Altercentric Bias: A Potential New Window into Implicit Theory of Mind

Dissertation

for the award of the degree

"Doctor of Philosophy" Ph.D.

Division of Mathematics and Natural Sciences

of the Georg-August-Universität Göttingen

within the doctoral program Behavior and Cognition

of the Georg-August University School of Science (GAUSS)

submitted by

Feride Nur Haskaraca Kizilay

from Manisa, Turkey

Göttingen, 2023

Thesis Committee

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

Prof. Dr. Ulf Liszkowski, Abteilung Entwicklungspsychologie, Institut für Psychologie, Universität Hamburg

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

Members of the Examination Board

Reviewer 1: Prof. Dr. Hannes Rakoczy, Abteilung Kognitive Entwicklungspsychologie, Georg-Elias-Müller-Institut für Psychologie, Georg-August-Universität Göttingen

Reviewer 2: Prof. Dr. Ulf Liszkowski, Abteilung Entwicklungspsychologie, Institut für Psychologie, Universität Hamburg

Other Members of the Examination Board:

Dr. Tanya Behne, Abteilung Kognitive Entwicklungspsychologie, Georg-Elias-Müller-Institut für Psychologie, Georg-August-Universität Göttingen

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

Prof. Dr. Nivedita Mani, Abteilung Psychologie der Sprache, Georg-Elias-Müller-Institut für Psychologie, Georg-August-Universität Göttingen

Prof. Dr. Annekathrin Schacht, Abteilung Kognition, Emotion und Verhalten, Georg-Elias-Müller-Institut für Psychologie, Georg-August-Universität Göttingen

Oral Examination Date 11.07.2023

Preliminary Note

The present thesis is a publication-based (cumulative) dissertation. It is based on three original articles, of which one has been published, and two are currently in preparation.

Haskaraca, F.N., Proft, M., Liszkowski, U., & Rakoczy, H. (2023). How robust are egocentric and altercentric interference effects in social cognition? A test with explicit and implicit versions of a continuous False Belief task. *Frontiers in Psychology, 14,* 1142302.

This article has been published in Frontiers in Psychology, published by Frontiers Media S.A.

Haskaraca, F.N., Proft, M., Liszkowski, U., & Rakoczy, H. (2023). *Testing Egocentric and Altercentric Biases in a Referential Communication Task* [Unpublished manuscript]. Department of Developmental Psychology. University of Göttingen.

Haskaraca, F.N., Proft, M., Liszkowski, U., & Rakoczy, H. (2023). *Measuring Altercentric Biases in Level-1 & Level-2 Perspective-Taking Tasks by the Means of Mouse-Tracking* [Unpublished manuscript]. Department of Developmental Psychology. University of Göttingen.

In my dissertation, I bring together three manuscripts within a general theoretical framework, and I provide an extensive literature review, a detailed summary of empirical findings, and a general discussion about the empirical and theoretical implications of the current projects. The manuscripts are attached to the dissertation in Appendices A, B, & C. I served or will serve as the first author in all published work based on these manuscripts. I contributed to the current projects by developing ideas; designing, conducting, and supervising studies; analyzing and interpreting the data; and writing up the manuscripts. My co-authors supported me with their helpful suggestions and comments throughout the whole process. I hereby declare that all parts of my dissertation were written by myself, the assistance of third parties was only accepted if scientifically justifiable and acceptable with regard to the examination regulations, and all sources have been cited.

Thank you…

To my supervisor Hannes Rakoczy for guidance, cooperation, and discussion of research ideas and for educating me on thinking bigger and writing *crispier*. To Marina Proft for her constructive feedback, great study ideas, and support in rough times.

To Ulf Liszkowski and Claudia Fichtel for their feedback and input in the thesis committee meetings and for support and understanding throughout the whole time.

To Tanya Behne, Nivedita Mani, and Annekathrin Schacht for agreeing to be examination board members.

To the whole Kindsköpfe lab and to Regina Zörner and Marlen Kaufmann for their help and assistance that made these projects possible. Also, to all student assistants and interns for their assistance that helped me in different phases of this dissertation.

To all students who have worked with me on my projects: Marilena Reinhardt, Tobias Wiegand, Jana Tatian, Finja Schaade, Nele Neuhäuser, Mona Gröneweg, and Tijana Lajic.

To Ulf Liszkowski, Charlotte Grosse Wiesmann, and Marie Luise Speiger for great collaborations.

To Britta Schünemann, Lydia Schidelko, Rowan Titchener, Natalie Bleijlevens, Isa Garbisch, and Saba Amirhaftehran for their unconditional support, all the German-English translations, and the inspirational coffee breaks.

To Isa Garbisch, Natalie Bleijlevens, and Saba Amirhaftehran for proofreading earlier drafts of this thesis. Your comments and suggestions significantly improved the quality of this work.

To Hande Ilgaz and Jed Allen for their endless support and mentorship.

To my dearest friends, Elif Dirik and Esra Şenol, for making rough times bearable.

To my parents, brothers, and sisters who made my Ph.D. journey possible with their moral support and created all the family drama which kept me in this field for so long.

To my nephew and my nieces for the endless inspiration.

To my cat, Türkan, for bringing color into my life.

And to my dear husband, my love, Hamza Kızılay, for being there for me *after all this time* and *always*.

And, finally, to all participants who have participated in my studies.

Table of Contents

0 Summaries

0.1 English Summary

It has been traditionally assumed that theory of mind (ToM) -our ability to ascribe mental states to others and ourselves- emerges around age four, as indicated in performance on standard explicit false belief (FB) tasks (Wellman et al., 2001). More recent studies assessing FB understanding with implicit measures suggested that some form of ToM may be present even in infancy (Scott & Baillargeon, 2017). However, many of these studies now face replicability issues and thus cannot serve as robust evidence for implicit ToM (e.g., Dörrenberg et al., 2018). One type of implicit task, namely altercentric bias, still constitutes a promising alternative to tap implicit perspective-taking abilities. Altercentric bias is an indicator of spontaneous and implicit mentalizing of others' (irrelevant) perspectives: people get slower and more error-prone in making first-order judgments about the world if another agent in the scene holds a diverging perspective, even if this perspective is irrelevant or detrimental to the task (Samson et al., 2010). This bias has been studied in different task formats so far (e.g., dot-perspective task by Samson et al., 2010; object-detection task by Kovács et al., 2010); however, the existing altercentric bias measures have been shown subject to reliability and validity issues (e.g., Santiesteban et al., 2014).

The current work takes a new approach to the altercentric bias and aims to develop and adapt reliable altercentric bias measures through the mediums of existing egocentric bias measures. Egocentric interferences refer to the interferences from our own perspectives on our judgments about others' perspectives. Egocentric bias measures easily lend themselves to analogous tasks that can be used to reveal altercentric biases. The current study tests three different tasks that could provide potential new ways to tap altercentric interferences, with the ultimate aim of using this bias as a window into implicit ToM. All studies constituting this dissertation have been conducted with adult participants via unmoderated online sessions (except for the two baseline studies of Project 1, which were conducted in live settings). The first project of the current thesis capitalizes on the so-called Sandbox task (e.g., Sommerville et al., 2013) as a means to tap altercentric biases. This measure has been originally developed to tap egocentric biases. This project also aimed to replicate the earlier Sandbox studies, which revealed robust egocentric interferences across the lifespan. Across five different Sandbox studies, we found no evidence for egocentric and altercentric biases. The second project focuses on the so-called Director Task (e.g., Samuel et al., 2019) as a potential new

way of tapping altercentric biases. Like the Sandbox task, this measure has predominantly targeted egocentric interferences in social cognition or egocentric heuristics in communication. Therefore, this project also aimed to replicate the earlier studies in terms of the egocentric interferences revealed. Across two studies, this task provided evidence for robust egocentric interference effects. The results, however, revealed inconsistent effects in terms of the altercentric version, raising critical issues regarding the validity of this task as a measure of altercentric biases. Finally, in Project 3, altercentric biases have been tested in typical level-I and level-II visual perspective-taking tasks (e.g., Flavell et al., 1981). In two studies, altercentric biases were revealed in both level-I and level-II visual perspective-taking tasks. However, the patterns revealed by the different measures drew a complicated picture and prevented us from arriving at a conclusion about the presence and nature of these biases.

The findings of this dissertation contribute to the methodological debates in the altercentric bias literature and the theoretical debates in the field of implicit ToM. The current projects showed that altercentric biases might not be revealed at all or may be contingent on alternative explanations in some task formats. The one task that remained to be a somewhat promising measure of altercentric interferences (as revealed by Project 3) showed that these biases were likely to occur due to implicit mentalizing, and they hinted at unified ToM abilities. However, these claims stayed at the level of speculation as none of the tasks used in the current thesis was free from methodological concerns.

0.2 Deutsche Zusammenfassung

Traditionell wurde angenommen, dass sich Theory of Mind (ToM) - unsere Fähigkeit, anderen und uns selbst mentale Zustände zuzuschreiben - im Alter von etwa vier Jahren herausbildet, wie das Lösen von expliziten Standardaufgaben zufalschen Überzeugungen (englisch: False Belief; FB) zeigt (Wellman et al., 2001). Neuere Studien, in denen das FB-Verständnis mit impliziten Maßen untersucht wurde, deuten darauf hin, dass eine Form von ToM bereits im Säuglingsalter vorhanden sein könnte (Scott & Baillargeon, 2017). Viele dieser Studien haben jedoch Probleme mit der Replizierbarkeit und können daher nicht als robuster Beweis für implizites ToM dienen (z. B. Dörrenberg et al., 2018). Eine Art von impliziten Aufgaben, nämlich die alterzentrische Verzerrung, stellt nach wie vor eine vielversprechende Alternative dar, um implizite Fähigkeiten zur Perspektivenübernahme zu erfassen. Die alterzentrische Verzerrung ist ein Indikator für das spontane und implizite Übernehmen von (irrelevanten) Perspektiven anderer: Menschen werden langsamer und fehleranfälliger bei Urteilen über die Welt, wenn ein anderer Akteur in der Szene eine abweichende Perspektive einnimmt, selbst wenn diese Perspektive für die Aufgabe irrelevant oder schädlich ist (Samson et al., 2010). Diese Verzerrungen wurde bisher in verschiedenen Aufgabenformaten untersucht (z.B. dot-perspective task von Samson et al., 2010; object-detection task von Kovács et al., 2010). Es hat sich jedoch gezeigt, dass die Messungen der alterzentrischen Verzerrungen Probleme mit der Replizierbarkeit und Validität haben (z.B. Santiesteban et al., 2014).

Die vorliegende Arbeit verfolgt einen neuen Ansatz für die alterzentrische Verzerrung und zielt darauf ab, verlässliche Maße für die alterzentrische Verzerrung mit Hilfe bestehender Maße für die egozentrische Verzerrung zu entwickeln und anzupassen. Egozentrische Interferenzen beziehen sich auf die Interferenzen unserer eigenen Perspektiven auf unsere Urteile durch die Perspektiven anderer. Egozentrische Verzerrungen lassen sich leicht mit analogen Aufgaben messen, die zur Aufdeckung alterzentrischer Verzerrungen verwendet werden können. In der aktuellen Studie werden drei verschiedene Aufgaben getestet, die potenziell neue Möglichkeiten zur Erfassung alterzentrischer Verzerrungen bieten könnten, mit dem Ziel, diese Verzerrungen als Fenster zum impliziten ToM zu nutzen. Alle Studien im Rahmen dieser Dissertation wurden mit erwachsenen Teilnehmern in unmoderierten Online-Sitzungen durchgeführt (mit Ausnahme der beiden Ausgangsstudien von Projekt 1, die in einer Live-Situation durchgeführt wurden). Das erste Projekt der vorliegenden Arbeit nutzt die so genannte Sandbox-Aufgabe als Mittel (z.B. Sommerville et al., 2013), um alterzentrische Verzerrungen zu erfassen. Diese Messung wurde ursprünglich entwickelt, um egozentrische Verzerrungen zu erfassen. Dieses Projekt zielte auch darauf ab, die früheren Sandbox-Studien zu replizieren, die robuste egozentrische Interferenzen über die gesamte Lebensspanne hinweg aufzeigten. In fünf verschiedenen Sandbox-Studien fanden wir keine Hinweise auf egozentrische und alterzentrische Verzerrungen. Das zweite Projekt konzentriert sich auf die so genannte Direktor-Aufgabe (z.B. Samuel et al., 2019) als potenzielle neue Methode zur Erfassung alterzentrischer Verzerrungen. Wie die Sandbox-Aufgabe zielt auch diese Messung in erster Linie auf egozentrische Interferenzen in der sozialen Kognition oder egozentrische Heuristiken in der Kommunikation ab. Daher zielte dieses Projekt auch darauf ab, die früheren Studien in Bezug auf die aufgedeckten egozentrischen Interferenzen zu replizieren. In zwei Studien lieferte diese Aufgabe Beweise für robuste egozentrische Interferenzeffekte. Die Ergebnisse zeigten jedoch inkonsistente Effekte in Bezug auf die alterzentrische Version, was kritische Fragen hinsichtlich der Validität dieser Aufgabe als Maß für alterzentrische Verzerrungen aufwirft. In Projekt 3 wurden schließlich alterzentrische Verzerrungen in visuellen Aufgaben zur Perspektivenübernahme auf Ebene I und Ebene II getestet (z.B. Flavell et al., 1981). In zwei Studien wurden alterzentrische Verzerrungen sowohl bei visuellen Aufgaben der Ebene I als auch der Ebene II festgestellt. Die Muster, die sich aus den verschiedenen Messungen ergaben, zeichneten jedoch ein kompliziertes Bild und hinderten uns daran, eine Schlussfolgerung über das Vorhandensein und die Art dieser Verzerrungen zu ziehen.

Die Ergebnisse dieser Dissertation tragen zu den methodologischen Debatten in der Literatur über alterzentrische Verzerrungen und zu den theoretischen Debatten im Bereich des impliziten ToM bei. Die aktuellen Projekte haben gezeigt, dass alterzentrische Verzerrungen bei einigen Aufgabenformaten möglicherweise überhaupt nicht aufgedeckt werden oder von alternativen Erklärungen abhängig sind. Die eine Aufgabe, die sich als einigermaßen vielversprechendes Maß für alterzentrische Störungen erwies (wie von Projekt 3 aufgedeckt), zeigte, dass diese Verzerrungen wahrscheinlich auf die implizite Perspektivübernahme zurückzuführen ist, und sie deutete auf einheitliche ToM-Fähigkeiten hin. Diese Behauptungen blieben jedoch spekulativ, da keine der in der vorliegenden Arbeit verwendeten Aufgaben frei von methodischen Einschränkungen war.

1 General Introduction

> "He can read minds?" said Harry quickly, his worst fears confirmed.
>
> "You have no subtlety, Potter" said Snape […] "Only Muggles talk of 'mind reading.' The mind is not a book, to be opened at will and examined at leisure. Thoughts are not etched on the inside of skulls, to be perused by any invader. The mind is a complex and many-layered thing, Potter… or at least, most minds are…" He smirked. "It is true, however, that those who have mastered Legilimency are able, under certain conditions, to delve into the minds of their victims and to interpret their findings correctly […]"
>
> Whatever Snape said, Legilimency sounded like mind reading to Harry and he did not like the sound of it at all. (Rowling, 2003, pp. 530-531)

People's minds and mental states are complex constructs that are unobservable, subjective, and private. These cannot be simply read but inferred from behavior and evaluated within the (social) context. The ability to ascribe mental states to other people and the self, or in other words Theory of Mind, is essential for social interactions. Research on Theory of Mind has occupied the socio-cognitive development literature for the past thirty years (see for reviews, Miller, 2012; Wellman, 2014). For a long time, it has been assumed that Theory of Mind shows major advances in the preschool years, specifically around age four. This assumption was primarily based on children's performance on verbal, standard false belief tasks that require children to explicitly ascribe beliefs to mistaken agents and predict their behavior according to these false beliefs (see for a review, Wellman et al., 2001). However, newer studies assessing false belief understanding with implicit, non-verbal measures, such as looking time paradigms and interactive tasks, revealed that some form of Theory of Mind may be present even in the first or second year of life (see for a review, Scott & Baillargeon, 2017). However, more recent studies using these implicit tasks failed to replicate earlier findings and could not provide robust evidence for Theory of Mind understanding in infancy (e.g., Dörrenberg et al., 2018).

One line of research in implicit Theory of Mind literature remains promising. It is a different type of implicit task called altercentric bias, which refers to the interferences from others' (irrelevant) perspectives on our own judgments and behavior. For example, it has been found that participants were slower in detecting and counting objects when another agent was present but had an incongruent perspective on the same objects (e.g., Kovács et al., 2010; Samson et al., 2010).

