not depend on Hank's perspective and allows the substitution of co-referential terms. Thus, in our task the action is unintentional under the mis-represented description on a *de dicto* level. In contrast, it is intentional on a *de re* level irrespective of the description (Hank intended to do what was in fact the reporting of secrets to Heisenberg). Of course, we assumed that the most intuitive approach to our task would be the *de dicto* reading. And indeed, the results from adults show that for them, the *de dicto* reading is natural and obvious. That it is not equally obvious for children may mean that they do not understand the *de dicto* reading in principle. But alternatively, it may mean that they understand both *de re* and *de dicto* readings but have different thresholds for when they find one rather than the other obvious. Perhaps the *de dicto* reading in the present tasks was simply not sufficiently salient, relevant and obvious to them?

Whether or not Hank intended to tell secret information to Heisenberg is without doubt highly salient and relevant (it will change Hank's life, in fact). The same is true for many decisions. Should you be angry with your neighbor who during your vacation killed your orchid, because she mis-took it for a plastic replica? Should a child be angry at her grandmother because she gave her a pink note pad for Christmas which she thought was the desired "tablet"? Examples like these show how significant the consideration of aspectuality of intentions often is. In contrast, whether or not the hedgehog intended to take the pen in our task is by far not of

comparable real-life relevance. Now, the rationale behind choosing the present scenarios was to test for an understanding of the aspectuality of intentions in comparable vignettes and formats that conform to standard procedures. However, it is possible that in these scenarios it did not become clear to the children that they were required to judge the intentionality from the agent's perspective. And as our aim is not to find when children develop an adult-like approach to aspectuality that comprises irrelevant cases, but when they develop an understanding of intentions in general. Accordingly, before we categorize understanding the aspectuality of intentions as part of protracted Theory of Mind development, future research needs to address this alternative explanation. This requires a task that clearly and conclusively asks for a *de dicto* reading. One such case would be child-friendly versions of the Walt/Heisenberg-type in which the aspectuality of intentionality is practically and morally highly obvious, salient and relevant.

Conclusion

Taken together, the present studies found that children have substantial difficulties appreciating that actions are not intentional under mis-represented descriptions. Such an appreciation of the aspectuality of intentions appears to develop late (not before six) and to go substantially beyond the ability to ascribe mis-representations in general. Even though ascribing mis-representations seems to be necessary, it alone is not sufficient. All of these are features that the appreciation of the aspectuality of intentions shares with other capacities that belong to protracted theory of mind development. Yet, future research needs to follow up on this. First, to check whether we measured children's genuine competence (or whether their true competence was masked by performance factors of the specific tasks). And second, in case the picture of a necessary but not sufficient developmental trajectory persists, to identify the crucial additional capacities that are necessary in the progression from basic to advanced Theory of Mind.

Acknowledgements

We thank Christiane Decker, Maike Holland-Letz and Lynn Weiß for help with data collection and Marlen Kaufmann and Konstanze Schirmer for assistance and organization of data collection. The first author was funded by Evangelisches Studienwerk Villigst and Studienstiftung des Deutschen Volkes.

References

- Anscombe, G.E.M. (1957). *Intention*. Oxford: Basil Blackwell.
- Astington, J. W. (2001). The paradox of intention: Assessing children's metarepresentational understanding. *Intentions and intentionality: Foundations of social cognition*, 85-103.
- Behne, T., Carpenter, M., Call, J., & Tomasello, M. (2005). Unwilling versus unable: infants' understanding of intentional action. *Developmental psychology*, 41(2), 328.
doi:10.1037/0012-1649.41.2.328
- Chandler, M. J., & Lalonde, C. (1996). Shifting to an interpretive theory of mind: 5-to 7-year-olds' changing conceptions of mental life. *The five to seven year shift: The age of reason and responsibility*, 111-139. doi
- 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). doi: 10.1037/0012-1649.17.1.99
- Flavell, J. H., Green, F. L., Flavell, E. R., Watson, M. W., & Campione, J. C. (1986). Development of knowledge about the appearance-reality distinction. *Monographs of the society for research in child development*, i-87. doi:10.2307/1165866
- Flavell, J. H., Shipstead, S. G., & Croft, K. (1978). Young children's knowledge about visual perception: Hiding objects from others. *Child Development*. doi:10.2307/1128761
- Frye, D., David Zelazo, P., & Burack, J. A. (1998). Cognitive complexity and control: I. Theory of mind in typical and atypical development. *Current Directions in Psychological Science*, 7(4), 116-121.
- Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature*, 415(6873). doi:10.1038/415755a
- Hadwin, J., & Perner, J. (1991). Pleased and surprised: Children's cognitive theory of emotion. *British Journal of Developmental Psychology*, 9(2). doi:10.1111/j.2044-835X.1991.tb00872.x
- Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and brain sciences*, 21(6), 803-831.
doi:10.1017/S0140525X98001769
- Harris, P. L., Johnson, C. N., Hutton, D., Andrews, G., & Cooke, T. (1989). Young children's theory of mind and emotion. *Cognition & Emotion*, 3(4).
doi:10.1080/02699938908412713

- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve?. *science*, 298(5598), 1569-1579.
doi:[10.1126/science.298.5598.1569](https://doi.org/10.1126/science.298.5598.1569)
- Jacob, P. (2019). Challenging the Two-systems Model of Mindreading. *Knowing Other Minds*, 79.
- Kamawar, D., & Olson, D. R. (2011). Thinking about representations: The case of opaque contexts. *Journal of experimental child psychology*, 108(4).
doi:[10.1016/j.jecp.2010.10.005](https://doi.org/10.1016/j.jecp.2010.10.005)
- Kay, M. (2020, March 19). *Effect Sizes with ART*. The Comprehensive R Archive Network. Retrieved May 26, 2020, from <https://cran.r-project.org/web/packages/ARTool/vignettes/art-effect-size.html>
- Kay, M., & Wobbrock, J. O. (2020, March 20). *ARTool: Aligned Rank Transform*. The Comprehensive R Archive Network. Retrieved May 26, 2020, from <https://cran.r-project.org/web/packages/ARTool/>
- Lang, B., & Perner, J. (2002). Understanding of intention and false belief and the development of self-control. *British Journal of Developmental Psychology*, 20(1), 67-76. doi: 10.1348/026151002166325
- Liddle, B., & Nettle, D. (2006). Higher-order theory of mind and social competence in school-age children. *Journal of Cultural and Evolutionary Psychology*, 4(3-4), 231-244. doi:[10.1556/JCEP.4.2006.3/4.3](https://doi.org/10.1556/JCEP.4.2006.3/4.3)
- Liu, D., Gelman, S. A., & Wellman, H. M. (2007). Components of young children's trait understanding: Behavior-to-trait inferences and trait-to-behavior predictions. *Child development*, 78(5), 1543-1558. doi:[10.1111/j.1467-8624.2007.01082.x](https://doi.org/10.1111/j.1467-8624.2007.01082.x)
- Mahy, C. E., Bernstein, D. M., Gerrard, L. D., & Atance, C. M. (2017). Testing the validity of a continuous false belief task in 3-to 7-year-old children. *Journal of experimental child psychology*, 160, 50-66. doi: 10.1016/j.jecp.2017.03.010
- McKay, T., & Nelson, M. (2014). Propositional attitude reports (Spring 2014 Edition). In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from <http://www.plato.stanford.edu/archives/spr2014/entries/prop-attitude-reports/>
- Melchers, P., & Preuß, U. (2009). Kaufman assessment battery for children 8. Unveränderte Auflage. *Frankfurt/Main: Pearson Assessment*.

- Meltzoff, A. N. (1995). Understanding the intentions of others: re-enactment of intended acts by 18-month-old children. *Developmental psychology*, 31(5), 838. doi:[10.1037/0012-1649.31.5.838](https://doi.org/10.1037/0012-1649.31.5.838)
- Michel, E., & Roebbers, C. M. (2008). Children in regular and special needs classes: Cognitive and non-cognitive aspects. *Swiss Journal of Psychology/Schweizerische Zeitschrift für Psychologie/Revue Suisse de Psychologie*, 67(4), 249. doi: 10.1024/1421-0185.67.4.249
- Müller, U., Kerns, K. A., & Konkin, K. (2012). Test–retest reliability and practice effects of executive function tasks in preschool children. *The Clinical Neuropsychologist*, 26(2), 271-287. doi:[10.1080/13854046.2011.645558](https://doi.org/10.1080/13854046.2011.645558)
- Osterhaus, C., Koerber, S., & Sodian, B. (2016). Scaling of advanced theory-of-mind tasks. *Child development*, 87(6), 1971-1991. doi:[10.1111/cdev.12566](https://doi.org/10.1111/cdev.12566)
- Perner, J., & Howes, D. (1992). " He thinks he knows": And more developmental evidence against the simulation (role taking) theory. *Mind & Language*. doi:[10.1111/j.1468-0017.1992.tb00197.x](https://doi.org/10.1111/j.1468-0017.1992.tb00197.x)
- Perner, J., Lang, B., & Kloo, D. (2002). Theory of mind and self-control: More than a common problem of inhibition. *Child development*, 73(3), 752-767. doi: [10.1046/j.1467-8624.2003.00640.x](https://doi.org/10.1046/j.1467-8624.2003.00640.x)
- 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.
- Ponitz, C. E. C., McClelland, M. M., Jewkes, A. M., Connor, C. M., Farris, C. L., & Morrison, F. J. (2008). Touch your toes! Developing a direct measure of behavioral regulation in early childhood. *Early Childhood Research Quarterly*, 23(2), 141-158. doi:[10.1016/j.ecresq.2007.01.004](https://doi.org/10.1016/j.ecresq.2007.01.004)
- Proft, M., & Rakoczy, H. (2019). The ontogeny of intent-based normative judgments. *Developmental science*, 22(2), e12728. doi:[10.1111/desc.12728](https://doi.org/10.1111/desc.12728)
- Proft, M., Schünemann, B., & Rakoczy, H. (2019). Children's understanding of the aspectuality of intentions. *Journal of experimental child psychology*, 181. doi: [10.1016/j.jecp.2018.12.001](https://doi.org/10.1016/j.jecp.2018.12.001)
- Quine, W. V. (1956). Quantifiers and propositional attitudes. *the Journal of Philosophy*, 53(5), 177-187. doi:[10.2307/2022451](https://doi.org/10.2307/2022451)

- Rakoczy, H., Bergfeld, D., Schwarz, I., & Fizke, E. (2015). Explicit theory of mind is even more unified than previously assumed: Belief ascription and understanding aspectuality emerge together in development. *Child Development, 86*(2). doi:[10.1111/cdev.12311](https://doi.org/10.1111/cdev.12311)
- Remijn, G. B., Kikuchi, M., Yoshimura, Y., Shitamichi, K., Ueno, S., Tsubokawa, T., ... & Minabe, Y. (2017). A near-infrared spectroscopy study on cortical hemodynamic responses to normal and whispered speech in 3-to 7-year-old children. *Journal of Speech, Language, and Hearing Research, 60*(2), 465-470. doi: 10.1044/2016_JSLHR-H-15-0435
- Schult, C. A. (2002). Children's understanding of the distinction between intentions and desires. *Child Development, 73*(6). doi:10.1111/1467-8624.t01-1-00502
- Searle, John R., and S. Willis. *Intentionality: An essay in the philosophy of mind*. Cambridge university press, 1983.
- Shultz, T. R., & Wells, D. (1985). Judging the intentionality of action-outcomes. *Developmental Psychology, 21*(1), 83.
- Sullivan, K., Zaitchik, D., & Tager-Flusberg, H. (1994). Preschoolers can attribute second-order beliefs. *Developmental Psychology, 30*(3), 395. doi:10.1037/0012-1649.30.3.395
- 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. doi:[10.1111/1467-8624.00304](https://doi.org/10.1111/1467-8624.00304)
- 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). doi:[10.1016/0010-0277\(83\)90004-5](https://doi.org/10.1016/0010-0277(83)90004-5)
- Wobbrock, J. O., Findlater, L., Gergle, D., & Higgins, J. J. (2011, May). The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 143-146). doi:10.1145/1978942.1978963
- Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69*(1), 1-34. doi:[10.1016/S0010-0277\(98\)00058-4](https://doi.org/10.1016/S0010-0277(98)00058-4)
- Zoelch, C., Seitz, K., & Schumann-Hengsteler, R. (2005). From rag (bag) s to riches: Measuring the developing central executive. *Young children's cognitive development*.

Interrelationships among executive functioning, working memory, verbal ability and theory of mind, 39-69.

Appendix A

Paper Pencil Version for Adults

To validate the task and ensure that our task analysis fits with mature folk psychology, we administered a paper pencil version of the task to adult participants. Adults watched presentations of the scenarios. They answered on questionnaires in multiple choice-format.

Method

Participants. Twenty-four adults (36-95 months, $M=66.26$ months, $SD=16.96$; 35 male) were recruited on Campus.

Design and Procedure. A 2 (Scenario: Two-Objects or One-Object) x 2 (Test Question: Belief or Intention) within-subjects design was conducted. Participants received eight trials in counterbalanced order, two per combination of Test Question and Scenario. As in Study 1b, the 2nd-order false belief test question was asked subsequent to the belief test question. Note that due to the paper pencil version all participants received the 2nd-order belief test questions. Yet, only when the respective 1st-order question was answered correctly we included the answer in our analyses.

Scenarios and Test Questions. Participants received the same scenarios as in Study 1. However, in the adult-version these were not acted out. Adults watched animated PowerPoint-presentations. The verbal descriptions of the plot were given as written descriptions in the presentations. The same control and test question as in Study 1 were asked as part of the presentation. Participants answered to these questions on a printed out questionnaire. The multiple choice-format of this questionnaire offered as answer opportunities the two boxes (e.g. “The pink box.” vs “The green box.”). For the control question we offered the choice between “Yes.” And “No.” and for the 2nd-order belief test question “Yes, I know that. “ or „No, I don’t know that.”.

Results and Discussion. Participants performed significantly above chance for all test questions (see Fig. A1; all $V_s=0$, all $p_s<.018$). All participants judged the actions under the mis-represented description as unintentional. This supports that our task analysis indeed fits with mature folk psychology.

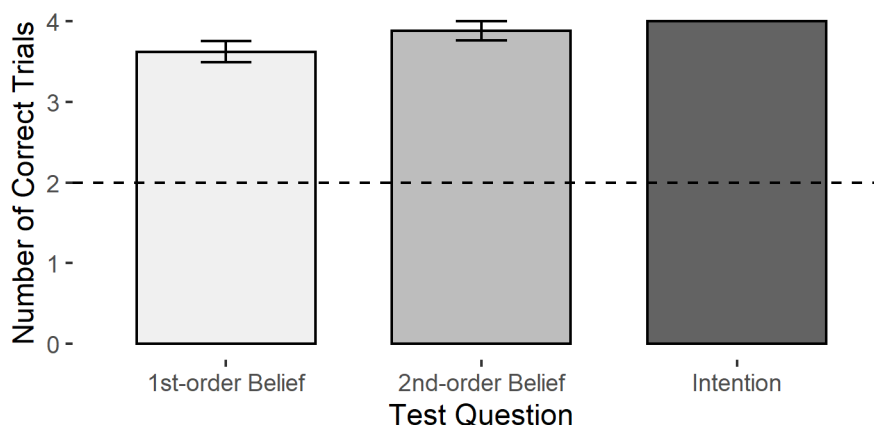


Fig. A1. Mean number of correctly answered trials out of four. The dashed line depicts chance performance. Error bars depict ± 1 Standard Error.