Altercentric interference effects have been documented in different settings using different task formats (e.g., Kampis & Kovács, 2022; Kovács et al., 2010; Samson et al., 2010; Van der Wel et al., 2014). However, the reliability of the existing altercentric bias tasks is an open question (e.g., Conway et al., 2017; Phillips et al., 2015). The current thesis aims to devise a task that can tap altercentric interferences reliably and to conduct proof-of-concept studies with adult participants testing altercentric bias as an implicit Theory of Mind measure. To this end, it capitalizes on the modifications of several tasks that have been used to reveal (the limits of) explicit perspective-taking abilities, i.e., egocentric biases (e.g., Keysar et al., 2003; Sommerville et al., 2013). Once reliable altercentric interference effects are shown, these tasks can also be used to further investigate the nature of the altercentric bias and the assumptions of different theoretical accounts on implicit Theory of Mind.

In this thesis, I will first introduce the concept of Theory of Mind and provide an overview of how it is measured and when and how it develops. Then I will introduce implicit Theory of Mind and discuss empirical and theoretical work on this concept. Afterward, I will introduce the concept of altercentric bias and review empirical and theoretical work focusing on this bias. Then I will address the possibility of tapping implicit Theory of Mind abilities by means of altercentric bias and the further investigations that this possibility would allow. The foundations for the three projects constituting this thesis lay on these ideas. After presenting the three studies conducted within the scope of this dissertation, I will explain and discuss what has been found in these studies and their empirical and theoretical implications for current and future research about implicit Theory of Mind and altercentric bias

2 Theory of Mind

Our perception and depiction of human beings differ from our perceptions of inanimate objects: We consider people (including ourselves) rational beings with subjective perspectives. As rational agents, we experience sensations and emotions, perceive our surroundings, hold beliefs that have different truth values, have desires, and intend to achieve our goals. We do not just act but perform intentional actions serving specific reasons. In order to understand rational agents and their actions, we ascribe certain beliefs, desires, emotions, intentions, and perceptions to these agents. This way of explaining human behavior is commonly known as Theory of Mind - hereafter ToM (Premack & Woodruff, 1978) and sometimes as folk psychology or belief-desire psychology. ToM is crucial to virtually every aspect of our social life: from communication and cooperation to competition and engagement with others. Without constantly monitoring other people's beliefs, knowledge states, desires, and emotions, we would be unable to interact effectively with them (Rakoczy, 2017). Due to this fundamental importance, ToM has become an important and fruitful area of research in psychology literature in the last 50 years, and a great deal of research has been conducted in order to understand the scope, limits, and cognitive foundations of ToM.

2.1 How is ToM measured?

At the heart of ToM lies the metarepresentation of propositional attitudes (Rakoczy, 2022): the ability to understand that other people have beliefs and desires (and other mental states) with propositional content and represent how they represent the world from their own perspectives. The tasks requiring participants to comprehend that agents perceive the world from their unique subjective perspectives, which may differ from those of others and may also be inaccurate, constitute compelling evidence for this ability (Bennett, 1978; Dennett, 1978). In light of this idea, False Belief tasks (hereafter FB task/s) have become the most extensively used tool to tap ToM abilities, and they have been regarded as the litmus test of ToM for decades. FB tasks assess subjects' understanding of a character's false representation by explicitly asking them to predict the character's behavior (e.g., Wimmer & Perner, 1983). In these tasks, subject learns that the agent believes that p and desires q; in reality, p deviates from the subject's own representation of reality; from p and q, the subject infers that the agent will x and therefore should anticipate x.

Expanding on this reasoning, Wimmer and Perner (1983) devised the *change-of-location* task that could be administered to young children in order to test their ToM understanding. The task unfolds as follows: An agent places an object in a box and leaves. While the agent is away, the object is relocated to another box (mostly by a second agent). Then the agent returns, and the children are asked to anticipate where the agent will search for the object. To accurately forecast that the agent will look for the object in the original location, namely the first box, children should attribute a false belief to the agent and meta-represent the agent's misrepresentation (i.e., "Object is in box 1").

Another paradigm that follows a similar rationale, but a different protocol, is the *unexpected content*s task (Hogrefe et al., 1986; Perner et al., 1987). In this task, children are shown a descriptive and familiar box (e.g., a candy box) and asked to guess what is inside the box ("candy"). The experimenter then discloses that the box actually contains a crayon. To evaluate their comprehension of subjective representations, children are then asked what an unknowledgeable third person would say the box contains if he or she has never seen the inside of the box. In another adaptation, children are asked what they themselves initially assumed was inside the box (Gopnik & Astington, 1988).

Numerous studies conducted over the past few decades using the standard FB tasks have yielded interestingly consistent patterns of results (Rakoczy, 2017): First and foremost, it has been found that children up to about four years of age tend to struggle with comprehending the false beliefs of agents, leading them to provide systematically incorrect answers in these tasks, i.e., they predict that the protagonist will act based on reality rather than their mistaken beliefs. Starting from age four, children provide systematically correct answers (i.e., they predict that the protagonist will act based on their mistaken beliefs rather than reality) (e.g., Wellman et al., 2001). This pattern has been dubbed *the 4-year (theoretical) revolution* (e.g., Perner, 1992; Rakoczy, 2022) and found to be very consistent across different languages and cultures (e.g., Liu et al., 2008) Secondly, children's performance in these FB tasks and related tasks that require an understanding of misrepresentation and divergent perspectives (e.g., visual perspective-taking tasks that require an understanding that the same object can appear differently to different observers who have different perspectives) emerge simultaneously: these seemingly divergent tasks provide highly consistent findings as evidenced by strong inter-task correlations (e.g., Perner & Roessler, 2012). Additionally, it has been found that the ability to complete these tasks is reliant on domain-general socio-cognitive abilities such as executive function and language (e.g., Astington & Jenkins, 1999; Devine & Hughes, 2014). Altogether, these findings point to the emergence of a unified conceptual capacity underlying all of these tasks -namely, the acquisition of an understanding of representations or metarepresentation- in the preschool years based on central cognitive and linguistic resources. However, these resources are not the only predictors of ToM, and ToM understanding does not emerge around four years of age out of nowhere. Instead, there are precursors to the 4-year revolution in ToM understanding.

2.2 Earlier Forms of ToM: What precedes FB understanding?

Social perception and engaging with others start in infancy, right after birth. For example, new-born infants have been found sensitive to faces and biological movements (e.g., Bushneil et al., 1989; Simion et al., 2008). Starting from two months of age, infants display turn-taking patterns in vocal interactions with their caregivers (e.g., Rochat et al., 1999) and imitate others' facial gestures (e.g., Meltzoff & Moore, 1977; 1983; but see Oostenbroek et al., 2016; Oostenbroek et al., 2018 for contradicting findings). However, none of these social perception behaviors can be considered as proper ToM, even in the broadest sense, since they do not involve or require attributing mental states. Infants start to show rudimentary forms of ToM understanding just before their first birthdays, or more specifically around 9 months, in the form of perception-goal psychology (Wellman, 2011): they start to follow the perceptions and goals of other individuals and form expectations about how they will act to achieve their objectives based on their perceptual understanding (e.g., Brooks & Meltzoff, 2002; Gergely & Csibra, 2003; Sodian et al., 2007; Warneken & Tomasello, 2006; 2007; Woodward, 1998). This milestone has sometimes been dubbed *the 9-month revolution* (Tomasello, 1999).

Although the 9-month revolution involves attributing simple mental states -such as perceptions and goals, and understanding rational actions based on these states- the perception-goal psychology acquired at this age is still limited compared to fully-fledged meta-representational ToM, which requires an understanding of mutually incompatible representations, false beliefs, and aspectuality, as well as diversity or difference (Rakoczy, 2022). This distinction between early perception-goal psychology and the later, fully-fledged ToM is clearly illustrated by level-I versus level-II perspective-taking abilities (Flavell et al., 1981; Flavell et al., 1986). Children acquire level-I perspective-taking, i.e., *what* is seen by different agents in a given scene may differ between agents, within the scope of perception-goal psychology as this ability requires only attending to the simple perspectival differences or diversity of agents. However, they progress to level-II perspective-taking, i.e., *how* something is seen by different agents in a given scene may differ between agents, only later when they acquire belief-desire psychology or meta-representational ToM, since this ability requires them to understand that different agents can have incompatible perceptions of the same thing (Surtees et al., 2012).

At this point, it is clear that basic and more rudimentary forms of ToM are present way before age four, even in the first year of life. These abilities are potential precursors of the subjective, meta-representational ToM. However, they are conceptually different from the fully-fledged ToM abilities; therefore, they cannot indicate the possibility of proper mental state understanding before age four. More recent research in ToM literature, though, has looked into the possibility of an earlier understanding of meta-representational ToM, or belief-desire psychology. These studies utilized non-verbal implicit FB tasks focusing on the behaviors of participants, instead of verbal responses. In the next chapter, I will go into the details of the so-called implicit ToM tasks and their findings, as well as the theoretical views on these findings.

3 Implicit ToM

The consensus that the meta-representational, subjective mental state understanding emerges and develops in the preschool years has been challenged by newer findings in the literature: studies assessing FB understanding with non-verbal, implicit measures revealed that some form of FB understanding is present before four years of age, even in the first or second year of life (see for a review, Scott & Baillargeon, 2017). More specifically, looking time paradigms and interactive measures employing the classical change of location scenarios have shown that even infants are somewhat sensitive to the mistaken beliefs of other agents as shown by their spontaneous looking behaviors and their behavioral interactions with other people, such as their helping actions (e.g., Buttelmann et al., 2009; Clements & Perner, 1994; Onishi & Baillargeon, 2005). In the following, the most common -*standard*- implicit FB tasks will be introduced, and their findings will be summarized.

3.1 Implicit Measures of ToM

3.1.1 Violation of Expectation (VoE)

In the classic VoE paradigm, participants are shown a change-of-location scenario; however, instead of asking for a verbal answer about where the agent will look for the object. The VoE focuses on the gaze duration during the agent's belief-congruent (i.e., approaching the box where the object was initially located) versus belief-incongruent (i.e., approaching the box where the object is currently located) actions. It has been shown that even 15-month-old infants showed extended gaze duration in the belief-incongruent condition compared to the belief-congruent condition (e.g., Onishi & Baillargeon, 2005; Surian et al., 2007). This finding has been interpreted as indicating surprise upon the agent's false-belief-incompatible action: infants expected the agent's action to be consistent with the agent's mistaken belief, thus they expected the agent to reach for the empty box, and they looked longer when the agent reached for the correct box.

3.1.2 Anticipatory Looking (AL)

In AL paradigms, again change-of-location scenarios are used. However, different than the VoE, in AL measures, the agent does not approach to any location but simply reappears in the scene; then the participant's gaze is investigated to see if the participants show anticipating saccades, i.e., if their gaze is directed at the empty box that held the object in the beginning as the agent believes the object to be in that box. The findings from this paradigm showed that even 25-month-old infants were able to anticipate the actions of another individual based on their mistaken beliefs, and they looked at the initial location of the object upon the agent's return (e.g., Southgate et al., 2007; Surian & Geraci, 2012).

3.1.3 Helping Paradigm

In this task, children are presented with a change-of-location scenario in a helping context: the agent returns after the object has been moved to the second box and approaches the first box; she attempts to open it but becomes unsuccessful and asks the child for assistance. Findings from these studies showed that, in this condition, even 18-month-old infants opened (or attempted to open) the second box and gave the object to the agent. In the control condition, where the agent observed the change of location, children opened the first box instead. This demonstrates that children, in the FB condition, understood that the agent had a mistaken belief about the object's location and attempted to open the wrong box in order to get the object. In contrast, when the agent correctly believed that the object was in the second box, children understood that she really desired to open the empty box (e.g., Buttelmann et al., 2009; Knudsen & Liszkowski, 2012; Southgate et al., 2010).

3.2 Implications of the Implicit Measures

The findings from the implicit FB measures have collectively shown that children are sensitive to belief-involving situations way before age four. However, the nature of this sensitivity and early belief reasoning is still debated. On the one hand, some researchers viewed the results as evidence for the claim that traditional FB tasks underestimated performance due to their high cognitive and linguistic demands. It has been suggested that children under four years of age are indeed capable of attributing false beliefs, and they can show their true performance once the superficial performance factors (i.e., high linguistic and cognitive demands) are eliminated from the FB tasks (e.g., Leslie, 2005; Scott & Baillargeon, 2017). On the other hand, skeptical views interpreted this early capacity as a mere sensitivity to beliefs, but not as the subjective, full-blown understanding measured by traditional (verbal) FB tasks (e.g., Apperly & Butterfill, 2009; Low et al., 2016). And finally, even more skeptical accounts explained children's reactions and behaviors through low-level behavioral cues, such as the direction of attention or preference for novelty (e.g., Heyes, 2014a; Perner & Ruffman, 2005). In the next section, these accounts will be discussed in more detail.

3.3 Theoretical Accounts on the Implicit Measures

The early research on ToM has long shown that children cannot pass the explicit verbal FB tasks until they are four years old. However, studies using implicit measures have shown that even much younger children show sensitivity to the mistaken beliefs of others. The discrepancy between the findings of these explicit and implicit measures calls for further explanation: why do children younger than four fail explicit FB tasks while even infants succeed in implicit FB tasks? Several different theoretical approaches have been proposed to integrate these results and answer this question. In this section, three of those will be examined in more detail: a modularist nativism account, a skeptical and deflationary submentalizing account, and a more middle-ground dual-system account.

3.3.1 Nativist Account

The nativist approach (e.g., Baillargeon et al., 2010; Leslie, 1994; Carruthers, 2013) assumes that domain-specific ToM mechanisms are present early in infancy so that young children already have the necessary concepts or core knowledge to represent the mental states of others. As children develop, the operation of the system becomes more efficient and organized, but representational capacities do not fundamentally change. According to some versions of this view, ToM takes on a domain-specific modular form that enables infants to explain actions by automatically and swiftly attributing mental states. This is realized by a module specialized for ToM understanding, which is usually called the ToM module or ToMM (e.g., Fodor, 1992; Leslie, 2005). It is suggested that this competence is obscured in explicit FB tasks due to extraneous performance factors in explicit tasks, such as the amount of information that needs to be processed (e.g., Baillargeon et al., 2010). Once children become able to handle these extraneous demands around four years of age through developing linguistic and cognitive capacities, they start to show success in standard verbal FB tasks too. Before that, their competence is revealed by implicit tasks that are free from interfering performance demands.

3.3.2 Submentalizing

The submentalizing account, proposed by Heyes (2014a), challenges the idea of a dedicated ToMM and instead suggests that social cognition arises from more domain-general cognitive processes that "simulate the effects of mentalizing in social contexts" (Heyes, 2014a, pp. 131). Heyes proposes that individuals use associative low-level learning mechanisms to make predictions about the behaviors of others, rather than explicitly inferring mental states. This submentalizing account suggests that social cognition may be more continuous across species and development, rather than being a distinct, innate module. Heyes argues that low-level behavioral cues, such as gaze direction and action observation, can account for the apparent mindreading abilities observed in traditional ToM tasks. For example, according to this view, the findings revealed by VoE and AL paradigms can be explained by perceptual and imaginal novelty (Heyes, 2014b). For example, when infants see an object being moved from one location to another, they form a perceptual representation of the object in the new location. However, when the agent returns and searches for the object in the new location, infants may find this perceptually novel and surprising, leading them to look longer at the agent's search behavior. This explanation then deflates the violation of expectation and mentalizing explanations for the longer looking times as it suggests that children's early success on FB tasks may not require them to fully -or even partly- understand mental states, but rather to track and predict others' behaviors based on observed events.

3.3.3 Two-System Accounts

Two-system accounts of ToM constitute a more moderate view than the aforementioned accounts, and they integrate the findings of the implicit and explicit tasks within two distinct systems: a limited but efficient, and evolutionarily and ontogenetically more ancient system (System I) which accounts for the findings of the implicit tasks; and a later-developing, complex but more flexible system (System II) which accounts for the full-blown performance in explicit FB tasks (e.g., Apperly & Butterfill, 2009; Butterfill & Apperly, 2013; Low et al., 2016). According to this theory, efficiency and flexibility work at each other's expense. More specifically, on the one hand, we sometimes reason about a person's belief automatically without further elaborating on her perspective, which might be highly efficient for quick and spontaneous -but not very flexible- responses. On the other hand, we sometimes show flexibility and ascribe mental states more adaptably and explicitly by considering all the available information.

The first system relies on simple representational capacities, which provide efficiency at the expense of flexibility. The second system relies on more complex capacities and, as a result, becomes less efficient but more flexible. The first system is already present at birth or very early in infancy, while the second system emerges around four years of age once children develop sufficient meta-representational skills and have better linguistic and cognitive capacities. In this approach, two systems can operate in tandem in adults and children who are older than four years of age, depending on the requirements of a certain situation. To be more specific, System I has signature limits: it cannot process propositional attitudes. Namely, it cannot track *how* an agent perceives reality, and it fails to resolve under which description the agent represents something. Rather, this simple system operates on belief-like states which are relational and can only track *what* an agent can or cannot perceive (Low & Watts, 2013). Hence, the first system is limited to level-I perspective-taking (Flavell et al., 1981), which is sufficient to guide the looking and helping behaviors in implicit FB tasks. In contrast, System II is more flexible, and it operates on proper propositional attitudes, meaning it can reason about how an agent perceives something and whether she represents reality in a different manner. Therefore, this system is needed to show success in level-II perspective-taking (Flavell et al., 1981) and explicit FB tasks.