Appendix B

Measures Linguistic and Cognitive Capacities

Vocabulary Subtest of the Kaufman Assessment Battery for Children II (KABC)

The vocabulary subtest of the KABC (Melchers & Preuß, 2009) consists of 39 different photographs of objects that are increasingly unlikely to be part of the child’s vocabulary (ranging from item 1 “Dog” to item 39 “Septum”). The experimenter shows the photographs to the child in a fixed order of increasing difficulty. For each presented photograph the experimenter asks “What is this?”. If the child’s answer is unambiguously correct or wrong the experimenter proceeds to the next item. For answers that are only close to the correct answer, the KABC offers a list of possible answers that have to be followed by pursuing questions as well as a list of answers that have to be coded as wrong directly. For example, for the item “Guitar” the answer “Instrument” requires a pursuing question, while the answer “Violin” has to be coded as wrong. As soon as the child gives for incorrect answers in a row, the session is terminated. The number of correct items gives the score. The objectivity of the KABC fulfills the TBS-TK¹ requirements to complete satisfaction. Reliability and validity requirements are largely fulfilled (Rollett & Preckel, 2011). The vocabulary subtest of the KABC has been administered in other studies to children of similar age-ranges (e.g., Perner

¹ Test Review Model of the German Federation of National Psychology Associations

et al., 2002; Remijn et al., 2017).

Color Span Backwards task (CSB)

The CSB (Schmid, Zoelch, & Roebbers, 2008) follows the same logic as the widely used digit span backwards task (Pickering & Gathercole, 2001) and requires the child to remember a sequence of visually presented colored disks and to reproduce them in the inverted order. Yet in contrast to the digit span backwards task, it avoids the impact a diverging acquaintance with digits can have. The original CSB-procedure tries to further enhance plausibility by introducing a dwarf who loses colors out of his bag. During piloting phase, we found that especially young children had difficulties to process this cover story in addition to the task itself. This might have been due to the length and high number of tasks of Study 1b. For this reason, we decided not to include the cover story in our study. Apart from that, we exactly followed the instructions of the CSB.

The experimenter explains that when she shows to the child the color yellow and the color green the child's task is to reproduce them the other way round, i.e. Y and X. This underlined with two cards showing a yellow and a green colored disc that are presented and then reversed. Then the procedure is continued on a screen on which the colors appear successively for the duration of 1s. The child receives three practice trials. If she fails a trial the experimenter will use cards again to illustrate the correct solution. If a child fails two or more practice trials the practice phase will be repeated. If a child also fails two or more trials of the repeated training phase the task will be aborted. After a successful practice phase the first six sequences with two colors are presented to the child. If the child correctly reverses at least four of the six sequences, the number of discs for the next six sequences will be increased by one. Otherwise, the test will be aborted. The score is the number of correctly reversed sequences. The reliability and validity of the CSB are appropriate (Schmid, et al., 2008). The CSB has been administered in other studies to children of similar age-ranges (e.g., Müller et al., 2012; Michel & Roebbers, 2008).

Head-Toes-Knees-Shoulders task (HTKS)

The HTKS (Ponitz et al., 2008) begins with an introduction phase in which the child is introduced to touching her head/toes upon an according oral command ("Touch your head/toes."). Next, the experimenter tells her to engage in a rather goofy activity and to react to each command in the opposite way: hence, to touch her toes when asked to touch her head

and vice versa. After four corrected training trials, the child receives ten test trials. If children react correctly to at least four of these ten trials they will learn about a second paired behavioral command (knees-shoulders). This new command is trained mixed with the first command in four more training trials, which are followed by ten more test trials (both paired commands). Again, if the child masters four of these test trials she will receive the third part of the HTKS in which the pairs are reversed (head-knees, shoulders –toes). In this third part there are no training trials. Instead, the 10 test trials start right after the introduction. Each mastered test trial earns 2 points. Test trials, in which the child first indicates a wrong reaction but the corrects herself, earn 1 point. Incorrect test trials earn 0 points. The sum of points indicated the child HTKS-score. The task has been found to be a valid and reliable measurement (McClelland et al., 2014). The HTKS has been administered in other studies to children of similar age-ranges (e.g., Mahy et al., 2017).