3.4 Replicability of Implicit Measures

The findings from the original implicit FB tasks have sparked interest and curiosity in ToM literature and led to a number of direct and conceptual replication studies (e.g., Király et al., 2018; Yott & Poulin-Dubois, 2012). However, a growing body of published and unpublished studies has failed to replicate the original results and raised doubts about their reliability (e.g., Dörrenberg et al., 2018; Kulke & Rakoczy, 2018; Kulke et al., 2018a; Kulke et al., 2019; Powell et al., 2018; Wenzel et al., 2020; Yott & Poulin-Dubois, 2016). More interestingly, some of these non-replications resulted from direct replications of the original tasks with no procedural differences but with even bigger samples, while the others were revealed by conceptual replications with minor procedural alterations (e.g., Schuwerk et al., 2018). As a matter of fact, even a self-replication study of the original AL study -a replication attempt conducted by the two of the original authors using the original materials and procedure with a bigger sample- failed to replicate the original findings and did not find evidence for sensitivity to false beliefs in infancy (Kampis et al., 2021).

In addition to the non-replications, some studies have found that the effects found by the original implicit tasks disappeared under more stringent test conditions, which raised questions about the construct validity of these implicit measures too. More specifically, original effects could be replicated, but they then disappeared once confounds had been removed from the design and suitable controls were added (e.g., Phillips et al., 2015). For instance, in a large-scale replication attempt conducted by Kulke and colleagues (2018b), only two conditions were shown to be replicable by multiple AL measures. Then, it was revealed that these two conditions were replicable only due to some procedural confounds (e.g., identical answers in practice and experimental trials, leading to a bias to look at a particular location), and the effects disappeared after these confounds were eliminated.

In addition to the construct validity issues, implicit tasks fail to show convergent validity too: there are no systematic correlations between the different types of implicit tasks that supposedly tap the same ability (e.g., Dörrenberg et al., 2018; Kulke et al., 2018a; Powell et al., 2018). Even the different tasks of the same type do not converge, and the performance on those tasks does not correlate with each other. These results show that the tasks that have been designed to tap the same underlying construct do not operate as expected and fail to provide robust evidence for an early ToM understanding.

3.5 Implicit ToM Measures: State-of-the-Art

The implicit ToM literature now suffers from a replication crisis: There is a serious discrepancy between the original studies, which showed significant effects hinting at an early understanding of false beliefs, versus the replication studies that either failed to replicate the original findings or found very fragile effects. These replication issues are now at the center of heated theoretical and empirical debates. On the one hand, some researchers argue that the initial positive findings of the implicit tasks provide sufficient evidence for the existence of implicit ToM (e.g., Baillargeon et al., 2018). On the other hand, there are more dubious accounts arguing whether the null findings revealed by the replication studies mean *absence of evidence* or *evidence of absence* in terms of implicit ToM (e.g., Poulin-Dubois et al., 2018). These accounts claim that more systematic replication studies are needed before arriving at a conclusion about the existence and nature of implicit ToM. Such an initiative has been started in the last years as a multi-lab project (Schuwerk et al., 2021): the researchers who conducted the original studies with implicit tasks now cooperate with many other labs around the world for a systematic replication attempt of implicit measures. The results of this project would say more about whether infants can really attribute false beliefs. Another line of research that can provide the implicit ToM literature with a more solid foundation is the so-called altercentric bias studies. In the next chapter, I will focus on this bias as a potential new window into the implicit ToM, and I will explore it in detail.

4 Altercentric Bias

Recent research has shown that implicit measures of FB understanding are prone to replication and validation issues; therefore, they cannot provide a reliable measure to tap implicit ToM understanding. However, one line of research still constitutes a promising alternative to tap implicit ToM: altercentric bias or altercentric interference. This socio-cognitive bias refers to the interferences from others' differing perspectives on our judgments and behavior (e.g., Kampis & Southgate, 2020; Kovács et al., 2010; Samson et al., 2010; Southgate, 2020). Even the mere presence of another agent can trigger this bias, meaning the agent's perspective does not have to be relevant to the task at hand. To exemplify, when I am to judge how many objects exist in my visual field, I become slower and more prone to error in my judgment if there is another agent present in the context who has a different perspective on the objects (e.g., I see 4 objects, but she sees 3 objects due to an occluder in her visual field) (e.g., Samson et al., 2010).

Altercentric interferences demonstrate that our own behavior and judgments can be modulated by how (we think) other agents perceive the world. This then indicates that we implicitly represent their perspectives even when those are completely irrelevant or interfering with our own task. These interference effects occur spontaneously and automatically, without intentional focus on the other agent. Therefore, from a theoretical point of view, altercentric biases can reflect implicit ToM processes, possibly more clearly than the typical implicit FB tasks studied so far (i.e., VoE, AL, and interaction tasks).

But why and how are people influenced by others' perspectives, especially when these perspectives are obviously irrelevant? Or in other words, what are the mechanisms behind the altercentric interference effects, and what do they say about the architecture of the minds? It is important to address these questions before dwelling on the empirical work on altercentric interference effects, as the mechanisms behind this bias are intriguing and somewhat counterintuitive (or contrary to the popular opinion about the architectural structure of representations). More specifically, it has been long assumed that our representations are partitioned from others' perspectives, and there is a so-called representational quarantine that separates our own representations from our representations of others' minds. For example, Leslie (1987) argued that the second-order (and meta-) representations are isolated from one's own first-order (or primary) representations, allowing children not to get confused by pretense uses of objects or to attribute any kind of mental state to others, including false beliefs. However, altercentric interference effects suggest the contrary and point in the direction of the lack of representational quarantine. This, of course, has implications about the configuration of representations: Are our own representations about the real state of affairs distinct from our representations of other's mind and perspectives, or do they interfere with each other? If altercentric biases indeed occur as a result of (implicitly) mentalizing others' differing perspectives and then being influenced by those, then the proposed architectural structure of the representations should be reconsidered. But how do others' perspectives (or our representations of others' perspectives) interfere with our own?

So far, this question has been investigated most extensively under the conceptual umbrella of altercentrism hypotheses (e.g., Kampis & Southgate, 2020). This account argues that although many cognitive processes are likely to contribute to altercentric cognition, few of those are particularly important and deterministic. First, altercentric interferences require us to assign value to others' choices in terms of their attentional and motor qualities. Being aware of others' choices has been indeed shown to cause changes in the neural mechanisms that are involved in assigning value (Zaki et al., 2011). Moreover, the fact that the actions of others can affect our own behavior suggests that our mental representations of others may use the same cognitive processes as our representations of ourselves (Kovács et al., 2010). In terms of motor activity, our motor system is involved not only in executing our own actions but also in observing and predicting the actions of others (Southgate & Vernetti, 2014), and through automatic motor mimicry, others' actions can influence our own actions (Brass et al., 2001). This provides support for a mechanism that jointly codes self and other actions and allows the motor representations derived from others to influence our own actions (Decety & Sommerville, 2003). When it comes to joint action, the concept of task co-representation suggests that our own and the other person's tasks are encoded as part of one integrated representation, which can affect our behavior (Sebanz et al., 2005). There are also common neural responses to the same stimuli experienced by self or others, indicating shared mechanisms (e.g., Forgács et al., 2019; Rueschemeyer et al., 2015). According to the altercentrism hypothesis, these examples suggest that if another's perspective is represented in a "quasi-perceptual" format (Westley et al., 2017), it may initially be treated as first-person input and processed by the same cognitive and neural mechanisms as our own perspective.

What are the potential benefits of this mechanism, though? First, this is assumed to be very useful in interpersonal relations: if group members are equipped with the same input, this would presumably facilitate communication and cooperation (e.g., Shteynberg, 2010). It is indeed possible that the altercentric bias originated out of the need to cope with the diversity of human cognitions (Leslie, 1987; Tomasello et al., 1993) and then became an effective tool for overcoming differences between minds. This tool also has developmental and evolutionary advantages. For example, infants make use of altercentric interferences when they are faced with the challenge of identifying the most relevant source of information in their social environment, and when learning about others and the world (Kampis & Kovács, 2022). Letting others' perspectives influence their own judgment and aligning their attention and motor movements with others can guide infants in terms of the information that needs to be acquired and the shared knowledge base that needs to be formed.

The idea that altercentric bias can actually tap spontaneous and automatic perspective-taking abilities and the potential benefits of this bias paved the way for many studies exploring altercentric interferences. As a result, this bias has been studied in various settings with infants, children, and adults, through different measures that include but are not limited to reaction times, error rates, looking times, and mouse trajectories (e.g., Kampis & Kovács, 2022; Kovács et al., 2010; Samson et al., 2010; Surtees et al., 2012; Van der Wel et al., 2014). In the rest of this chapter, I will first review four of these task settings and their findings. Then I will discuss the reliability and replicability of these measures. Finally, I will introduce the theoretical approaches to altercentric interference effects, which vastly differ in their interpretations of the existing findings and the amount of mentalizing attributed to the altercentric bias effects.

4.1 Altercentric Bias Measures

4.1.1 Dot-Perspective Task

In the dot-perspective task, Samson and colleagues (2010) asked adult participants to judge the number of dots that were visible to them or to an avatar who had either a congruent or an incongruent perspective on the dots. In the task, participants were presented with a diagram of a room in which an avatar could be seen looking at a wall. The avatar was positioned in a way that allowed him to see only one wall of the room, as the other wall was behind the avatar. In each trial, a number of red dots were shown on the walls of the room. In the congruent trials, the dots were shown only on the wall at which the avatar was looking, so that the avatar and the participant saw the same number of dots. In the incongruent trials, the dots appeared on both walls, so that the avatar saw fewer dots than the participants. In each trial, participants were required to make either SELF or OTHER judgments. In SELF trials, participants were asked to judge whether they saw a certain number of dots themselves: first, a number was presented on the screen (e.g., 3), and then the room with the avatar and the dots was shown; then participants judged if the number of dots which were visible to them matched the previously shown number by pressing buttons (one button for YES-matched and one button for NO-unmatched answer). The OTHER trials were very similar to SELF trials, except participants were asked to judge if the number shown in the beginning matched the number of dots that were visible to the avatar. The results of this study revealed that participants made more errors and answered more slowly in incongruent SELF trials than in congruent SELF trials. The authors interpreted this result as evidence of an altercentric bias since it showed that the avatar's perspective interfered with the participant's judgments, even though it was not relevant to the task at hand. In addition to the altercentric bias, the OTHER trials also revealed an egocentric bias, i.e., interferences from one's own perspective or knowledge on their judgments about others' perspectives: participants made more errors, and they responded slower in the incongruent OTHER trials than the congruent OTHER trials. This bias will be introduced more comprehensively in the *Conclusions and the Open Questions* section.

The findings of this study have been supported by other dot-perspective task studies conducted with children and adults (e.g., Michael et al., 2018; Surtees & Apperly, 2012).

4.1.2 Number-Verification Task (6-9 Task)

The original dot-perspective task targeted the altercentric biases in level-I perspective-taking situations in which *what* was seen by the participant and the agent differed. Later, a modification of the dot-perspective task has also been used to investigate the altercentric biases in level-II perspective-taking tasks where *how* an object was seen by the participant and the agent differed (e.g., Surtees et al., 2012). In this task, adults and children (6- to 11-year-old) were presented with numbers, instead of dots, which were either ambiguous in the sense that they appeared as different numbers from opposite perspectives (e.g., 6 and 9) or unambiguous in the sense that they appeared the same regardless of the perspective (e.g., 8 and 0). The numbers were projected either on a wall (level-I version of the task) or a table at which an agent was seated (level-II version of the task). In such a situation, the ambiguous numbers that appeared on the table looked differently to the agent and the participant (e.g., 6 from the participant's point of view and 9 from the agent's), whereas the numbers that appeared on the wall and unambiguous numbers always looked the same to the participant and the agent. The rest of the experiment was analogous to the dot-perspective task of

Samson and colleagues (2010): the agent had visual access to only one wall so that the numbers presented on the other wall were invisible to the agent; there were SELF and OTHER trials where participants had to verify the number either from their own perspective or from the avatar's perspective, and participants' reaction times and error rates were measured. The results revealed no interference from the agent's differing perspective on the participant's reaction times (or error rates) in SELF trials of the level-II perspective-taking task; hence no altercentric bias was present. However, it has been shown that participants reacted slower and made more errors in OTHER trials if their own perspective differed from the agent's. This result hints at an egocentric bias. This study revealed similar results for adult and child participants.

It is important to note that a later study (Elekes et al., 2016) using a similar set-up revealed spontaneous perspective-taking for level-II perspective-taking, too, only when the identity of the number was relevant to the task partner. In this study, there were two participants instead of a participant and a computer-generated agent. In one condition, both participants performed the number-verification task; therefore, *how* a number looked was relevant for both participants. In the other condition, only one participant performed the number-verification task while the other participant performed another -irrelevant- task. Results revealed that participants were slower and more error-prone in SELF trials when both participants were supposed to do the number-verification task, but not when the other participant had an irrelevant task.

4.1.3 Object-Detection Task (Smurf Task)

In the so-called object-detection task, Kovács and colleagues (2010) investigated whether the presence of an avatar, who is irrelevant to the task, could trigger altercentric bias when detecting an object. In the task, the participants were presented with animated materials in which a ball disappeared behind an opaque screen, then reappeared, and then either disappeared again (behind the screen) or rolled out of the scene. The screen was then dropped, and the participants were asked to judge whether the ball was still behind the screen or not. Adult participants responded by pressing a button, whereas in the infant version of the task, participants' looking times were measured. In some trials, the participant's expectation was violated (no ball appeared behind the screen, although it was expected based on the ball's latest motion). In those trials involving an expectation violation, participants displayed prolonged reaction (adults) and looking (infants) times, compared to the trials where the anticipation of the participant matched the presence of the ball. More interestingly, a comparably slow reaction and looking times were displayed when an irrelevant avatar (a Smurf) was temporarily absent during the movement of the ball (therefore, had a false belief about the presence of the object) even though the participants themselves witnessed all changes and had a matched anticipation at the end of the task. Based on this finding, Kovács and colleagues (2010) argued that in the trials where an (irrelevant) agent had a mistaken belief about the presence of an object, participants' behaviors resembled the trials where they themselves had mistaken beliefs. The authors interpreted these findings as evidence of an altercentric bias because participants' behaviors were automatically and spontaneously influenced by others' beliefs.

The findings of the object-detection task have later been supported by the conceptual replications conducted in other labs with neurologically atypical groups, as well as neurotypical samples (e.g., Deschrijver et al., 2016; El Kaddouri et al., 2020; Nijhof et al., 2017) and neuroscience studies investigating the brain regions involved in ToM (e.g., Bardi et al., 2017; Kovács et al., 2014; Nijhof et al., 2018).

4.1.4 Mouse-Tracking Measures

Kovács and colleagues' (2010) object-detection task has also been used in a mouse-tracking study where altercentric bias was deduced from participants' mouse movements, instead of reaction or looking times (Van der Wel et al., 2014). In their study, Van der Wel and colleagues (2014) asked adult participants to choose one of two possible locations of an object and investigated participants' mouse trajectories (i.e., the paths drawn by their mouse cursors). This measure was integrated into a version of the object-detection task, which differed from the original task in two regards: 1) there were two visual stimuli in this version -a ball, as in the original study, and a cube (the ball was the target object while the cube played a role as a distractor) and 2) the ball did not roll out of the scene but swapped places with the cube. Similar to Kovács et al. (2010), some trials involved an irrelevant avatar who had either a false or true belief about the ball's final location. Also, as in the original study, some trials violated the participants' expectations since the ball did not appear behind the screen even though it was last spotted disappearing there, causing the participants to have a false belief about the location of the ball in some trials. At the end of each trial, participants heard a tone, and they were asked to choose the current location of the ball with their mouse cursors. The occluder was lowered as soon as the participants moved their cursors upward on the screen, revealing the ball's actual location. Van der Wel and colleagues (2014) found that participants' mouse movements showed a detour in the direction of the incorrect response when they themselves had a false belief about the final location of the object. Surprisingly, the same effect was found in those trials where the participant themselves had a true belief, but the avatar had a false belief, indicating an influence from the avatar's perspective on participants' judgments.