Ranges of Scores of the Measures obtained in Study 1

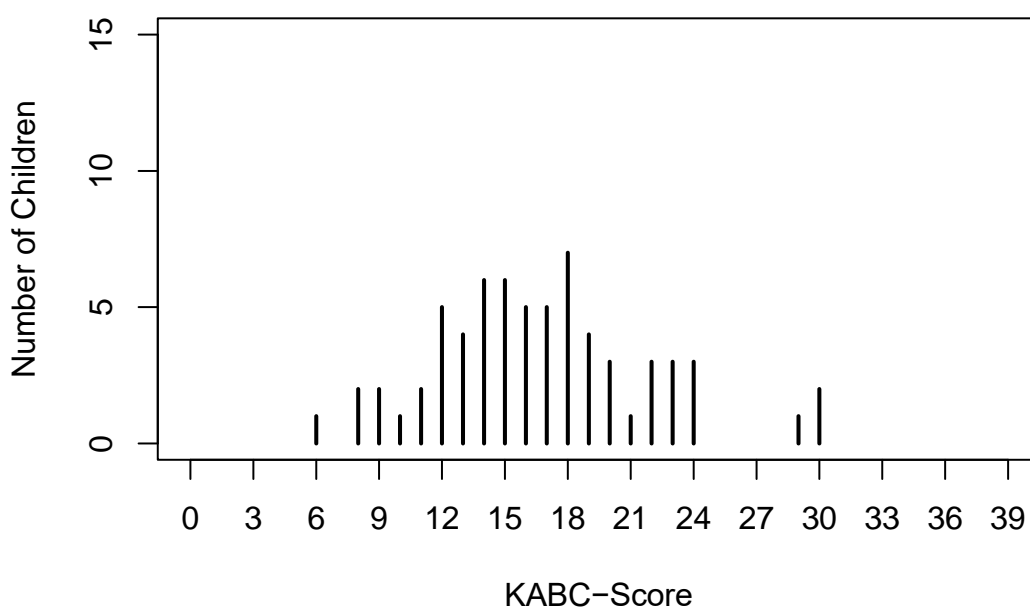


Fig. B1. Number of children per verbal intelligence score in Study 1a

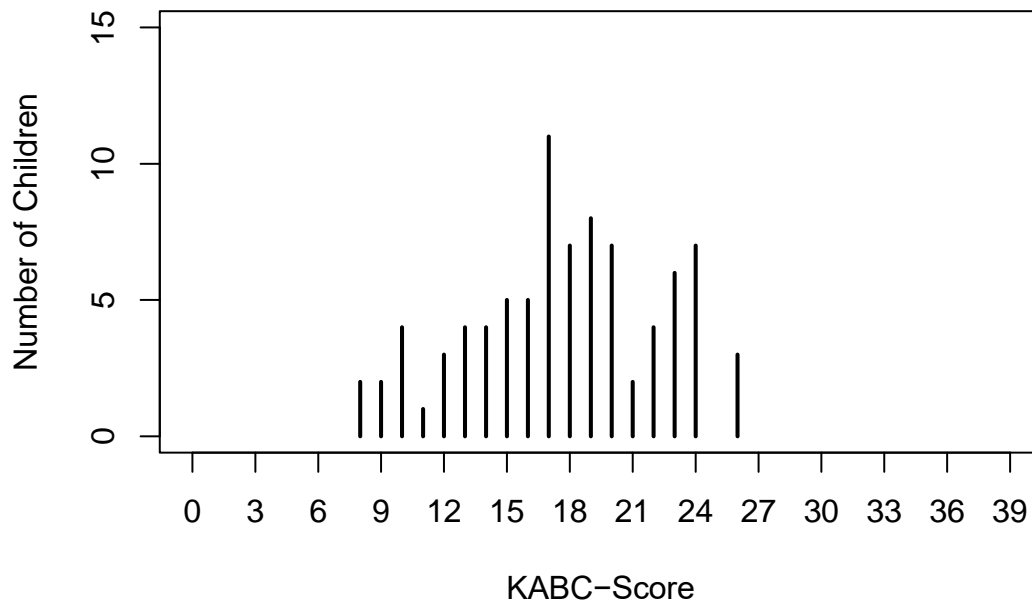


Fig. B2. Number of children per verbal intelligence score in Study 1b

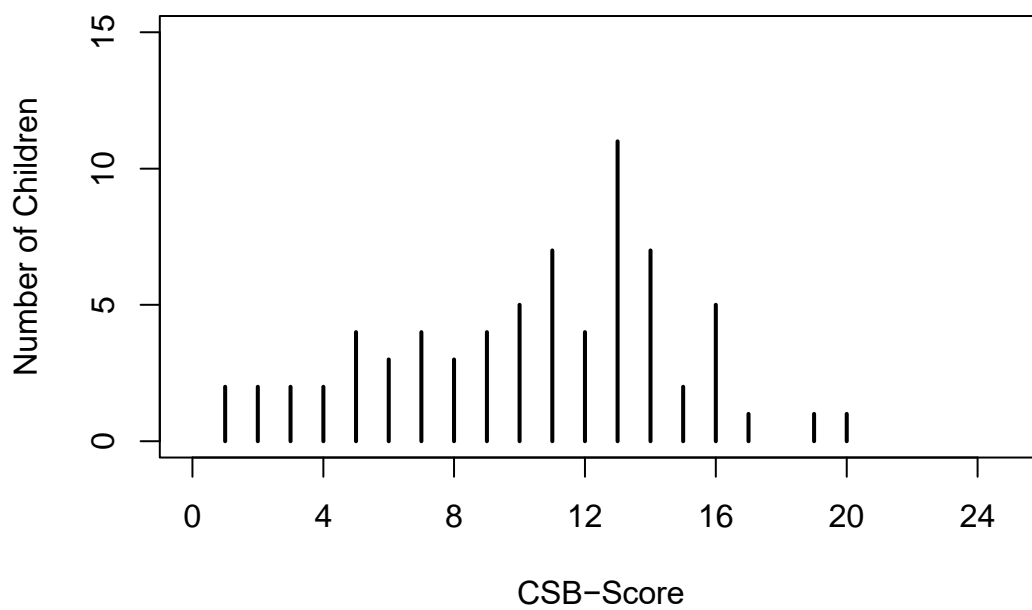


Fig. B3. Number of children per working memory score in Study 1b

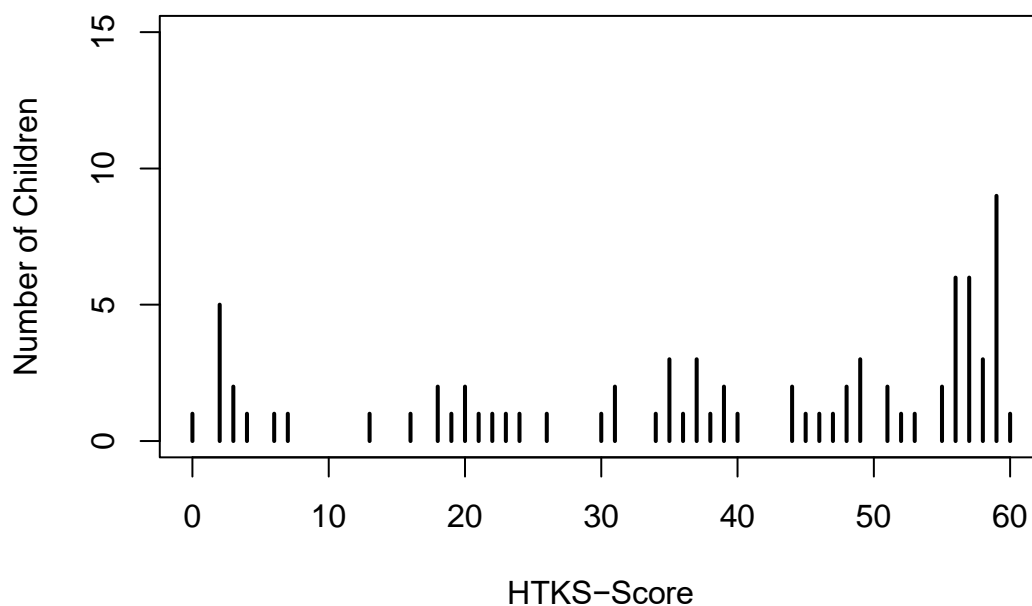


Fig. B4. Number of children per inhibitory control score in Study 1b

Order of Measures in Study 1b

The verbal intelligence task is less frustrating than the other tasks. We did not correct wrong answers and stated that items which the child could not name were indeed very difficult. This was not possible for the working memory and the inhibitory control task. Here children always realized when they started to struggle which often led to frustration. For this reason, we decided to always start with the verbal intelligence task and to move one of the frustrating tasks to the end of session. This way they were less overwhelming and exhausting. To control for possible impact of the task, we randomized which task was administered first.

Appendix C***Complete Output ANOVAS******Study 1a***

Table A1

4 (Age) x 2 (Test Question) x 2 (Scenario) ANOVA

	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Scenario	3.63	1, 186	.058	.01
Test Question	22.22	1, 186	<.001	.08
Age	2.06	3, 62	.114	.05
Scenario*Test Question	0.22	1, 186	.640	.00
Scenario*Age	0.61	3, 186	.605	.00
Test Question*Age	3.93	3, 186	.009	.04
Scenario*Test Question*Age	0.92	3, 186	.433	.00

Note. The dependent variable is number of correct trials.

Study 1b

Table A2

5 (Age) x 2 (Test Question) x 2 (Scenario) ANOVA

	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Scenario	0.28	1, 240	.598	.00
Test Question	62.81	1, 240	<.001	.00
Age	9.85	4, 80	<.001	.02
Scenario*Test Question	1.08	1, 240	.300	.00
Scenario*Age	1.15	4, 240	.332	.00
Test Question*Age	9.36	4, 240	<.001	.00
Scenario*Test Question*Age	0.84	4, 240	.499	.00

Note. The dependent variable is number of correct trials.

Study 2

Table A3

4 (Age) x 2 (Scenario) ANOVA

	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Scenario	0.00	1, 75	.926	.01
Age	8.04	3, 75	<.001	.19
Scenario*Age	1.44	3, 75	.237	.01

Note. The dependent variable is number of correct trials.