4.2 Implications of the Altercentric Bias Measures

The findings from the altercentric bias measures showed that adults and children become slower and more error-prone in their own judgments in the presence of an agent who has an incongruent perspective of the situation. There have been also other altercentric bias tasks that showed similar sensitivities in infants as young as 18-month-old as revealed by their searching times (e.g., the so-called object-identification task by Kampis & Kovács, 2022). However, as is the case for the implicit tasks, the nature of the altercentric biases and the mechanism behind these interferences are still debated. One question that has been inquired the most is related to whether these tasks indeed show that people are subject to interferences from others' differing perspectives and they mentalize how others perceive the world even though it is irrelevant (e.g., El Kaddouri et al., 2020; Kovács et al., 2010; Qureshi et al., 2010; Samson et al., 2010; Sebanz et al., 2003), or whether there are lower-level behavioral explanations for these interference effects which argue against the need for mentalizing (e.g., Conway et al., 2017; Heyes, 2014a; 2014b; Santiesteban et al., 2014; Vestner et al., 2022). In the next section, these alternative explanations will be investigated in more detail.

4.3 Theoretical Accounts on the Altercentric Bias Measures

A considerable amount of research has been allocated to understand if the altercentric interferences result from implicit mentalizing or if they are a product of domain-general mechanisms that do not depend on mentalizing in social situations but rather on domain-general attentional cues, such as the submentalizing approach. In this section, I will first discuss the alternative theoretical accounts that attempt to explain the altercentric interference effects, and then I will review the research that has been produced by the non-mentalizing accounts.

4.3.1 Mentalizing Account

Mentalizing refers to thinking about one's own or others' mental states and perspectives. Traditionally mentalizing has been thought of as requiring conscious deliberation. However, recently it has been suggested that infants, children, and adults are capable of implicit mentalizing, namely they can represent mental states unconsciously and automatically (Frith & Frith, 2012). Altercentric bias tasks constitute important support for the implicit mentalizing claims as these biases putatively provide evidence for automatic representations of what others see, believe, and intend (e.g., Kovács et al., 2010; Michael et al., 2018; Qureshi et al., 2010; Samson et al., 2010; Sebanz et al., 2003). Further research showed that these effects were stronger in social conditions compared to less social conditions, and they appeared only when the participant knew the avatar had visual access to the stimuli. These findings provide support for the claim that others' minds and perspectives are indeed decisive in these tasks.

To exemplify, Nielsen and colleagues (2015) used the dot-perspective task in social (involving a social agent), semi-social (involving an arrow), and non-social (involving a dual-colored block) conditions. They found altercentric intrusion effects across all conditions, but these effects were stronger for the social condition compared to the semi-social and non-social conditions. Moreover, in another version of the dot-perspective task, Furlanetto and colleagues (2016) included an avatar who wore colored goggles. Participants were told that the goggles were either transparent (i.e., the avatar had visual access to the stimuli) or opaque (the avatar did not see a thing). The results revealed altercentric inferences only in the transparent-goggles condition, but not in the opaque-goggles condition, showing that participants were influenced by the avatar's perspective only when they knew that the avatar had a differing perspective.

If altercentric bias effects indeed occur as a consequence of implicit mentalizing, this would provide tacit support for the theoretical accounts that were suggested to explain implicit ToM findings. On the one hand, evidence from adult studies supports the dual-system accounts (e.g., Apperly & Butterfill, 2009; Butterfill & Apperly, 2013; Low et al., 2016), which argue that humans have two cognitive systems for mentalizing: an early-developing, automatic, quick, and efficient -but limited- system (System I), and a later-developing, controlled, slow but flexible system (System II). According to this account, the altercentric bias effects operate under System I, reflecting limited but efficient implicit mentalizing abilities; but they do not manifest themselves in the tasks that target System II (e.g., level-II PT tasks). This claim has been supported by empirical studies that disclosed altercentric bias in level-I but not in level-II perspective-taking task (e.g., Surtees et al., 2012; Surtees et al., 2016) and also by theoretical reviews (e.g., Schneider et al., 2017).

On the other hand, evidence from infant altercentric bias studies can be evaluated as supporting the nativism accounts (e.g., Baillargeon et al., 2010; Leslie, 2005). These accounts suggest that infants are sensitive to others' minds and they mentalize their perspectives but cannot show this ability yet in the explicit tasks due to the task demands. According to the nativist accounts, altercentric bias effects can actually reflect proper, fully-fledged mentalizing abilities. However, considering the replication crisis in implicit ToM literature, this claim should be approached with caution. Also, studies that revealed an altercentric bias only in level-I tasks constitute counterevidence for this claim as the nativist accounts assume unified ToM abilities, but the altercentric bias studies uncovered signature limits that allow the altercentric interferences to appear in level-I perspective-taking tasks but not in level-II tasks.

Recently, another account -the altercentrism hypothesis- has been offered to explain the altercentric interferences shown by infants (e.g., Southgate, 2020). As mentioned before, this account is similar to nativism accounts in the sense that it expects very early sensitivity to others' minds. It even carries the claim one step further and argues that altercentrism is the default way of experiencing the world for young infants (Kampis & Southgate, 2020; Southgate, 2020). According to this account, infants are influenced by others' perspectives until they develop self-representation, which then allows them to separate their own perspectives from others' perspectives so that they become more egocentric and less altercentric.

4.3.2 Submentalizing

The submentalizing account (e.g., Heyes, 2014a, 2014b), which has been introduced in detail under the theoretical accounts of implicit ToM, has also been proposed to explain the altercentric interference effects in an alternative, non-mentalistic way. According to this account, the findings of the altercentric bias studies do not constitute conclusive evidence for an implicit mind-reading or perspective-taking system, as these findings can be resolved via domain-general cognitive processes such as attentional processes, object recognition, and spatial perception (e.g., Conway et al., 2017; Phillips et al., 2015; Santiesteban et al., 2014). These processes produce the same results as the mentalizing processes but via a different route: they use cognitively nonspecific mechanisms to solve the altercentric bias tasks, instead of perspective calculation, perspective taking, or mentalizing. For example, the effects found in the dot-perspective task can be explained by the spatial orientation of the avatar and the attentional cues it produced. Or in the object-detection task, the positive findings can be due to retroactive inference effects: when the avatar, i.e., Smurf, returns in the last phase of this task after the last appearance of the ball on the screen. All this creates an intensive event that is loaded in terms of the processing demands, which would then create memory problems and cause participants to slow down.

The skeptical submentalizing approach stimulated a wave of replication and validation studies with the automatic-visual perspective tasks. These studies tested potential alternative explanations for the altercentric bias effects by either adding control conditions or altering the stimuli. In the next session, a more detailed overview of these studies will be presented along with their findings.

4.4 Replicability of the Altercentric Bias Measures

A number of experiments have been conducted to test the competing perspective-taking and submentalizing accounts by using mostly the dot-perspective task. In this section, several of these studies will be investigated in terms of how they differed from the original study and what their results indicate.

The submentalizing account has been initially tested by Santiesteban and colleagues (2014) in a version of the dot-perspective task. The researchers claimed that the avatar employed in the dot-perspective task has directional properties due to its position and gaze. They stated that these properties cued the participants in the direction of the stimuli that were visible from the avatar’s perspective, causing interference in the participant’s response. In light of this claim, they tested the mentalizing account against the submentalizing hypothesis by replacing the avatar with an arrow in some trials of the task: they expected the arrow to cause similar interferences as the avatar and the arrow had the same directional properties. Otherwise, the experiment proceeded as in the original study by Samson et al. (2010). As expected, their results revealed similar interferences in avatar and arrow trials, supporting the claim that the altercentric bias is not driven by the avatar’s visual perspective but rather by its canonical orientation.

The study by Santiesteban and colleagues (2014) paved the way for other studies with different controls. For example, in a study by Cole and colleagues (2016), the avatar’s visual field was occluded by a barrier: open barrier allowed the avatar to see the stimuli, and a closed barrier blocked the visibility. The expected consistency effect was found in this study, however, in both conditions, suggesting that the effect was not driven by mental state attribution, and it was simply a result of the directional effect of the avatar. Similar results were found in a study by O'Grady and colleagues (2017), who replaced the stimuli of Cole and colleagues' experiment with less ambiguous materials (e.g., they used clear pictures involving concrete materials, which made the seeing vs not-seeing conditions more distinguishable). In a more recent study, Langton (2018) found similar findings (i.e., no difference between seeing vs not-seeing conditions) even when the avatar was replaced with photo-realistic stimuli or socially co-present individuals. Following a similar logic, some studies moderated the avatars' visual access by using transparent versus opaque goggles (e.g., Furlanetto et al., 2016). Although some studies found that the effects only appeared when the avatar wore transparent goggles, a later study by Conway and colleagues (2017) failed to replicate this finding in both direct and conceptual replications of the task and found similar effects for transparent versus opaque goggle condition.

The seminal control study by Santiesteban and colleagues (2014) has also been the target of some criticism due to its use of arrows. In their discussion of the results, even the authors themselves acknowledge the possibility that they found similar results for avatars and arrows because participants attributed quasi-visual perspectives to arrows. This is not a far-fetched possibility: decades of research have shown that humans attribute psychological properties to inanimate objects (e.g., Heider & Simmel, 1944), and arrows are specifically subject to perspective attributions due to the derived intentionality they display (i.e., they are used for communicative purposes, and they carry a semantic content) (Searle, 1983). This criticism has been targeted in further studies that used other inanimate objects instead of arrows in the dot-perspective task. For example, in a recent study by Vestner and colleagues (2022), the arrow was replaced with desk fans -objects that are known to cue attention. They found that desk fans produced interferences that were similar to the altercentric bias, and they also showed that participants who were more susceptible to the effects caused by desk fans were also more susceptible to the effect caused by human avatars. Due to similar concerns, Westra and colleagues (2021) manipulated how the avatar/object was recognized by using a novel entity paradigm (e.g., Johnson et al., 1998): depending on the participant's belief, the object was recognized as either animate or inanimate. As a result, they found that the altercentric bias effects were not modulated by participants' beliefs about the animacy of the central stimulus, indicating that the errors and the slowdown did not occur due to mental state attribution.

4.5 Altercentric Bias Measures: State-of-the-Art

The picture shown by the altercentric interference effects is somehow complicated and equivocal. On the one hand, there have been studies showing that the self-consistency effects found in the altercentric bias measures were due to implicit mentalizing of others' visual perspectives and beliefs. On the other hand, many direct and conceptual replications found these effects in non-social situations with inanimate objects, indicating that alternative explanations, such as attentional cueing, are at play and causing these effects. Therefore, on the basis of the existing literature, it is very hard to arrive at a final conclusion about the nature of the altercentric bias effects. There have been also some studies that tried to reconcile the two accounts. For example, Marshall and colleagues (2018) argued that the altercentric interferences that occur in response to social stimuli (e.g., an avatar) are driven by perspective-taking mechanisms, whereas the interferences that occur in response to non-social stimuli (e.g., an arrow) are driven by low-level domain-general mechanisms, such as attention cueing. In another non-binary approach, Westra and colleagues (2021) claimed that neither account is fully correct, and they are also not mutually exclusive, so they should be investigated together to understand the mechanism of the altercentric bias better. Finally, a meta-analysis of the dot-perspective task (Holland et al., 2021) found evidence for both implicit mentalizing and directional/attentional cueing hypotheses. However, they also found that the effects of directional/attentional cueing were significantly larger. Overall, these results point to a complicated and incomplete picture of altercentric bias measures. Whether these biases exist, and if they actually reflect truly implicit ToM abilities, are still open questions.

5 Conclusions and Open Questions

Decades of research on explicitly measured meta-representational ToM abilities have consistently shown that children acquire a subjective, fully-fledged understanding of minds around four years of age. This view has been challenged in the last twenty years by the findings of the implicit measures of ToM understanding, e.g., VoE, AL, and interaction paradigms. These intriguing findings have led to various theoretical accounts that aimed to integrate the implicit ToM findings with the years of research on explicit tasks. More recently, though, a growing body of non-replications casted doubts on the reliability and validity of implicit FB measures as they failed to show robust findings with these tasks. On the contrary, altercentric bias measures are still promising, and they can provide an additional -probably more reliable- potential window into implicit ToM abilities. This bias corresponds to the interferences from others' perspectives that are manifested in our own judgments and behavior, even when those perspectives are irrelevant or interfering with our own task.

Altercentric interferences have been studied in different settings with adults, children, and infants. Among these, the most influential task has been the so-called dot-perspective task. The original study that used this task in social cognition research for the first time (i.e., Samson et al., 2010) has shown that humans get slower and more error-prone in their own judgments if there is another agent in the situation, who has a different perspective on the stimuli that need to be judged by the participants. This task initiated later replication attempts and has been at the center of heated theoretical discussions in the last years. To briefly remind, two lines of interpretations emerged from the findings of this measure: mentalizing accounts argued that altercentric interferences are the consequences of implicit mentalizing of others' perspectives, whereas non-mentalizing alternatives such as the submentalizing account claimed that these interferences are due to lower-level domain-general processes such as attentional/directional cues. Different versions of the dot-perspective task provided mixed evidence for these views and rendered the dot-perspective task a not-so-reliable measure of the altercentric bias effects. However, the mixed findings also made it difficult to entirely rule out the possibility of genuine altercentric interferences that occur due to implicit mentalizing. Indeed, many studies tried to incorporate the positive and null findings in more nuanced theoretical accounts.

As a result, there is still a need for a reliable altercentric bias task so that further claims about the scope, limits, and foundations of the altercentric bias (and implicit ToM) can be tested. In this case, the obvious solution would be to turn to other tasks that have been used to tap altercentric interferences so far. However, these tasks are either vulnerable to the previously mentioned concerns (e.g., object-detection task) or have been yet tested regarding their reliability and validity (e.g., object-identification task). Therefore, a novel methodological approach to the altercentric bias is warranted.

Such an approach may be found in related literature that focuses on a different kind of sociocognitive bias, namely the egocentric bias. Egocentric bias refers to the interferences from our own perspective and knowledge on our judgments about others' perspectives or our understanding of others' beliefs, even when we are explicitly asked to focus on the third person and judge *their* perspective (e.g., Apperly et al., 2007; Apperly et al., 2011; Back & Apperly, 2010; Bernstein et al., 2011a; Birch & Bloom, 2004; 2007; Chambers et al., 2008; Keysar et al., 2003; Samson et al., 2010; Sommerville et al., 2013). Egocentric interferences are thus negative indicators of explicit perspective-taking abilities: stronger egocentric bias means more limited perspective-taking or vice versa.

Egocentric biases have several advantages over standard explicit tasks, e.g., FB tasks. For instance, these biases provide continuous, rather than binary, measures of perspective-taking. More specifically, the standard FB tasks capitalize on pass/fail responses, whereas the egocentric biases are tapped through more fine-grained continuous measures. Moreover, egocentric bias tasks easily lend themselves to implicit versions where altercentric biases can be tapped. Namely, the egocentric bias tasks allow us to modify the task in a way that the two biases could be measured within the same task format, with only small -but crucial- differences in task prompts (e.g., asking about the other-perspective in the egocentric bias condition and asking about the self-perspective in altercentric bias condition). This has two advantages. First, it would create diverse task formats that target altercentric bias. Considering the issues with the existing altercentric bias tasks, this would then increase the chances of finding a robust altercentric bias measure. Second, tapping both biases within the same task format would allow us to measure explicit and implicit ToM abilities by using analogous versions, which would eliminate the superficial differences between explicit and implicit tasks. This would then render the performance factors irrelevant and enables more straightforward comparisons of these abilities and clearer investigations of their relation to each other.

Some tasks already tested both biases together, e.g., the dot-perspective task. However, as discussed above, this task has been shown to be susceptible to reliability and validity issues.

Therefore, different measures are needed. Given that egocentric bias measures can be adapted to tap both biases together in elegant and analogous ways, we then search for the candidate task within the egocentric bias literature. Therefore, the first question of the current work is focused on which egocentric bias (or explicit perspective-taking) task would constitute a more suitable alternative for an altercentric version adaptation (and consequently for tapping both biases together).

If a satisfying answer to this question can be revealed, meaning if altercentric biases are reliably gauged by modifying egocentric bias tasks, the assumptions and theoretical accounts pertaining to altercentric interferences can be tested further by using this reliable task. For example, we can then investigate if altercentric biases are indeed products of implicit mentalizing or do they occur because of lower-level domain-general processes. If the former, this would then bring back the discussion about the representational architecture of the mind: Are the representations quarantined, or do they interfere with each other? Furthermore, if altercentric bias effects indeed occur due to representing others' perspectives automatically and spontaneously, this would mean that they naturally tap implicit ToM abilities, and they can be then used to further investigate the implicit ToM reliably. This would enable us to explore the origin, developmental trajectories, and foundations of implicit ToM as well as to test the theoretical accounts of implicit ToM against each other (e.g., nativism versus two-systems account).

6 Aims of the Dissertation

ToM literature needs tasks that can reliably reveal implicit ToM abilities. This dissertation aims to address this gap. To this end, the present work capitalizes on altercentric bias as a means to tap implicit ToM. Given the issues with the existing altercentric bias tasks, the current work brings a different approach. It aims to devise or adapt new altercentric bias tasks by modifying several existing egocentric bias (or explicit perspective-taking) measures. With adult proof-of-concept studies, this dissertation separately and systematically tests these task formats in terms of their suitability and validity concerning altercentric biases and their reliability and replicability concerning egocentric biases. If this initial aim is accomplished, meaning once a task is shown to tap altercentric biases reliably, further investigations using this task will target the theoretical assumptions of the altercentric bias and implicit ToM. More specifically, whether altercentric biases result from implicit mentalizing will be investigated, and the theoretical assumptions of the implicit ToM will be tested.