Appendix D***Online Version for 8- and 9-year-olds***

To determine the age of onset of children's spontaneous appreciation of aspectuality of intentions, we administered the main task of Study 1 and a 2nd-order belief test question to a sample of seventeen 8- and seventeen 9-year-olds (96-118 months, $M=106.59$ months, $SD=7.03$; 15 male) participated in an online version of the main task. One further 9-year-old was tested but excluded because of technical issues. Out of pandemic-related reasons these data had to be collected online. We therefore used the material originally used in the paper pencil version for adults (see Appendix A) and transformed it into videos. These videos were presented to participants via video chat. Apart from that, design and procedure were similar to Study 1's main task.

This online adaptation appears to have worked well for the Two-Objects Condition. The adaptation of the One-Object Condition however had some flaws. This condition builds on sounds (bell ringing, rattle rattling) and effects (torch shines, eraser rubs out). Depending on factors as internet connection, these sounds and effects sometimes occurred with a time delay. Moreover, the rather abstract presentation of the transfer hidden in a hand is also less clear in an online video. This was also mirrored in some children's comments (e.g., "I'm not sure whether [the hedgehog] can see this."). Mostly, these difficulties appear to have had an impact on the One-Object Belief Condition. Here, children performed significantly worse than in the Two-Objects Belief Condition ($V=88$, $p=.002$). This was not the case for the Intention

Condition ($V=18.5$, $p=.105$). The One-Object Belief Condition was also the condition in which most adults failed.

For these reasons, we refrained from including these data in the analyses of the data of the acted-out version. Nevertheless, our data show that even with the described material-related challenges 8- and 9-year-olds performed proficiently on the intention test question.

Appendix C: Schünemann, Bleijlevens, Proft, & Rakoczy (2021)

Schünemann, B., Bleijlevens, N., Proft, M., & Rakoczy, H. (2021). *Children's meta-representational notion of intentions – Understanding the subjectivity of intentions* [Unpublished manuscript]. Department of Developmental Psychology. University of Göttingen.

Children's meta-representational notion of intentions
– Understanding the subjectivity of intentions

Britta Schünemann, Natalie Bleijlevens, Marina Proft, and Hannes Rakoczy

Department of Developmental Psychology, University of Göttingen, Göttingen, Germany

Abstract

A genuine meta-representational notion of intentions should account for the subjectivity of intentions. But when do children develop a subjective understanding of intentions? One way in which intentions are subjective is that they are aspectual. Whether or not an action is intentional depends on the description or aspect the agent represents. Oedipus, for example, intentionally married Yocasta. However, as he did not know she was also his mother he did not marry her intentionally under the description “his mother”. In the present study, we tested when children begin to appreciate the aspectuality of intentions. Children of age four to six learned about an agent who harmed someone else. The agent, however, mis-represented her action to be helping. Thus, she did not harm the other agent intentionally. This was contrasted to an agent who correctly represented her action as harming and thus did so intentionally. Children were asked to evaluate the agent’s action by rewarding or punishing it accordingly. We found that 5- and 6-year-olds were more likely to punish the agent who correctly represented her action to be harmful. They seem to have appreciated that intentions are aspectual, and an action is not intentional under mis-represented descriptions. In contrast, 4-year-olds show such differentiation in their evaluations. These results suggest that children develop a subjective understanding of intentions around the age of five.

The ability to understand intentional action is one of the most fundamental capacities of human social cognition. But when and how does such an understanding develop? From an ontogenetic perspective, understanding intentional action develops in a rather protracted and stagewise manner (Astington, 2001): A rather basic and rudimentary form is already present in infancy. From early on, infants understand that actions are directed towards goals and they can distinguish intentional actions from unintentional behavior (Behne et al., 2005; Gergely & Csibra, 2003; Meltzoff, 1995; Woodward, 1998). In contrast, a more sophisticated notion develops substantially later. Only around the age of five, children begin to clearly differentiate between intentions and desires (Baird & Astington, 2005). At this age, they comprehend that intentions, in contrast to desires, commit us to performing a particular action (Bratman, 1984; Shultz & Wells, 1985). Likewise, children at that age begin to appreciate that intentions are causally self-referential (Astington, 2001; Schult, 2002). Intentions are only fulfilled if the intention itself causes the action. If an agent desires that her uncle is dead (so she can inherit his fortune), this desire will be fulfilled when her uncle dies, irrespective of how he dies. In contrast, the intention to run over her uncle to kill him will only be fulfilled if it the intention that actually causes her this action. If she does not pay attention (because she is absorbed in her murder-plans) and accidentally runs over a man who happens to be her uncle, her intention is not realized (Searle, 1983). Thus, it seems that from early on children understand intentional action based on observable features. But it is not until much later they can consider not-directly observable features as commitment to action and causal self-referentiality. This more sophisticated reasoning capacity has been interpreted to indicate a meta-representational notion of intentions (Astington, 2001; Baird & Astington, 2005): Children understand that intentions are subjective representations of the reality. These representations are separate from reality and can mis-represent it. Thus, to reason about an agent's intentional action, it is necessary to relativize to the agent's subjective standpoint and meta-represent her representation.

However, a genuine meta-representational notion of intentions should not only be able consider less openly accessible feature. It would have to be reflected in a profound subjective understanding. One way in which intentions are subjective is that they are aspectual (Anscombe, 1957; Searle, 1983). Action can have numerous applicable descriptions (I run a marathon, I am doing sports, I move one leg after the other). Typically, we perform an action intentionally under some specific description or aspect but not under others. Under which description an action is intentional depends on how the agent represents the action. It will be unintentional

under descriptions the agent does not represent (or attend) or mis-represents. Consider the following example: In the TV series *Breaking Bad*, the detective Hank tries to catch the Meth cook everyone knows as Heisenberg. Hank believes that Heisenberg has to be some threatening mafia boss-like gangster. Accordingly, he sees no problem in sharing insights and plans of his investigation with his seemingly boring and inconspicuous brother-in-law, Walt. Unfortunately, Walt is in fact the very same person as the mysterious Heisenberg. Did Hank intentionally share secret information about how to catch Heisenberg with Heisenberg himself? In some broader sense, one could say that the action of talking about secret information was intentional. However, in some other stricter sense, one strongly reject the claim that Hank intentionally gave away secret information to Heisenberg. Why is this the case? “Giving away secret information to Heisenberg” is a correct description of Hank’s action. However, Hank mis-represents this description. He represents his action as “Giving away secret information to inconspicuous Walt”. This makes his action intentional under the latter description. But the action is unintentional under the first description, which Hank mis-represent. To come to this conclusion, we need to appreciate that intentions are aspectual. We need to represent under which description Hank represents his action.

In this study, we aim to test at what age children develop such a subjective understanding of intentional action. When do they appreciate that actions are unintentional under mis-represented descriptions? First evidence indicates that such an understanding develops rather late (Schünemann et al., 2021): Children observed an agent who mis-represented a box to contain only a ball but not a pen. The agent then took the box in order to obtain the ball. Children failed to appreciate that while the agent intentionally took the ball, she did not intentionally take the pen until the age of eight. Only when children were specifically reminded of the agent’s mis-representation children succeeded by the age of six. At first glance, this questions the interpretation that children develop a meta-representational notion of intentions by the age of five. But do these results really demonstrate a limitation of competence or only of performance?