The aim of the Project 1 was to test the so-called Sandbox task as a potential measure of altercentric bias. This task has been initially developed to tap egocentric biases, and it capitalizes on a standard FB task but with continuous- rather than discrete- locations (e.g., Begeer et al., 2012; Bernstein et al., 2011b; Coburn et al., 2015; Mahy et al., 2017; Sommerville et al., 2013). In Project 1, we aimed to test the replicability of the earlier egocentric bias findings with a novel online version of this task and investigate if this task can reveal altercentric interferences. In this project, one study also aimed to reveal potential cross-cultural differences between more independent German- and more interdependent Turkish-speaking adults in terms of the amount of egocentric and altercentric biases they display. The second study conducted within the scope of Project 1 investigated if incorporating mouse-tracking measures into the Sandbox task could reveal the implicit processes more reliably. Finally, the third study of this project investigated if previously found altercentric bias effects were due to general switching costs instead of mentalizing. This project also involved two baseline studies where adult participants were tested on a previously used live version of this task (i.e., the paper-pencil version). These baseline studies aimed to control if there were differences between in-person versus online versions of the task.

Project 2 aimed to test if the so-called Director Task could be adapted to tap altercentric biases. Director Task is a referential communication task in which participants are asked to move or find objects according to a semi-knowledgeable third person's (i.e., the director's) directions. This task was initially used to reveal the effects of mutual- versus self-knowledge in social communication situations, and then it has been utilized as an indicator of egocentric biases in adults and children (e.g., Cane et al., 2017; Dumontheil et al., 2010; Keysar et al., 2000; Keysar et al., 2003; Legg et al., 2017; Samuel et al., 2019; Symeonidou et al., 2016). The first study of the Project 2 aimed to replicate the previously found egocentric bias effects in a completely online version of this task and test if the Director Task can be adapted to tap altercentric interference effects too. The second study conducted as a part of Project 2 aimed to self-replicate the egocentric bias results found in Study 1 and further investigate the nature of the altercentric interferences found in this task. More specifically, this study aimed to investigate whether altercentric biases are the product of spontaneous mentalizing in social situations but not low-level situational factors. This investigation compared the interferences shown in social (with an irrelevant agent present) versus non-social (without any agent) versions of the altercentric bias condition. If altercentric biases are found only in the social version, this would indicate that an irrelevant agent triggered an interference, and the effect is not solely due to the task demands. However, if similar effects are found in both social and non-social versions, this would imply that the interferences are not altercentric but they result from situational demands.

Finally, Project 3 measured altercentric biases by using visual perspective-taking tasks. These tasks are explicit measures of (lack of) perspective-taking abilities: they ask the subjects to report what is visible to a person who has a divergent perspective (level-1 perspective-taking) and how something is seen from a different perspective (level-2 perspective-taking) (e.g., Flavell et al., 1981; Flavell et al., 1986). These measures are generally based on binary responses. However, with the addition of behavioral data such as reaction times and mouse trajectories, they can become continuous and reveal potential interferences in fine-grained ways. The first study of Project 3 aimed to reveal robust and reliable altercentric interference effects in visual perspective-taking tasks using different measures such as error rates, reaction times, and mouse trajectories. This study also included the egocentric bias version of the task as a more explorative investigation. Another aim of this study was to investigate the assumptions of different theoretical accounts on implicit ToM. Whether and to what degree altercentric biases are found in different levels of perspective-taking tasks has theoretical implications.

On the one hand, two-system accounts expect biases only in level-I tasks capturing automatic System-I but not in level-II tasks tapping effortful System-II. Therefore, the results would support the dual-system approaches if significant biases are found only for level-I tasks. On the other hand, nativism accounts assume unified perspective-taking abilities and predict no differences between the two levels of the task. Therefore, if biases occur in both level-I and level-II tasks, this would support the nativist approaches. There is indeed some research supporting the former alternative (e.g., Surtees et al., 2012; Surtees et al., 2016); however, that line of research almost exclusively depended on the dot-perspective task, which is questionable in terms of its reliability and validity. The second study of this project aimed to replicate the altercentric interferences found in the first study and also tested a non-social version of the altercentric bias task with a similar rationale as the second study of Project 2, i.e., in order to ensure that the altercentric biases revealed by this task triggered by the presence of an agent, but not situational or attentional factors.

7 Empirical Findings

In this section, I will summarize the findings of three projects that I conducted in the course of this dissertation: 1) Haskaraca, Proft, Liszkowski, & Rakoczy (2023) – How robust are egocentric and altercentric interference effects in social cognition? A test with explicit and implicit versions of a continuous False Belief task; 2) Haskaraca, Proft, Liszkowski, & Rakoczy (2023) – Testing Egocentric and Altercentric Biases in a Referential Communication Task; 3) Haskaraca, Proft, Liszkowski, & Rakoczy (2023) – Measuring Altercentric Biases in Level-1 & Level-2 Perspective-Taking Tasks by the Means of Mouse-Tracking. In the following, I will focus on the experimental design and the main results of the studies. For further details (e.g., participants, design and procedure, analyses, and detailed results), please refer to the original manuscripts (Appendices A, B, and C).

7.1 Project 1: Haskaraca, Proft, Liszkowski, & Rakoczy (2023)

The current study investigates if the so-called Sandbox Task can be a way to measure altercentric interferences reliably. The Sandbox task was initially developed to measure egocentric bias in a modified standard false belief scenario with continuous -rather than discrete- locations: when asked where an agent believes an object is, participants deviate from the correct answer somewhat in the direction of where they themselves know the object to be –and the degree of deviation reflects the magnitude of the egocentric bias (e.g., Sommerville et al., 2013). More specifically, in this task, participants witness a false belief scenario where an agent hides an object at Location 1 in a continuous area, such as a sandbox. However, then, the object is relocated to Location 2 (in the same sandbox) by another agent in the absence of the first agent. Participants are then asked to indicate either where the first agent believes the object is (experimental trials) or where the agent hid the object before relocation (control trials). If participants are affected by their own perspective or knowledge when judging others' behaviors (in experimental trials), we expect their answer to be somewhere around Location 1 (correct answer) but biased toward Location 2 (the object's actual location). The deviation is expected to be smaller or absent in the control trials.

One advantage of this task is that it can be easily adapted to measure altercentric interferences: it requires no change in the task material but only one slight alteration in the task question. More specifically, instead of the agent's belief, in the altercentric bias condition, participants are asked where the object currently is. Altercentric interference effects suggest that participants would deviate from the correct answer in the direction of where the agent believes the object to be –and the degree of deviation reflects the magnitude of the altercentric bias. To be more precise, in this version, participants again see a similar scenario but are asked a different question: namely, they are asked to indicate where the object really is. In the experimental trials, agents have a false belief about the object's current location, and agents' and participants' perspectives differ. In the control trials, agents witness the relocation; thus, the agents' and participants' perspectives are the same. If participants engage in implicit perspective-taking and represent the agent's mistaken -although irrelevant- perspective, their judgments are expected to become subject to altercentric interferences such that their answers deviate from the object's actual location (Location 2) toward the agent's belief location (Location 1) in experimental trials. The deviation is expected to be smaller or absent in the control trials.

In three studies, both versions were tested with adult participants using online versions of the original and adapted versions of the Sandbox task. In Study 1, 94 German and 94 Turkish adults were tested on either the egocentric or altercentric bias version. They also completed a self-construal scale, which informed us about how independent and interdependent the individuals from different cultures were. Results revealed neither an egocentric nor an altercentric bias in either group. Also, German and Turkish adults did not differ from each other in terms of the magnitude of the biases. Turkish adults were shown to be more interdependent; however, the level of interdependence (and also the independence) never correlated with any bias.

In Study 2, we made changes to the materials so that they would become more engaging and also to the platforms where the study was running and advertised so that fewer technical limitations could be expected. In this study, we also incorporated the mouse-tracking measures into the Sandbox task to have more spontaneous behavioral measures. Mouse trajectories have previously been used to document altercentric bias (e.g., Van der Wel et al., 2014). When subjects were asked to move their mouse cursors to the target object's location, they took a little detour on their way to their answers when another agent in the scenario had a belief that differed from their own. However, when the participant and the agent shared a belief, participants followed a more direct route while moving their mouse cursors to mark the target location. Therefore, the area between the detour in the direction of the wrong answer and the direct line from the starting point to the target location indicated whether and to what degree participants engaged in altercentric bias (Van der Wel et al., 2014). In the current study, we included this measure (in addition to the final location measure inhabited in the Sandbox task) for both biases. We investigated if participants' mouse movements approached the wrong location more in the experimental trials compared to the control trials. Ninety-four German- and 94 English-speaking participants were tested in this study. These groups were targeted due to their availability in the study distribution platform, and no cross-cultural difference was expected (and analyzed). The results revealed neither an egocentric nor an altercentric bias for neither group (i.e., German- or English-speaking) and with neither measure (i.e., final location or mouse trajectories).

Study 3 aimed to investigate one possible explanation for the consistent null results: Socio-cognitive biases may be sensitive to test designs and procedures and reveal themselves only under specific circumstances. One such design could be within-subject studies where participants are tested on both biases in blocks. Existing evidence is compatible with the possibility that altercentric biases only arise in such mixed-block designs if they are presented after the egocentric bias version (Furlanetto et al., 2016; Speiger et al., 2022) but not in single-block designs like the one used in Study 1 and Study 2 (e.g., Conway et al., 2017). Study 3 investigated this possibility with 54 German-speaking adults who were tested on both biases in a within-subject design. Altercentric and egocentric bias versions were presented in blocks, and the order of blocks was counterbalanced across participants. The results revealed neither an egocentric nor an altercentric bias, and the presentation order of the blocks did not have an effect.

In addition to the three main studies, two in-person baseline studies were also conducted within the scope of Project 1. These baseline studies aimed to replicate earlier Sandbox task results (which revealed egocentric interference effects) using the previously validated paper-pencil version of the task and the original materials from this version. We tested 54 German and 54 Turkish adults in two separate studies employing the paper-pencil version of the Sandbox task. The results failed to replicate the original findings and revealed no egocentric bias. Overall, these findings raise critical issues regarding the replicability and validity of the Sandbox task as a measure of egocentric and altercentric biases.

7.2 Project 2: Haskaraca, Proft, Liszkowski, & Rakoczy (2023)

The current study investigates if the so-called Director Task can be a way to measure altercentric interferences. The Director Task was initially developed to reveal heuristic strategies in communication and then also used in social-cognition research to tap egocentric biases. In this task, participants are instructed to select and move objects on a grid shelf consisting of multiple compartments. The grid is placed between the participant and a computer-generated avatar -the director- who gives the instructions. Due to several occlusions in the grid, the director has a restricted view of the objects. Hence, an instruction from the director to select an object ("click on the biggest box") in the grid is ambiguous if the instruction cannot refer to the actual biggest box: it is visible from the participant's side but invisible from the avatar's point of view. Across different versions of the Director Task (e.g., in-person and computerized), robust egocentric biases were revealed: participants made more errors and were slower in the ambiguous trials compared to the unambiguous trials where the instructions of the director could describe only one object as it was seen from both perspectives (e.g., Apperly et al., 2010; Dumontheil et al., 2010; Dumontheil et al., 2012; Keysar et al., 2000; Keysar et al., 2003; Legg et al., 2017; Samuel et al., 2019).

The Director Task easily lends itself to altercentric bias adaptations. The current study achieved this by changing the director's side and adding an irrelevant agent to the scene. More specifically, in the altercentric bias version adapted within the scope of Project 2, we aligned the perspectives of the participants and the directors so that participants could act upon any instruction from the director by simply considering their own perspectives. This alignment was achieved by moving the director to the same side as the participants. In this version, participants also saw an agent on the other side of the grid. However, this agent was clearly differentiated from the director (in terms of appearance) and was not mentioned in the study; she simply stood behind the grid. The ambiguous trials in this version corresponded to the ones where the target object was not visible from the irrelevant agent's perspective due to an occluder. In the unambiguous trials, the instructions referred to the objects that were visible from both sides. If participants are influenced by others' differing yet irrelevant perspectives, meaning if they are subject to altercentric biases, they would be slower and more error-prone in the ambiguous trials than the unambiguous trials.

Across two online studies, we tested German-speaking adults on either the existing egocentric or the newly adapted altercentric bias version of the Director Task. The accuracy and reaction times were recorded. In Study 1, 126 participants were tested within the scope of a master's degree work. The results of this study revealed egocentric interference effects: participants made more errors and became slower in the ambiguous trials compared to the unambiguous trials. One intriguing point is that the two reaction time measures used in this study (mean and median reaction times) revealed somewhat different findings: median reaction times were higher for the ambiguous trials than the unambiguous trials, but no difference was found in the mean reaction times. The altercentric bias findings were not as straightforward: participants made more errors in the ambiguous trials, indicating a potential altercentric bias effect; however, they also responded faster in the ambiguous trials, indicating a reverse altercentric bias effect. Mean and median reaction times provided similar results in the altercentric bias version.

In Study 2, we aimed to re-conduct the first study with a more controlled data set, so we used an established data collection platform instead of convenience sampling to reach out to participants. Other than this change, the second study mimicked the first study with an addition of a non-social control condition for the altercentric bias task. In this non-social condition, the irrelevant agent was removed from the scene. This condition aimed to ensure that the altercentric biases were attributable to the irrelevant agent's perspective instead of low-level domain-general situational factors, as suggested by submentalizing accounts. One-hundred-eight German-speaking participants were tested either on the previously used egocentric or altercentric bias versions. An additional 54 participants were tested in the non-social condition of the altercentric bias version. The findings of Study 2 revealed a similar pattern as Study 1. More specifically, in the egocentric bias version, participants made more mistakes and were slower in the ambiguous trials than in unambiguous ones. In Study 2, mean and median reaction times point to the same direction of results. In the altercentric bias version, participants made more mistakes in the ambiguous trials, in which they also turned out to be faster. This puzzling result pattern was also observed for the non-social condition of the altercentric bias version. Also, no main effect of condition was found. These results imply that the contradictory interferences shown in the altercentric bias version of the Director Task are not on account of implicit mentalizing but could be attributed to domain general processes triggered by the task design. Overall, these findings provide further support for the reliability of the Director task as a measure of egocentric biases but raise issues regarding the validity of this task as a measure of altercentric biases.

7.3 Project 3: Haskaraca, Proft, Liszkowski, & Rakoczy (2023)

The current study investigates the presence of altercentric biases in level-I and level-II visual perspective-taking tasks. These tasks have been extensively used in sociocognitive development research in order to explore when children come to understand that what they see may differ from *what* others see in the same context (level-1 visual perspective-taking) and that *how* they and others see the same thing may differ (level-II visual perspective-taking) (e.g., Moll & Meltzoff, 2011; Moll & Tomasello, 2010). For example, in one of the earliest standard measures tapping level-I visual perspective-taking abilities, children are presented with a card displaying two different animals on its two sides. The card is held vertically between the experimenter and the participant. Then the participant is asked which animal is seen by the experimenter. In the level-II version of this task, participants are shown a picture of an animal (e.g., a turtle). The picture is laid horizontally on a table between the experimenter and the participant. Then the participant is asked how the experimenter sees the turtle: on its back or feet (Flavell et al., 1981).

The current study used a similar logic to create computer-based level-I and level-II visual perspective-taking tasks. In the level-I task, adult participants were shown a scene where a computer-based avatar was looking at a table, which was located between the participant and the avatar. On the table, a piece of paper was displayed. There were two symmetrical numbers/letters on the paper (e.g., 0, 8, o, s). In congruent trials, the two numbers were the same. In the incongruent trials, the two numbers were different. Once both numbers were shown to the participant, an occluder dropped into the scene, limiting the participant's and the avatar's visual access to one number. Then participants were asked to indicate which number they were still seeing. In the level-II task used in the current study, a similar setup was utilized. However, for this level, the paper only displayed one number/letter that looked different from opposite points of view (e.g., 6/9, u/n). No occluder appeared in this task. Participants were asked which number they saw on the table. For both levels, they chose among the two options presented on the screen: the correct answer and a distractor (i.e., the number/letter seen from the agent's perspective in the incongruent trials and a number/letter that looked like the correct answer in the congruent trials). Participants' mouse movements and reaction times during the answering phase, as well as the accuracy of their final answers, were used to deduce bias. Altercentric bias effects would predict participants to show bigger detours in the direction of the wrong answer, extended reaction times, and less accuracy in the incongruent trials compared to the congruent trials.

The current study is not merely about developing a task that can tap altercentric biases in a novel format. Instead, whether altercentric biases are found in both levels or only in level-I perspective-taking tasks has theoretical implications too. More specifically, dual-system accounts define the level-II visual perspective-taking as the signature limit for the implicit, automatic, and spontaneous perspective-taking abilities (which are operated by the fast, efficient, but limited System 1). These accounts expect this kind of perspective-taking, such as altercentric interferences, to occur in level-I tasks, but not in level-II tasks (which are solved via the operations of more flexible but less efficient System 2). On the other hand, nativist accounts do not make such a distinction and presume unified ToM abilities across different levels of perspective-taking.

Across two online studies, adult participants were tested on entirely web-based level-I and level-II visual perspective-taking tasks. The first study aimed to investigate whether these tasks revealed any altercentric interference effects and whether the altercentric biases were found for both levels or not. Fifty English-speaking participants were tested in the aforementioned altercentric bias. Study 1 also tested an additional 50 participants on an egocentric bias version of this task, where participants were asked about what the agent saw in similar congruent and incongruent trials. Egocentric bias version aimed to explore whether the current task format could be used to tap both biases analogously. No theoretical assumptions were based on the findings of this version in the current study. Results of the study revealed altercentric biases for both level-1 (as revealed by reaction time and accuracy measures) and level-2 (as revealed by reaction time measure) tasks. Egocentric biases were revealed only for level-II tasks (as revealed by mouse-tracking and reaction time measures).

Study 2 aimed to investigate if the altercentric biases revealed by the present task format were robust and reliable. To do that, a non-social condition has been tested along with the existing (social) condition of the altercentric bias task. In this non-social version, similar to the Director Task, the agent was removed from the scene. This condition aimed to explore if the interferences found in the current study were indeed altercentric in the sense that they occurred due to the presence of a differing -yet irrelevant- perspective or whether the interference effects were byproducts of domain-general interferences that required no mentalizing. If the effects were actually altercentric, then altercentric biases were expected to occur only in the social version. Otherwise, similar patterns were expected in both non-social and social conditions. The current study also scrutinized which measure (i.e., mouse tracking, reaction time, or accuracy) provided a more reliable means to tap altercentric biases in the current task format by investigating which measures continued revealing biases in the second study and which measure showed differential effects across two studies. The within-subject analyses revealed that participants showed altercentric interferences in the level-I (as revealed by the reaction time measure) and level-II (as revealed by the accuracy and reaction time measures) tasks of the social condition, but no bias has been revealed in the non-social condition. Moreover, mixed-effect analyses revealed that the difference between participants' accuracy in congruent versus and incongruent trials depended on the condition: the difference was bigger in the social condition compared to the non-social condition. These results imply that the interferences shown in the current study were indeed due to implicitly mentalizing others' perspectives. Regarding the theoretical assumptions of nativist and dual-system accounts, the current study did not differentiate between the two levels of perspective-taking and revealed biases for both levels. However, it should be noted that the measures employed in the current study failed to reveal consistent effects across experiments. They were also shown to be influenced by other factors, such as what kind of device was used to do the clicking in the study (mouse or touchpad). Therefore, further studies are needed before the findings of the current study can be used to make inferences about different theoretical accounts.

8 General Discussion

This dissertation aimed to develop and adapt novel task formats that can tap altercentric biases reliably, with the final objective of utilizing these tasks in order to study implicit ToM further. To this end, this dissertation tested three different task formats as potential measures of altercentric (and egocentric) biases in three different projects. In the following, I will first present a brief summary of the empirical findings and then discuss the implications of these findings. Then I will relate these findings to existing theoretical work on altercentric bias and implicit ToM. Finally, I will mention the limitations of the current projects and provide suggestions for potential future research that can extend the impact of the current projects and address open questions.

8.1 Brief Summary of the Findings

In three projects, we investigated if existing egocentric bias (or explicit perspective-taking) measures could be adapted in a way that allows us to study altercentric interference effects, too. In Project 1, we devised an altercentric bias version of the so-called Sandbox task, a standard FB task with continuous locations (e.g., Sommerville et al., 2013). In three web-based studies, we tested German-, English-, and Turkish-speaking adults either on the existing egocentric bias version or on the newly adapted altercentric bias version of the Sandbox task (except for one study where participants were tested on both versions). None of these studies revealed a bias, regardless of the type of bias measured. Incorporating more implicit, spontaneous measures into the Sandbox task (e.g., mouse-tracking) did not yield any significant changes to the null outcomes. Finally, two baseline studies were also conducted in order to test the replicability of the previously found egocentric interference effects. In these two baseline studies, German- and Turkish-speaking adults responded to an in-person version of this task (i.e., paper-pencil version). The results consistently indicated a lack of bias.

In Project 2, we adapted an altercentric bias version of the so-called Director Task, a referential communication task which has been used to reveal egocentric biases in communication and social cognition before (e.g., Keysar et al., 2000; Samuel et al., 2019). In two online studies, we tested German-speaking adults on either the existing egocentric or newly adapted altercentric bias version of this task. The results revealed consistent egocentric bias effects, which were reflected in both accuracy and response times of participants. The findings from the newly adapted altercentric bias version revealed more complicated result patterns, where participants showed altercentric interference effects in the accuracy measure but anti-altercentric interference effects in response times. Moreover, the presence or absence of a divergent perspective did not cause a difference in these effects, casting doubt on the altercentric nature of the observed effects found in the new adaptation of the Director Task.

In Project 3, altercentric biases were studied by using classic level-I and level-II perspective-taking tasks (e.g., Flavell et al., 1981), with the addition of mouse-tracking measures (e.g., Van der Wel et al., 2014). Across two studies, altercentric biases have been revealed in both level-I and level-II perspective-taking, but somehow inconsistently: the mouse-tracking measure never revealed an altercentric bias, the accuracy measure indicated differential bias patterns for two levels in different studies, and the reaction times fell short of capturing the effect of condition (social versus non-social) on the biases, which was evident when biases were analyzed separately for two conditions. This project also tested egocentric biases in an analogous task format. However, egocentric bias version was abandoned after the first study because the results indicated unexpected bias patterns across two levels, which probably were caused by confounding factors rather than (lack of) perspective-taking.

8.2 Implications of the Current Null Findings for Altercentric (& Egocentric) Biases: *Absence of Evidence* or *Evidence of Absence*?

As summarized above, the three tasks used in our studies revealed very different findings in terms of egocentric and altercentric biases. The first project employing the Sandbox task revealed no bias at all, the second project employing the Director Task only revealed egocentric biases, and the third project employing the level-I/level-II perspective-taking tasks only revealed altercentric biases. Do these results mean that if one type of bias has not been revealed by a task format, this format should be given up as a potential measure of this bias? Or is it possible that the absence of evidence simply reflects false negatives? In the following, each task will be scrutinized separately to find out if they failed to provide evidence for a type of bias because they were not suited to tap this bias or because other factors were in effect.

8.2.1 Sandbox Task

In our studies, the Sandbox task repeatedly revealed null results for both egocentric and altercentric bias versions. This result pattern is especially intriguing in the case of egocentric bias as the Sandbox task has been repeatedly shown to reveal egocentric interferences in the original studies, and not just for children but also for adults too (e.g., Begeer et al., 2012; Bernstein et al., 2011; Coburn et al., 2015; Mahy et al., 2017; Sommerville et al., 2013).

General Discussion

However, our results were not the first to challenge the positive findings of the original studies. Some relatively recent replication attempts with the Sandbox task (Samuel et al., 2018a; 2018b) also either failed to find any egocentric interference effect with this task or they found that the effects were attributable to a general difficulty about false representation reasoning rather than false belief understanding. For example, Samuel and colleagues (2018a) have found that participants showed similar egocentric interferences when they were asked to indicate where a false film would depict an object versus where a protagonist with a false belief would look for the object. The findings of the current study are more compatible with the findings of these unsuccessful replication attempts of the Sandbox task as a measure of egocentric biases.

But why is there a contrast between the results of the original Sandbox studies and later replication attempts? At this point, a methodological confound requires our attention: the format of the task. The original Sandbox studies and the later replication studies differ from each other in terms of the task format. The earlier studies always used a lab-based, in-person versions of this task (e.g., with real-life sandboxes or paper-pencil materials); whereas the later studies were computer-based and sometimes completely online (as was the case in our studies). It is possible that the computer-based online formats were less sensitive in capturing egocentric bias effects, leading to the differences between the findings of the in-person versus online studies.

There are several factors that can make the online versions less sensitive, such as video-deficit effect or decreased attention and motivation during online studies. For example, video-deficit effect has been shown to hinder children's performance on FB tasks before (e.g., Reiß et al., 2019). Or decreased attention and motivation during online testing have been shown to have hampering effects in memory tasks (e.g., Finley & Penningroth, 2015). Whether these factors could affect adult participants' performance in social cognition tasks is still unknown. So far, systematic comparisons were conducted with children, and they provided mixed findings. For example, Sheskin and Keil (2018) found that children performed poorer in online FB tasks compared to the live versions, whereas, more recently, Schidelko and colleagues (2021) found no difference between lab-based versus web-based versions of standard FB tasks. This kind of systematic comparisons are still lacking for adult studies in general and for the Sandbox task specifically; therefore, we cannot know the effect of the task format for sure in the current study. However, we speculate that the task format cannot fully explain the null results in the current project as the two baseline studies that we conducted with the live paper-pencil version of the Sandbox task also failed to reveal any egocentric bias. Therefore, the reliability of the Sandbox task as a measure of egocentric bias is still an open question.

Regarding the altercentric bias findings, the results provided by the Sandbox task is even more difficult to interpret because of two reasons: 1) the reliability of the original Sandbox task is still debated, jeopardizing also the reliability and validity of any adaptation of this task; and, 2) the altercentric bias adaptation of the Sandbox task is completely original and there is no existing body of positive or null findings that we can use to compare our results with. Still, several factors emerge as potential causes of the null results in this version, which can be grouped as superficial methodological factors and substantial theoretical factors. Methodological factors include, but are not limited to, the online format, the simplicity of the task, and technical limitations. For example, it is possible that participants did not pay enough attention to the task because of the unmoderated online format or, when they did, they found the task too easy, hence no difference was observed between trials.

Technical limitations in the current project mainly pertain to the mouse-tracking measures which were integrated into the Sandbox task. Mouse-tracking is supposed to be a more automatic and spontaneous measure than marking an answer on the screen as normally required by the Sandbox task. It is because this measure occurs spontaneously without any explicit trigger, and it is manifested in automatic motor responses. However, this measure's spontaneity and automaticity were compromised in the current study due to technical limitations of the study creation platform. Specifically, participants could not move their mouse cursors freely starting from the beginning of the trials; they always had to click on a "record button" first, which ensured that their cursors were positioned at a standardized start point in the beginning. Although this step had an important role in the comparability of mouse trajectories across trials and participants, it is possible that it also interfered with potential altercentric bias effects as it gave participants more time to reflect on their answers, and probably let them overcome any bias that could have been influential otherwise.

There also more substantial factors with theoretical bases that prevent the current task format from revealing any altercentric bias. For example, altercentric biases are hypothesized to be automatic, spontaneous, and implicit processing of others' perspectives (Southgate, 2020). It is then possible that these biases manifest themselves in more spontaneous temporal measures such as reaction times, but not in fine-grained spatial measures which are searched within the content of the judgments. In other words, reaction times are more suitable for quick, spontaneous, and automatic judgments whereas the location-marking judgments require more extended processing due to the answer format and the preceding scenarios. One other factor can be related to the extent of the agent's irrelevancy in the Sandbox studies. Altercentric biases have been shown to occur even when the agent in the scene was completely irrelevant for the participant's mission (e.g., Samson et al., 2010); however, these task formats utilized simpler setups in which participants were asked to report their visual perspectives. In the contrary, Sandbox task involves relatively elaborative stories where the location of an object is changed by two different agents and the judgment depends on memory. It is possible that, in a demanding task like this, participants do not register the agent's perspective at all as it is not relevant and as they have to reserve their working memory resources to be able to remember the final location of the task.

8.2.2 Director Task

In our studies, the Director Task provided robust evidence for egocentric bias effects but mixed results in terms of the altercentric biases. Egocentric interferences are not the main focus of this thesis, and the results of our Director Task were in line with earlier findings. Therefore, egocentric interferences found by the Director Task will not be discussed further within the scope of this thesis. It is just important to point out that our study was the first to reveal egocentric bias effects in a completely online version of the Director Task. This implies that this task can be a useful measure to reveal fine-grained differences in adults' perspective-taking abilities effortlessly in web-based studies. However, systematic comparisons with online versus in-person versions of this task is needed before the online version of this task can be employed as a reliable measure of explicit perspective-taking abilities.

The newly adapted altercentric bias version of the Director Task provided mixed findings across measures. More specifically, on the one hand, accuracy measure revealed that participants were more error-prone in their first-order searching judgements when the situation involved another agent who did not have a visual access to the searched object compared to the situation where they both had a visual access. This finding is in line with the expectations of altercentric interference effects: participants implicitly represent the other person's perspective, and their judgments become susceptible to errors as they go for objects that are visible for the agents even though it is irrelevant and even detrimental for their decision. On the other hand, reaction time measures provided support for the opposite of what altercentric biases would predict; namely, participants were faster when the situation involved an agent with a differing perspective compared to the situations where the participant and the agent had similar perspectives. This effect cannot be explained by altercentric interference effects since this type of bias would have expected participants to be slower in incongruent-perspectives trials rather than the congruent-perspectives trials. Therefore, this result call for alternative explanations.

One explanation that can account for the mismatch between accuracy and reaction time measures is the speed-accuracy trade-off. This phenomenon refers to the situation where either the speed and efficiency with which the task is completed, or the accuracy and precision of the task is compromised in the quest of completing a task. So, it is possible that participants compromised the accuracy in the current task for the sake of responding fast. However, this possibility is not a likely explanation in the current study for several reasons. First, in the present task, participants were not required to make very quick judgments; they had a four-seconds time window in which they could give their answers. They did not receive any discouragement (e.g., an alarming tone or a warning) if they had not responded in the allocated time; the study simply proceeded with a new trial. So, there was no reason for them to rush through the study such that the accuracy would have been affected by their haste. Also, the post-hoc analyses of the accuracy and response times revealed no relationship between these two measures on trial level, meaning participants were not particularly slower in the trials where they responded accurately, and they were not faster when they responded inaccurately. Based on these findings, the speed-accuracy trade-off can be given up as an alternative explanation for the mismatch that occurred between measures in the altercentric bias version of the Director Task.

The more likely alternative explanation that can enlighten the mismatch between the accuracy and reaction time measures pertains to low-level task-related features. More specifically, in the current altercentric bias task, the correct answer in the incongruent-perspective trials always corresponded to a grid which was covered by a black occluder from behind. It is possible that the black occluder decreased the visibility of the objects in the incongruent-perspective trials, diverting participants' attention to the objects that stood in front of lighter backgrounds and causing them to answer incorrectly but more quickly. This possibility would then suggest that the interferences occurring in the newly adapted version of the Director Task are not due to any kind of spontaneous mentalizing, but rather the product of low-level attentional features. The fact that similar results were revealed in the non-social version of the task also supports this possibility as it indicates that the agent's differing perspective did not play a role in the displayed effects. Therefore, it is unlikely that the effects occurred due to implicit mentalizing; instead, task-related factors might have been at play.

8.2.3 Level-I and Level-II Perspective-Taking Tasks (with Mouse-Tracking)

In the third project, the level-I and level-II perspective-taking tasks revealed somewhat robust altercentric interference effects, but not with the mouse-tracking measure. Contrary to our expectations, mouse cursor trajectories did not provide a novel, reliable measure to tap altercentric interferences. A similar result has also been found in our first project, where mouse-tracking measures were integrated into the Sandbox measures. The potential methodological limitations and theoretical bases for null results have been already discussed under the *Sandbox Task* section. We suspect that similar methodological concerns apply to the mouse-tracking measures employed in level-I and level-II perspective-taking tasks, casting doubt on the reliability of this measure in the current task format. Given this, mouse-tracking measure will not be included in further empirical and theoretical discussion of this project's results. The robustness of the altercentric interference effects revealed by accuracy and reaction times will be discussed in the following section, i.e., *Implications of the Current Positive Findings for Altercentric Biases*.

Project 3 revealed puzzling part-null results when it comes to the egocentric interference effects. Namely, egocentric biases have been found in level-II tasks, but not in level-I tasks. However, previous literature has shown these biases in both levels so far (e.g., Samson et al., 2010; Surtees et al., 2012). Why the current task format revealed egocentric biases exclusively in level-II task is still a mystery. Before further exploring these intriguing findings, it is important to clarify the nature of the interference effects revealed by the current task format and show that they are indeed due to one' own perspective or knowledge, but not due to confounding task related factors. An example for such a factor can be the demand for mental rotation in the level-II perspective-taking task. To elaborate further, in the current task format, participants were asked to identify numbers and letters from the agent's (opposite) perspective. In level-I task, and in the congruent versions of the level-II task, the numbers and letters were symmetrical, meaning they looked the same from participant's and agent's perspectives. In this case, participants did not have to rotate the numbers in their mind in order to calculate how they would be seen from the other's perspective. In the incongruent trials of the level-II, however, the presented numbers were not symmetrical, meaning they looked different from the opposite perspectives. In these trials, in order to find the number/letter that is seen by the agent, participants had to engage in mental rotation and imagine the stimuli from the opposite perspective. This process is not trivial, and it has been differentiated from perspective-taking abilities (e.g., Gunia et al., 2021; Hegarty & Waller, 2004). It is likely that the demand for mental rotation in the incongruent-perspective trials of the level-II task created effects that resembled egocentric biases but were actually due to general processing requirements. As the egocentric biases have been studied only for explorative reasons in the current project, this version was omitted after the first study. However, future research should take these findings into account and provide a solution for the potential mental rotation difficulties caused by the present task format.