One substantial linguistic performance factor of this task is the reading children apply to the test question. We can apply different readings to intentions (Jacob, 2019; Nelson, 2019; Quine, 1956). One form of reading is the *de dicto* reading. When an ascriber applies a *de dicto reading* (about what is said), her intention report is only true if the agent herself would report her intention in this way. In consequence, we cannot substitute coreferential terms because this potentially changes the report’s truth value. For instance, Hank would agree to the report “Hank

intends to give away information to Walt”. But he would not agree if we exchanged Walt by Heisenberg, “Hank intends to give away information to Heisenberg”. This is also the reading the test question aimed at and which adults seem to apply in such tasks (Schünemann et al., 2021). Yet, one can also apply a *de re* reading of intentions: “The action (in itself) of giving away information to Walt/Heisenberg is such that Hank intends it”. In contrast to *de dicto* reading, this *de re* reading allows for the substitution of coreferential terms. Thus, under a *de re* reading, Hank’s action is also intentional under the description “giving away information to Heisenberg”.

Nevertheless, in Hank’s example, it is fairly obvious to apply a *de dicto* reading to ascribe his intentions. After all, his credibility as a detective will be put into question if we suspect him to give away information to Heisenberg. In contrast, much less depends on whether or not the agent intentionally took the pen. Possibly, children even younger than six can appreciate the aspectuality of intentions. But they simply fail to recognize the necessity for a *de dicto* reading in cases which are not sufficiently relevant and salient.

Support for this possibility comes from two studies on the interplay of ignorance and intentional action. In an irrelevant context, children did not consider the agent’s ignorance when they evaluated the agent’s intention until age eight. They falsely stated that an agent intentionally gave keys to Cathy’s dad although she was completely unaware that this man was Cathy’s dad. In contrast, in a highly relevant context even 5-year-olds correctly considered the agent’s ignorance.

For this reason, this study tested for the age of onset at which children appreciated that actions are unintentional under mis-represented descriptions in a relevant context. We confronted children with an agent who either intentionally or unintentionally harmed another agent. Children were asked to evaluate the agent’s actions as either bad, good or neither good nor bad by punishing or rewarding the agent accordingly. To tap their appreciation of the aspectuality of intentions, we tested whether they considered the agent’s intentions in their evaluations. In the experimental condition, the agent mis-represented the description of her action. In consequence, she harmed the other agent unintentionally. In contrast, in the control condition, the agent correctly represented her action as harmful and thus, harmed the other agent intentionally. Children who appreciate the aspectuality of intentions should distinguish accordingly between both conditions in their evaluations.

Methods

Participants

Seventy-five 4-, 5- and 6-year-olds (25 children per age group, 49-83 months, $M=65.72$ months, $SD=9.49$; 45 male) were recruited online and from a databank of children whose parents had previously given consent to experimental participation. A minimal sample size of 72 children was calculated via G*Power 3.1.9.2 in order to have a test power of at least $1-\beta = 0.8$ for conducting an analysis of variance for a mixed design (one three-level between-subject factor and one two-level within-subjects factor) assuming a moderate effect size ($\eta^2=.03$) of the interaction effect. To obtain valid data we introduced an inclusion criterion that ensured that children understood the dependent measure (see evaluation measure). We tested twenty-two additional children who had to be excluded from analyses because they did not meet this inclusion criterion. Six more children had to be excluded because of technical issues ($N=3$), uncooperative behavior ($N=2$) and parents' interference with their children's responses ($N=1$). Children were from mixed socioeconomic backgrounds.

Procedure

Children were tested online. Child and experimenter interacted via video chat (platform: BigBlueButton). Children saw the testing material (picture stories and videos) and a small video of the experimenter and themselves. To facilitate answering, we color-coded answers: Answer-options were illustrated and highlighted in different colors. Children could either answer test questions directly or by referring to the respective color (for a detailed explanation of the advantages of this approach, see Sheskin & Keil, 2018).

All test sessions began with a warm-up in which children were introduced to answering by referring to colors using simple examples (e.g., "On which color is the animal that can fly?"). This was followed by a familiarization with the main task's dependent measure. Then, children received the intention task.

Familiarization with Evaluation Measure

We used behavior evaluation to measure how children's intention ascribed intentions. Children were asked to distribute resources according to their evaluation of the agent's behavior: Give a marble to the agent for good behavior, take a marble away for bad behavior and do not change the number of marbles for behavior that is neither good nor bad. The logic is the following, if the agent does something bad intentionally, she should be punished while she should not be

punished for doing something bad unintentionally. We included the last option (not changing the number of marbles) to not require children to combine negative outcomes with positive consequences.

We used a three-stage procedure to familiarize with this measure. In all three stages, children were told about a protagonist “Peter”. This always followed the pattern: “Here, Peter does X (un)intentionally. Here, Peter is good/bad/neither good nor bad.” Respectively, X was either intentional and had a good outcome, intentional and had a bad outcome or unintentional and had a bad outcome. All stories were presented together with a drawing of Peter doing X. Following each Peter-story, children saw the color-coded answering options (see Figure 1).

In the first stage, the experimenter then evaluates the behavior herself: “That is why we give a marble to Peter/take a marble away/ do nothing with his marbles. So, we choose [the according color].”

In the second stage, the experimenter asked the child to evaluate. “How did Peter just behave? Do you do something with his marbles?”. Here, incorrect answers were corrected.

In the third stage, the experimenter again asked the child to evaluate Peter’s behavior. However, in this stage, incorrect answers were not corrected. Children who failed to correctly assign marbles in this stage were excluded from analyses.

Figure 1

Color Coding of the Dependent Evaluation Measure



Intention task

Children watched videos of an agent A whose action had a negative outcome for another agent B. In the experimental condition, A misrepresents what she is doing. Accordingly, she does not bring about the negative outcome intentionally. In the control condition, A holds no

misrepresentation and thus brings about the negative outcome intentionally. Children then were asked to evaluate A's behavior using the evaluation measure.

Experimental Condition

Children saw A and B (see Fig. 2). B wants to do something (e.g., draw something) but she lacks a necessary tool for this action (e.g., pencil). A third character C enters the scene and puts the desired object in one box and an undesirable object (e.g., piece of paper) in the other box (1). B asks A to give her the desired object. A leaves the scene in order to get a stool to be able to reach the boxes. In A's absence, C swaps the content of the two boxes. Children are asked the control question:

“Did A see that this [the content] has been swapped?”

Incorrect answers were corrected (2). A returns now holding a mis-representation about the object's location (3.). She takes the box which she falsely believes to contain the desired object but which in fact now contains the undesirable object (4). Thus now, A misrepresents the real description of her action (*give box containing undesirable object*) to be *give box containing desired object*. This makes her action intentional under the latter but not under the mis-represented description.

Children then received the test question:

“How did A just behave? Do you do something with her marbles?”

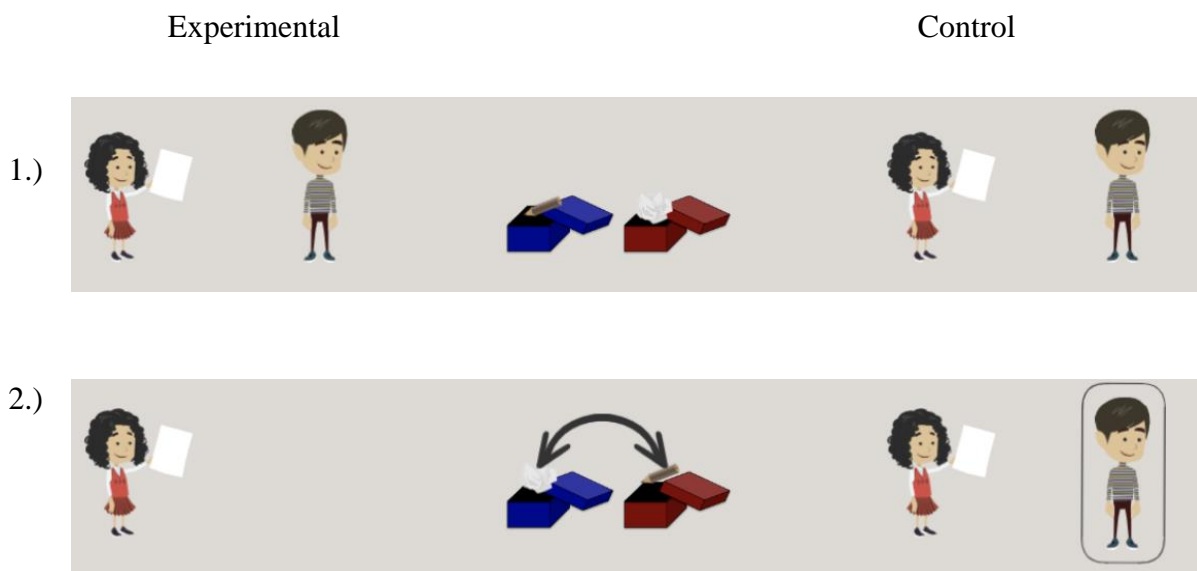
Correct answers had to consider that the agent did not bring about the negative outcome intentionally. Thus, it was correct to either give a marble to the agent or do nothing with her marbles.

Control Condition

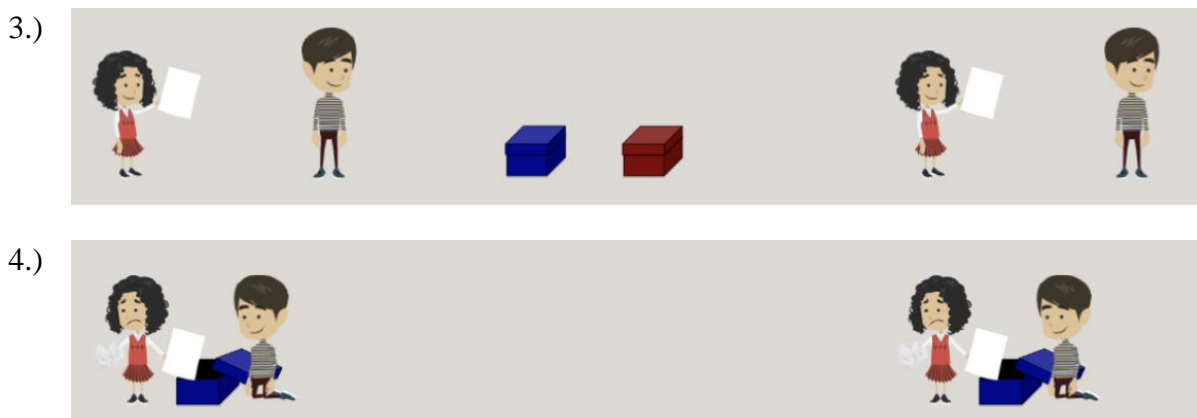
The control condition was similar to the experimental condition with one exception (in 4). A returns before C swaps the boxes' contents and observes which object was put in which box. She still gives the box containing the undesirable object to B. Hence in the control condition, A correctly represents her action under the description *give box containing undesirable object*. This means she brings about the negative outcome intentionally. Thus, here the correct answer to the test question was to take a marble away.

Figure 2

a) Procedure Experimental and Control Condition



Control Question: Did A see that this has been swapped?



Test Question: How did A just behave? Do you do something with her marbles?

b) Basic Set-up of the Intention Task's Videos



Order of Trials

Children received two trials of each condition in a block. Order of condition was counterbalanced. Before the first test trial, children received two filler trials to familiarize them with the procedure. In filler trials, the content of boxes was not swapped. In the first filler trial, A gave the desired object to B and in the second, the undesirable. In between test trial-blocks, children received a third filler trial, in which A gave the desirable object to B. The reason for the third filler trial was to interrupt the structure of A constantly giving the undesirable object. After each filler trial, children evaluated A's behavior. Each test and filler trial had different protagonists and object.

Results

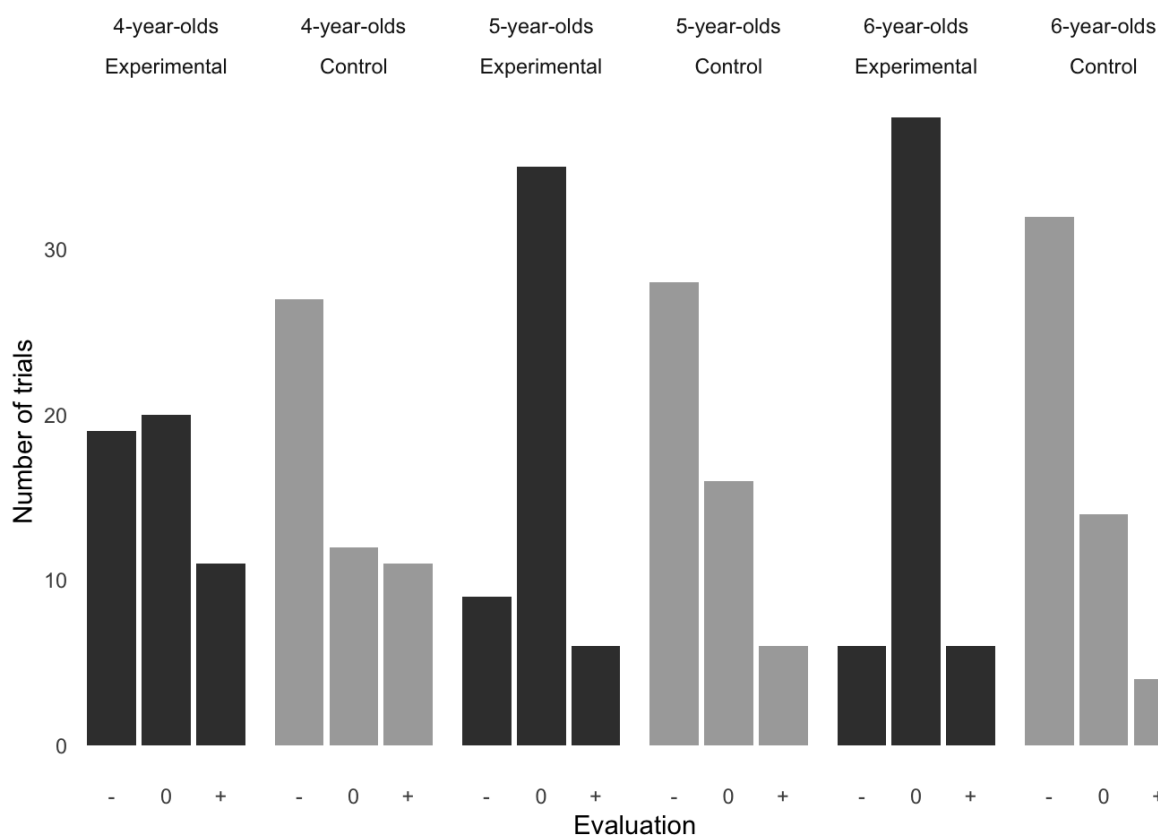
Coding

For each trial we coded the action the child chose to evaluate the agent's action. Children could either take a marble away (-), do not change the number of marbles (0) or give a marble to the agent (+).