8.2.4 Conclusion

Overall, our studies casted doubts on a) the Sandbox task as measure of both egocentric and altercentric biases; b) the Director Task as a measure of altercentric bias; and c) the modified level-I and level-II perspective-taking tasks as a measure of egocentric biases. When these findings were discussed in relation to the existing empirical and theoretical work, it was seen that there were several factors that could make these tasks unsuitable for tapping altercentric (and egocentric) biases, promoting the possibility of evidence of absence. However, several methodological limitations that could potentially account for the null results have also been identified, precluding eliminating the possibility of absence of evidence. Therefore, it is necessary for future research to test these factors more systematically before the proposed task formats are given up as potential measures to tap altercentric (and egocentric) biases.

8.3 Implications of the Current Positive Findings for Altercentric Bias

In three projects testing three different task formats as potential measures of altercentric interference effects, only one task remained promising. This task tapped the altercentric biases in the context of level-I and level-II perspective-taking via reaction time and accuracy measures (also through mouse-tracking but this measure will not be discussed any further due to the reasons presented above). In the following, the reliability and robustness of these measures, as well as what the findings indicate to in terms of altercentric bias and implicit ToM will be discussed.

8.3.1 Reconciling Accuracy or Reaction Time Measures in Project 3

In Project 3, in two studies, altercentric biases have been found in level-I and level-II perspective-taking tasks through both accuracy and reaction time measures. Among these measures, reaction times revealed consistent altercentric biases for both levels in all studies, whereas the accuracy indicated different bias patterns in two studies. Namely, accuracy revealed altercentric interferences only for level-I task in the first study and only for level-II task in the second study. This difference raised the question of which of these measures provided a more reliable means to reveal altercentric biases in the current task format. There are advantages and disadvantages to both measures. On the one hand, reaction times revealed more consistent results, however they were heavily influenced by the device that was used to do the marking (i.e., a mouse or touchpad). Reaction times also failed to reveal the effect of condition (i.e., social versus non-social) even though these conditions differed in the level of within-subject comparisons (i.e., altercentric bias was revealed in the social condition but not in the non-social condition). On the other hand, accuracy was sensitive to the condition and revealed an effect of condition that interacted with the trial types. Also, accuracy was not influenced by device selection. However, it failed to reveal biases when reaction times showed a bias. I suspected that this might be due to the lack of variation in this measure. In order to check this possibility, I re-conducted the within-subject analyses by combining all data collected for the second study of this project. Results revealed altercentric biases for both levels as indicated by the accuracy measure. This finding suggests that with increased power, accuracy measures also provide evidence in both levels. Therefore, the further theoretical discussion will be based on the premise that the altercentric biases are present in both levels.

8.3.2 Theoretical Implications

Reconciling the accuracy and reaction time measures provided us with a more solid empirical base on which the further theoretical discussion can build on. The results from these two measures across two different studies revealed fairly robust evidence for altercentric bias effects. A non-social control condition supported these results in the sense that the absence or presence of the agent's perspective played a role in the existence of these interference effects, hinting at the possibility that these effects indeed occur as a result of spontaneous and automatic calculations of others' perspectives. This kind of approach to the current results are in line with the implicit mentalizing interpretations of the altercentric biases (e.g., Samson et al., 2010). These approaches argue that the altercentric intrusions come about from computing others' perspectives quickly and implicitly in a given situation. These computations do not have to be registered as "perspectives" or be recognized consciously in order to be incorporated into one's judgments.

The mechanisms behind this incorporation and the potential reasons and advantages of this process remained to be studied. So far, the altercentrism hypotheses made some speculations about the mechanisms and reasons behind the altercentric intrusions (e.g., Kampis & Southgate, 2020). However, this account almost exclusively focused on early years of life and based their predictions on infants. However, in my work, as well many other earlier studies, these biases have been revealed in adult samples too (e.g., Kovács et al., 2010; Samson et al., 2010). Therefore, more comprehensive frameworks pertaining to the potential benefits and mechanisms of altercentric biases are needed.

This necessity becomes especially relevant for the discussion about the structural architecture of the mind. In other words, if altercentric biases actually stem from implicit and spontaneous mentalizing, what would be the implications of this finding for the idea of representational quarantining? The representational quarantining hypothesis argues that one's second-order representations (e.g., of others' mind, or descriptions of particular aspect of reality) are kept separate from one's own first-order or primary representations. However, altercentric intrusion effects, assuming they indeed result from implicit mentalizing, suggest that one's own first-order representations are not completely partitioned from one's second-order representations of others' perspectives. Instead, altercentric bias indicates that the first-order representations are subject to interferences from the representations of other's perspectives, even when there is no apparent need to represent what others are perceiving or representing. But how does this process happen and what are the potential advantages of this representational intrusion? Any explanatory framework on altercentric interference effects should clarify this question in future research.

And finally, if the current task format is indeed a reliable measure of altercentric biases, what would be the implications of its findings -which revealed altercentric biases in both level-I and level-II tasks- in terms of the theoretical approaches to implicit ToM, i.e., dual-system accounts, nativist approaches, and submentalizing hypothesis? The findings of the present studies did not reveal that the implicit and spontaneous perspective-taking was limited to level-I judgments; instead, it appeared in both level-I and level-II tasks in the form of altercentric biases. Therefore, these results cannot be explained by dual-system accounts (e.g., Apperly & Butterfill, 2009), which proposes signature limits on efficient and implicit theory of mind processes, such as level-II perspective-taking. This claim has been tested in previous work and some earlier studies have already demonstrated these signature limits for implicit perspective-taking abilities (e.g., Surtees et al., 2012; Surtees et al., 2016). However, these studies used slightly different task designs in order to investigate level-I versus level-II judgments. Therefore, it is possible that potential confounds led to the differences between the two levels in those studies. For example, it is possible that, in the study by Surtees and colleagues (2016), level-I judgments were affected by the directional cues generated by participant's gaze and position, whereas no such cues were present for the level-II tasks. In the current study, however, no such differences were observed between the stimuli of the two levels. Future work should take these potential confounds and differences into account and systematically retest the biases in two levels.

If we assume, for now, that our task revealed reliable and robust altercentric biases for both levels of judgment, where do these findings lead us theoretically? Earlier in this thesis, nativism and submentalizing accounts have been introduced as the other alternative theoretical approaches to the implicit ToM. As the findings of the current study did not reveal a difference between level-I versus level-II judgements in terms of the implicit perspective-taking, do they necessarily indicate unified ToM abilities as suggested by nativism accounts? Perhaps. However, to be able to make supporting claims for nativist account, first the alternative submentalizing explanations of the present findings should be dismissed.

But are the findings of the current studies sufficient to disqualify the possible submentalizing explanations or is it still possible that they reflect interferences occurring due to low-level, domain-general processes triggered by the task design? The submentalizing account (e.g., Heyes, 2014a; Santiesteban et al., 2014) suggests that, for example, the "altercentric" interferences found in the dot-perspective task and its variations occur due to directional features of the agent's position and gaze, which cue participants in the direction of what is seen by the agent. These cues then cause interferences that are similar to the altercentric biases but not altercentric in the sense that they occur through implicit mentalizing of others' perspectives. A similar concern applies to the altercentric biases that were revealed in the current study. That is why a non-social version of the task was tested in the second study of Project 3. Even though this condition revealed no bias, indicating that the effects were dependent on the presence of the agent, the results can still be explained by submentalizing arguments. For example, it is possible that agent's orientation and fixed gaze on the table made participants to fixate on the numbers/letters that were visible from this person's perspective, or on how they were seen from this person's perspective. This directional cue would have then diverted participants' attention, causing them to become more error prone and slower in their judgments due to a general cognitive processing burden, but not because implicit mentalizing of the other's perspective. As this effect is also dependent on the agent's gaze, it is expected that it will disappear once the agent is removed from the scene. Therefore, this alternative can also explain the null results found in the non-social condition of this task. Given that, in order to be able to claim that the current task format is reliably measuring the altercentric biases, follow-up studies with strict control conditions are needed.

8.4 Limitations

The projects that were conducted within the scope of this dissertation were not free from methodological and conceptual limitations. Some of these methodological limitations have already been mentioned. For example, the diminished spontaneity of the mouse-tracking measures and the variation in the sensitivity and responsivity of the response devices (i.e., mouse and touchpads) are important limitations that potentially contributed to the null findings. One other limitation is a more general issue that apply all projects of this thesis: the lack of standardization and moderation in the testing sessions. All studies that are presented in this thesis have been conducted in completely online formats as unmoderated studies (except the two baseline studies in Project 1), meaning we had no control over the surroundings of the participants and potential distractors. We tried to overcome this limitation by hiding attention checks within the studies, however, this precaution does not guarantee quality data. Also, the lack of experimenter and any type of supervision and feedback might have influenced participant's performance in the online studies through decreased attention and motivation.

Conceptual limitations pertain to the salience and reality of the agent in the altercentric bias measures. In these tasks, we deliberately refrained from making any reference to the irrelevant agents in order not to prompt any perspective-taking attempt from the participant's side. However, it is possible that this effort went beyond its purpose and caused participants to ignore the agent entirely, making its perspective go completely unrecognized and unprocessed. Apart from the salience, the (un)reality of the agent might be a potential limitation of the current studies. In our experiments, the irrelevant agents were computer-generated virtual characters. Even though we tried to make them as realistic as possible by using animated materials that were supported through real-life scenarios, it is still possible that participants did not think of these characters as real agents with perspectives. Animated characters and figurines have been extensively used in social-cognition literature, and their properties did not have an effect on, for example, the performance in FB tasks (e.g., Liu et al., 2008; Wellman et al., 2001). However, in FB tasks, participants' attention is explicitly attracted to these agents' beliefs or behaviors, which would eventually help participants attribute perspectives to these agents. However, this was not the case in the current studies.

Therefore, it is still possible that the participants did not even attribute any perspectives to the agent to begin with.

8.5 Future Directions

Besides eliminating the aforementioned limitations and confounds from the current task designs, the scope of this thesis can be extended via two lines of future studies. One of these directions would be to keep testing altercentric biases as means to tap implicit ToM in different task formats. In the current study, we aimed to make use of the relatively extensive literature on egocentric bias and explicit perspective-taking tasks, therefore we relied on three of these measures in order to have altercentric bias adaptations. However, the literature provides us with many other feasible options that have been already used to target altercentric biases (e.g., manual search tasks, Kampis & Kovács, 2022) or that can be adapted to tap altercentric intrusions with small modifications (e.g., number-approximation tasks, Odic & Starr, 2018).

The second line of potential future directions builds on the first line of future studies and depends on the premise that a task format is proved to be a reliable and valid way of tapping altercentric biases, and, thus, implicit ToM abilities. Once a suitable task format is identified, developmental studies would be needed in order to investigate the ontogeny and developmental trajectory of the altercentric interference effects and implicit ToM. Numerous studies have been conducted in explicit ToM literature so far, and our knowledge about the ontogeny, development, and lifelong trajectory of this ability is well-based by now (even though the interpretations of these findings might vary). However, implicit ToM literature lacks this kind empirical basis, mostly due to the replication crisis in this literature. Therefore, adapting a developmental perspective together with a novel task can enrich the implicit ToM literature to a great extent.

8.6 Conclusion

Altercentric biases provide a potential new window into the implicit ToM abilities, which have lacked robust empirical support in recent years. As the typical altercentric bias measures have also been in the center of methodological and theoretical criticism, new measures that could reliably tap these intrusions were much needed. The current thesis explored the altercentric biases in three different novel task formats and evaluated their results in relation to the existing empirical and theoretical work. Among three measures, only one measure, where altercentric biases were tapped in level-I and level-II perspective-taking tasks, persisted as a potential new measure of altercentric intrusions. Although this measure is not completely foolproof yet, it revealed consistent altercentric intrusions in two different studies and provided support for the implicit mentalizing interpretations of altercentric interference effects. Overall, these results provided a promising first step to a reliable altercentric bias measure and laid the foundations for further studies that would extend the use of altercentric bias tasks as measures of implicit ToM.

9 References

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review, 116*(4), 953-970. https://doi.org/10.1037/a0016923

Apperly, I. A., Back, E., Samson, D., & Grant, L. (2007). The cost of thinking about false beliefs: Evidence from adult performance on a non-inferential theory of mind task. *Cognition, 106*(3), 1093–1108. https://doi.org/10.1016/j.cognition.2007.05.005

Apperly, I. A., Carroll, D. J., Samson, D., Humphreys, G. W., Qureshi, A., & Moffitt, G. (2010). Why are there limits on theory of mind use? Evidence from adults' ability to follow instructions from an ignorant speaker. *Quarterly Journal of Experimental Psychology*, *63*(6), 1201-1217. https://doi.org/10.1080/17470210903281582

Apperly, I. A., Warren, F., Andrews, B. J., Grant, J., & Todd, S. (2011). Developmental continuity in theory of mind: Speed and accuracy of belief-desire reasoning in children and adults. *Child Development*, *82*(5), 1691–1703. https://doi.org/10.1111/j.1467-8624.2011.01635.x

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

Back, E., & Apperly, I. A. (2010). Two sources of evidence on the non-automaticity of true and false belief ascription. *Cognition, 115*(1), 54-70. https://doi.org/10.1016/j.cognition.2009.11.008

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

Baillargeon, R., Scott, R. M., & 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

References

Bardi, L., Desmet, C., Nijhof, A., Wiersema, J. R., & Brass, M. (2017). Brain activation for spontaneous and explicit false belief tasks overlaps: new fMRI evidence on belief processing and violation of expectation. *Social Cognitive and Affective Neuroscience, 12*(3), 391–400. https://doi.org/10.1093/scan/nsw143.

Begeer, S., Bernstein, D. M., van Wijhe, J., Scheeren, A. M., & Koot, H. M. (2012). A continuous false belief task reveals egocentric biases in children and adolescents with autism spectrum disorders. *Autism*, *16*(4), 357-366. https://doi.org/10.1177/1362361311434545

Bennett, J. (1978). Some remarks about concepts. *Behavioral and Brain Sciences*, *1*(4), 557–560. https://doi.org/10.1017/S0140525X00076573

Bernstein, D. M., Erdfelder, E., Meltzoff, A. N., Peria, W., & Loftus, G. R. (2011a). Hindsight bias from 3 to 95 years of age. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(2), 378–391. https://doi.org/10.1037/a0021971

Bernstein, D. M., Thornton, W. L., & Sommerville, J. A. (2011b). Theory of mind through the ages: Older and middle-aged adults exhibit more errors than do younger adults on a continuous false belief task. *Experimental Aging Research*, *37*(5), 481-502. https://doi.org/10.1080/0361073X.2011.619466

Birch, S. A., & Bloom, P. (2004). Understanding children's and adults' limitations in mental state reasoning. *Trends in Cognitive Sciences, 8*(6), 255-260. https://doi.org/10.1016/j.tics.2004.04.011

Birch, S. A., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science*, *18*(5), 382–386. https://doi.org/10.1111/j.1467-9280.2007.01909.x

Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica*, *106*(1-2), 3-22. https://doi.org/10.1016/S0001-6918(00)00024-X

References

Brooks, R., & Meltzoff, A. N. (2002). The importance of eyes: How infants interpret adult looking behavior. *Developmental Psychology, 38*(6), 958–966. https://doi.org/10.1037/0012-1649.38.6.958

Bushneil, I. W. R., Sai, F., & Mullin, J. T. (1989). Neonatal recognition of the mother's face. *British Journal of Developmental Psychology*, *7*(1), 3-15. https://doi.org/10.1111/j.2044-835X.1989.tb00784.x

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., & 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

Cane, J. E., Ferguson, H. J., & Apperly, I. A. (2017). Using perspective to resolve reference: The impact of cognitive load and motivation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(4), 591–610. https://doi.org/10.1037/xlm0000345

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

Chambers, J. R., Epley, N., Savitsky, K., & Windschitl, P. D. (2008). Knowing too much: Using private knowledge to predict how one is viewed by others. *Psychological Science*, *19*(6), 542-548. https://doi.org/10.1111/j.1467-9280.2008.02121.x

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

Coburn, P. I., Bernstein, D. M., & Begeer, S. (2015). A new paper and pencil task reveals adult false belief reasoning bias. *Psychological Research*, *79*, 739-749. https://doi.org/10.1007/s00426-014-0606-0