Plan of Analyses

When we pre-registered this study, we expected that, as a default, children would punish the agent's harmful behavior. We expected children to only refrain from punishing when they understand that the agent did not perform the harmful action intentionally in the experimental condition. Figure 3 depicts the frequency of each evaluation as a function of age group and condition. Contrary to our expectations, children were more likely to not punish the agent. Children only punished the agent in 40% of trials. For this reason, we applied the following logic: Children who understand that the agent acted intentionally in the control condition but unintentionally in the experimental condition should differentiate in their evaluation in the following way: They should deviate from the default "not punishing" more in the control condition than in the experimental condition. Children who cannot yet consider the agent's intention should not differentiate between conditions in their evaluations.

We conducted McNemar-tests to test whether children were more likely to punish the agent in the control than in the experimental condition. To this end, we combined the response options "+" and "0" to the category "not punishing" and compared it to "punishing (" - "). To determine at which age children correctly considered the agent's intentions, we conducted separate tests for each age group.

Figure 3*Number of Trials per Evaluation as a Function of Age Group and Condition*

We found that 5- and 6-year-olds more likely to punish the agent in the control condition than in the experimental condition (5-year-olds: $\chi^2(1) = 5.452$, $p = .020$); 6-year-olds: $\chi^2(1) = 6$, $p = .014$, see also table 1). In contrast, 4-year-olds were not more likely to punish the agent in the control condition ($\chi^2(1) = 0.381$, $p = .537$). Thus, while 5- and 6-year-olds considered the agent's intention in their evaluation, 4-year-olds did not.

Table 1*Number of Trials in which Children Punished and did Not Punish the Agent by Age Group and Condition*

	4-year-olds		5-year-olds		6-year-olds	
	Experimental	Control	Experimental	Control	Experimental	Control
Not Punishing	31	23	41	22	44	18
Punishing	19	27	9	28	6	32

Discussion

This main aim of this study was to identify the age at which children understand the subjectivity of intentions. To this end, we tested when children begin to appreciate that intentions are aspectual. To ensure that children recognize the necessity to take the agent's perspective, we embedded the task in a morally relevant context. Children observed an agent who harmed another agent. Yet, she mis-represented her harmful action as helping. In consequence, the agent did not intentionally harm the other agent. This was contrasted to an agent who represented her harmful action correctly and thus, harmed the other agent intentionally. Children were asked to evaluate these actions by punishing or rewarding the agent accordingly.

We found that overall children were highly reluctant to punish the agent. Nevertheless, 5- and 6-year-olds distinguished in their evaluation between intentional and unintentional harmful behavior. They were more likely to punish the agent, when she represented her action correctly and thus harmed the other agent intentionally. In contrast, 4-year-olds did not consider the agents' intentions in their evaluation. These results suggest, that by the age of five children understand that an action is not intentional under a description the agent mis-represents. They seem to understand that intentions are subjective in that they are aspectual. This supports the claim that children at this age develop a meta-representational notion of intentions (Astington, 2001; Baird & Astington, 2005). This claim has built on evidence that children around the age of five grasp complex features of intentions which are not directly observable (such as commitment to action and causal self-referentiality; Schult, 2002; Shultz & Wells, 1985). Our findings extend on this evidence in a fundamental way. They indicate that 5-year-olds really meta-represent how the agent's intention represents her action.

Moreover, our findings suggest that children are more inclined to relativize to the agent's standpoint in more salient, relevant contexts. It seems that children's difficulties to appreciate the aspectuality of intentions in previous work reflect a performance limitation (Schünemann et al., 2021). Children might simply not have recognized the necessity to consider the agent's perspective (apply a *de dicto* reading). Future research needs to systematically test the impact relevance has on children's performance in subjective perspective taking.

Furthermore, future research needs to address alternative explanations that result from the nature of our dependent measure. We chose evaluation as our dependent measure for two reasons: First, it increased the salience of the agent's subjective perspective (you do not want

to harm someone who had no intention to harm). Second, it allowed us to refrain from using the word intentionally, which has a rather negative connotation in German language. However, we cannot preclude that children younger than five were simply reluctant to punish the agent. And although they understood which agent performed the harmful action intentionally and who did not, they just did not want to punish anyone at all. Also, our dependent measure posed high inferential demands. Children had to infer the agent's mis-representation from the situational context, then the agent's intention, then how to evaluate the action, and then the according reward or punishment. Previous work shows that such inferential chains impact children's performance in the context of intention ascription (Schünemann et al., 2021) and moral evaluations (Proft & Rakoczy, 2018).

In conclusion, this study found that children appreciate that intentions are aspectual by the age of five. They seem to have developed a subjective understanding of intentions. This supports the assumption that children develop a meta-representational notion of intentions around this age. Moreover, our findings indicate that younger children only rely on their competence to take the agent's subjective perspective into account when a certain threshold in terms of relevance and salience is met. Future research needs address the exact impact of relevance on children's performance.

References

- Anscombe, G. E. M. (1957). *Intention*. Basil Blackwell.
- Astington, J. W. (2001). The paradox of intention: Assessing children's metarepresentational understanding. In *Intentions and intentionality: Foundations of social cognition* (Vol. 4, pp. 85–103). The MIT Press.
- Baird, J. A., & Astington, J. W. (2005). The development of the intention concept: From the observable world to the unobservable mind. *The New Unconscious*, 256–276.
- Behne, T., Carpenter, M., Call, J., & Tomasello, M. (2005). Unwilling Versus Unable: Infants' Understanding of Intentional Action. *Developmental Psychology*, 41(2), 328–337. <https://doi.org/10.1037/0012-1649.41.2.328>
- Bratman, M. (1984). Two Faces of Intention. *The Philosophical Review*, 93(3), 375. <https://doi.org/10.2307/2184542>
- Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences*, 7(7), 287–292. [https://doi.org/10.1016/S1364-6613\(03\)00128-1](https://doi.org/10.1016/S1364-6613(03)00128-1)
- Jacob, P. (2019). Intentionality. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2019). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/spr2017/entries/intentionality/>
- Meltzoff, A. N. (1995). Understanding the Intentions of Others: Re-Enactment of Intended Acts by 18-Month-Old Children. *Developmental Psychology*, 31(5), 838–850. <https://doi.org/10.1037/0012-1649.31.5.838>
- Nelson, M. (2019). Propositional Attitude Reports. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2019). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/spr2019/entries/prop-attitude-reports/>
- Proft, M., & Rakoczy, H. (2018). The ontogeny of intent-based normative judgments. *Developmental Science*, e12728. <https://doi.org/10.1111/desc.12728>
- Quine, W. V. (1956). Quantifiers and propositional attitudes. *The Journal of Philosophy*, 53(5), 177–187.

- Schult, C. A. (2002). Children's Understanding of the Distinction between Intentions and Desires. *Child Development*, 73(6), 1727–1747. <https://doi.org/10.1111/1467-8624.t01-1-00502>
- Schünemann, B., Proft, M., & Rakoczy, H. (2021). *Children's developing understanding of the subjectivity of intentions – a case of “advanced Theory of Mind.”*
- Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge university press.
- Sheskin, M., & Keil, F. (2018). TheChildLab. Com a video chat platform for developmental research.
- Shultz, T. R., & Wells, D. (1985). Judging the intentionality of action-outcomes. *Developmental Psychology*, 21(1), 83.
- Woodward, A. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, 69(1), 1–34. [https://doi.org/10.1016/S0010-0277\(98\)00058-4](https://doi.org/10.1016/S0010-0277(98)00058-4)

Curriculum Vitae

This information is available in the printed version of the manuscript.