References

Cole, G. G., Atkinson, M., Le, A. T., & Smith, D. T. (2016). Do humans spontaneously take the perspective of others?. *Acta Psychologica*, *164*, 165-168. https://doi.org/10.1016/j.actpsy.2016.01.007

Conway, J. R., Lee, D., Ojaghi, M., Catmur, C., & Bird, G. (2017). Submentalizing or mentalizing in a Level 1 perspective-taking task: A cloak and goggles test. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(3), 454–465. https://doi.org/10.1037/xhp0000319

Decety, J., & Sommerville, J. A. (2003). Shared representations between self and other: a social cognitive neuroscience view. *Trends in Cognitive Sciences*, *7*(12), 527-533. https://doi.org/10.1016/j.tics.2003.10.004

Dennett, D. (1978). Beliefs about beliefs [P&W, SR&B]. *Behavioral and Brain Sciences*, *1*(4), 568-570. https://doi.org/10.1017/S0140525X00076664

Deschrijver, E., Bardi, L., Wiersema, J. R., & Brass, M. (2016). Behavioral measures of implicit theory of mind in adults with high functioning autism. *Cognitive Neuroscience*, *7*(1–4). 192-202. https://doi.org/10.1080/17588928.2015.1085375

Devine, R. T., & Hughes, C. (2014). Relations between false belief understanding and executive function in early childhood: A meta-analysis. *Child Development*, *85*(5), 1777-1794. https://doi.org/10.1111/cdev.12237

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

Dumontheil, I., Apperly, I. A., & Blakemore, S. J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science*, *13*(2), 331-338. https://doi.org/10.1111/j.1467-7687.2009.00888.x

Dumontheil, I., Hillebrandt, H., Apperly, I. A., & Blakemore, S. J. (2012). Developmental differences in the control of action selection by social information. *Journal of Cognitive Neuroscience*, *24*(10), 2080-2095. https://doi.org/10.1162/jocn_a_00268

References

El Kaddouri, R., Bardi, L., De Bremaeker, D., Brass, M., & Wiersema, J. R. (2020). Measuring spontaneous mentalizing with a ball detection task: putting the attention-check hypothesis by Phillips and colleagues (2015) to the test. *Psychological Research, 84*(6), 1749-1757. https://doi.org/10.1007/s00426-019-01181-7

Elekes, F., Varga, M., & Király, I. (2016). Evidence for spontaneous level-2 perspective taking in adults. *Consciousness and Cognition, 41*, 93-103. https://doi.org/10.1016/j.concog.2016.02.010

Finley, A. J., & Penningroth, S. L. (2015). Online versus in-lab: Pros and cons of an online prospective memory experiment. In A. M. Columbus (Ed.), *Advances in psychology research* (pp. 35–161). New York, NY: Nova.

Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the Level 1–Level 2 distinction. *Developmental Psychology, 17*(1), 99–103. https://doi.org/10.1037/0012-1649.17.1.99

Flavell, J. H., 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, 51*(1), i–87. https://doi.org/10.2307/1165866

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

Forgács, B., Parise, E., Csibra, G., Gergely, G., Jacquey, L., & Gervain, J. (2019). Fourteen-month-old infants track the language comprehension of communicative partners. *Developmental Science, 22*(2), e12751. https://doi.org/10.1111/desc.12751

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

Furlanetto, T., Becchio, C., Samson, D., & Apperly, I. (2016). Altercentric interference in level 1 visual perspective taking reflects the ascription of mental states, not submentalizing. *Journal of Experimental Psychology: Human Perception and Performance, 42*(2), 158–163. https://doi.org/10.1037/xhp0000138

References

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

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

Gunia, A., Moraresku, S., & Vlček, K. (2021). Brain mechanisms of visuospatial perspective-taking in relation to object mental rotation and the theory of mind. *Behavioural Brain Research*, *407*, 113247. https://doi.org/10.1016/j.bbr.2021.113247

Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, *32*(2), 175-191. https://doi.org/10.1016/j.intell.2003.12.001

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, *57*(2), 243–259. https://doi.org/10.2307/1416950

Heyes, C. (2014a). Submentalizing: I Am Not Really Reading Your Mind. *Perspectives on Psychological Science*, *9*(2), 131–143. https://doi.org/10.1177/1745691613518076

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

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

Holland, C., Shin, S. M., & Phillips, J. (2021). Do you see what I see? A meta-analysis of the Dot Perspective Task. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *43*. https://escholarship.org/uc/item/7cs5r2xq

Johnson, S. C., Slaughter, V., & Carey, S. (1998). Whose gaze will infants follow? The elicitation of gaze-following in 12-month-olds. *Developmental Science*, *1*(2), 233–238. https://doi.org/10.1111/1467-7687.00036

References

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

Kampis, D., & Kovács, Á. M. (2022). Seeing the world from others' perspective: 14-month-olds show altercentric modulation effects by others' beliefs. *Open Mind*, *5*, 189-207. https://doi.org/10.1162/opmi_a_00050

Kampis, D., & Southgate, V. (2020). Altercentric cognition: How others influence our cognitive processing. *Trends in Cognitive Sciences*, *24*(11), 945-959. https://doi.org/10.1016/j.tics.2020.09.003

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, *11*(1), 32-38. https://doi.org/10.1111/1467-9280.00211

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

Király, I., Oláh, K., Csibra, G., & Kovács, Á. M. (2018). Retrospective attribution of false beliefs in 3-year-old children. *Proceedings of the National Academy of Sciences*, *115*(45), 11477-11482. https://doi.org/10.1073/pnas.1803505115

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

Kovács, Á. M., Kühn, S., Gergely, G., Csibra, G., & Brass, M. (2014). Are all beliefs equal? Implicit belief attributions recruiting core brain regions of theory of mind. *PloS One*, *9*(9), e106558. https://doi.org/10.1371/journal.pone.0106558.

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

References

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

Kulke, L., Reiß, M., Krist, H., & Rakoczy, H. (2018a). How robust are anticipatory looking measures of Theory of Mind? Replication attempts across the life span. *Cognitive Development*, *46*, 97–111. https://doi.org/10.1016/j.cogdev.2017.09.001

Kulke, L., von Duhn, B., Schneider, D., & Rakoczy, H. (2018b). Is implicit theory of mind a real and robust phenomenon? Results from a systematic replication study. *Psychological Science*, *29*(6), 888-900. https://doi.org/10.1177/0956797617747090

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

Langton, S. R. (2018). I don't see it your way: The dot perspective task does not gauge spontaneous perspective taking. *Vision*, *2*(1), 6. https://doi.org/10.3390/vision2010006

Legg, E. W., Olivier, L., Samuel, S., Lurz, R., & Clayton, N. S. (2017). Error rate on the director's task is influenced by the need to take another's perspective but not the type of perspective. *Royal Society Open Science*, *4*(8), 170284. https://doi.org/10.1098/rsos.170284

Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind." *Psychological Review*, *94*(4), 412–426. https://doi.org/10.1037/0033-295X.94.4.412

Leslie, A. M. (1994). ToMM, ToBy, and Agency: Core architecture and domain specificity. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 119–148). Cambridge University Press. https://doi.org/10.1017/CBO9780511752902.006

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

References

Liu, D., Wellman, H. M., Tardif, T., & Sabbagh, M. A. (2008). Theory of mind development in Chinese children: A meta-analysis of false-belief understanding across cultures and languages. *Developmental Psychology*, *44*(2), 523–531. https://doi.org/10.1037/0012-1649.44.2.523

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

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

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. https://doi.org/10.1016/j.jecp.2017.03.010

Marshall, J., Gollwitzer, A., & Santos, L. R. (2018). Does altercentric interference rely on mentalizing?: Results from two level-1 perspective-taking tasks. *PloS One, 13*(3), e0194101. https://doi.org/10.1371/journal.pone.0194101

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

Meltzoff, A. N., & Moore, M. K. (1983). Newborn infants imitate adult facial gestures. Child Development, *54*(3), 702-709. https://doi.org/10.2307/1130058

Michael, J., Wolf, T., Letesson, C., Butterfill, S., Skewes, J., & Hohwy, J. (2018). Seeing it both ways: Using a double-cuing task to investigate the role of spatial cuing in Level-1 visual perspective-taking. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(5), 693–702. https://doi.org/10.1037/xhp0000486

Miller, S. A. (2012). *Theory of mind: Beyond the preschool years.* Psychology Press.

References

Moll, H., & Meltzoff, A. N. (2011). How does it look? Level 2 perspective-taking at 36 months of age. *Child Development, 82*(2), 661-673. https://doi.org/10.1111/j.1467-8624.2010.01571.x

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

Nielsen, M. K., Slade, L., Levy, J. P., & Holmes, A. (2015). Inclined to see it your way: Do altercentric intrusion effects in visual perspective taking reflect an intrinsically social process?. *Quarterly Journal of Experimental Psychology, 68*(10), 1931-1951. http://dx.doi.org/10.1080/17470218.2015.1023206

Nijhof, A. D., Bardi, L., Brass, M., & Wiersema, J. R. (2018). Brain activity for spontaneous and explicit mentalizing in adults with autism spectrum disorder: An fMRI study. NeuroImage: *Clinical, 18*, 475–484. https://doi.org/10.1016/j.nicl.2018.02.016.

Nijhof, A. D., Brass, M., & Wiersema, J. R. (2017). Spontaneous mentalizing in neurotypicals scoring high versus low on symptomatology of autism spectrum disorder. *Psychiatry Research, 258*, 15–20. https://doi.org/10.1016/j.psychres.2017.09.060.

Odic, D., & Starr, A. (2018). An introduction to the approximate number system. *Child Development Perspectives, 12*(4), 223-229. https://doi.org/10.1111/cdep.12288

O'Grady, C. J., Scott-Phillips, T., Lavelle, S., & Smith, K. (2017). The dot perspective task revisited: Evidence for directional effects. In Cogsci 2017: *Proceedings of the 39th annual meeting of the cognitive science society* (pp. 2821–2826). https://pdfs.semanticscholar.org/90a2/0cdfbc0bd26c0b977 d6294d7f5ade726f7db.pdf

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

Oostenbroek, J., Redshaw, J., Davis, J., Kennedy-Costantini, S., Nielsen, M., Slaughter, V., & Suddendorf, T. (2018). Re-evaluating the neonatal imitation hypothesis. *Developmental Science, 22*, e12720. https://doi.org/10 .1111/desc.12720

References

Oostenbroek, J., Suddendorf, T., Nielsen, M., Redshaw, J., Kennedy-Costantini, S., Davis, J., Clark, S., & Slaughter, V. (2016). Comprehensive longitudinal study challenges the existence of neonatal imitation in humans. *Current Biology, 26*(10), 1334–1338. https://doi.org/10.1016/j.cub.2016.03.047

Perner, J. (1992). Grasping the concept of representation: Its impact on 4-year-olds' theory of mind and beyond. *Human Development, 35*(3), 146-155. https://doi.org/10.1159/000277146

Perner, J. (2010). Who took the cog out of cognitive science? Mentalism in an era of anti-cognitivism. In P. A. Frensch & R. Schwarzer (Eds.), *Cognition and neuropsychology: International perspectives on psychological science* (Vol. 1, pp. 241–261). Psychology Press.

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

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

Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology, 5*(2), 125–137. https://doi.org/10.1111/j.2044- 835x.1987.tb01048.x

Phillips, J., Ong, D. C., Surtees, A. D. R., Xin, Y., Williams, S., Saxe, R., & Frank, M. C. (2015). A Second Look at Automatic Theory of Mind: Reconsidering Kovács, Téglás, and Endress (2010). *Psychological Science, 26*(9), 1353–1367. https://doi.org/10.1177/0956797614558717

Poulin-Dubois, D., Rakoczy, H., Burnside, K., Crivello, C., Dörrenberg, S., Edwards, K., ... & Ruffman, T. (2018). Do infants understand false beliefs? We don't know yet–A commentary on Baillargeon, Buttelmann and Southgate's commentary. *Cognitive Development, 48*, 302-315. https://doi.org/10.1016/j.cogdev.2018.09.005

References

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

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

Qureshi, A. W., Apperly, I. A., & Samson, D. (2010). Executive function is necessary for perspective selection, not Level-1 visual perspective calculation: Evidence from a dual-task study of adults. *Cognition, 117*(2), 230-236. https://doi.org/10.1016/j.cognition.2010.08.003

Rakoczy, H. (2017). Theory of Mind. In B. Hopkins, E. Geangu, & S. Linkenauger (Eds.), *The Cambridge Encyclopedia of child development* (pp. 505–512). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/9781316216491.081

Rakoczy, H. (2022). Foundations of theory of mind and its development in early childhood. *Nature Reviews Psychology*, *1*(4), 223-235. https://doi.org/10.1038/s44159-022-00037-z

Reiß, M., Krüger, M., & Krist, H. (2019). Theory of mind and the video deficit effect: Video presentation impairs children's encoding and understanding of false belief. *Media Psychology*, *22*(1), 23-38. https://doi.org/10.1080/15213269.2017.1412321

Rochat, P., Querido, J. G., & Striano, T. (1999). Emerging sensitivity to the timing and structure of protoconversation in early infancy. *Developmental Psychology*, *35*(4), 950–957. https://doi.org/10.1037/0012-1649.35.4.950

Rowling, J. K. (2003). *Harry Potter and the Order of Phoenix*. Bloomsbury.

Rueschemeyer, S. A., Gardner, T., & Stoner, C. (2015). The Social N400 effect: how the presence of other listeners affects language comprehension. *Psychonomic Bulletin & Review*, *22*, 128-134. https://doi.org/10.3758/s13423-014-0654-x

References

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

Samuel, S., Legg, E. W., Lurz, R., & Clayton, N. S. (2018a). Egocentric bias across mental and non-mental representations in the Sandbox Task. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 71*(11), 2395–2410. https://doi.org/10.1177/1747021817742367

Samuel, S., Legg, E. W., Lurz, R., & Clayton, N. S. (2018b). The unreliability of egocentric bias across self-other and memory-belief distinctions in the Sandbox Task. *Royal Society Open Science, 5*(11), 181355. https://doi.org/10.1098/rsos.181355

Samuel, S., Roehr-Brackin, K., Jelbert, S., & Clayton, N. S. (2019). Flexible egocentricity: Asymmetric switch costs on a perspective-taking task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(2), 213-218. https://doi.org/10.1037/xlm0000582

Santiesteban, I., Catmur, C., Hopkins, S. C., Bird, G., & Heyes, C. (2014). Avatars and arrows: Implicit mentalizing or domain-general processing? *Journal of Experimental Psychology: Human Perception and Performance, 40*, 929–937. https://doi.org/10.1037/a0035175

Schidelko, L. P., Schünemann, B., Rakoczy, H., & Proft, M. (2021). Online testing yields the same results as lab testing: a validation study with the false belief task. *Frontiers in Psychology, 12*, 4573. https://doi.org/10.3389/fpsyg.2021.703238

Schneider, D., Slaughter, V. P., & Dux, P. E. (2017). Current evidence for automatic Theory of Mind processing in adults. *Cognition, 162*, 27-31. https://doi.org/10.1016/j.cognition.2017.01.018

Schuwerk, T., Kampis, D., Baillargeon, R., Biro, S., Bohn, M., Byers-Heinlein, K., et al. (2021). Action anticipation based on an agent's epistemic state in toddlers and adults. https://doi.org/10.31234/osf.io/x4jbm.

References

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

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.

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Your Task is my Task: Shared Task Representations in Dyadic Interactions. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *25*. https://escholarship.org/uc/item/3tz2979t

Sebanz, N., Knoblich, G., & Prinz, W. (2005). How Two Share a Task: Corepresenting Stimulus-Response Mappings. *Journal of Experimental Psychology: Human Perception and Performance, 31*(6), 1234–1246. https://doi.org/10.1037/0096-1523.31.6.1234

Sheskin, M., & Keil, F. (2018, December 30). TheChildLab.com A Video Chat Platform for Developmental Research. https://doi.org/10.31234/osf.io/rn7w5

Shteynberg, G. (2010). A silent emergence of culture: The social tuning effect. *Journal of Personality and Social Psychology, 99*(4), 683–689. https://doi.org/10.1037/a0019573

Simion, F., Regolin, L., & Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proceedings of the National Academy of Sciences, 105*(2), 809-813. https://doi.org/10.1073/pnas.0707021105

Sodian, B., Thoermer, C., & Metz, U. (2007). Now I see it but you don't: 14-month-olds can represent another person's visual perspective. *Developmental Science, 10*(2), 199-204. https://doi.org/10.1111/j.1467-7687.2007.00580.x

Sommerville, J. A., Bernstein, D. M., & Meltzoff, A. N. (2013). Measuring beliefs in centimeters: Private knowledge biases preschoolers' and adults' representation of others' beliefs. *Child Development*, *84*(6), 1846-1854. https://doi.org/10.1111/cdev.12110

Southgate, V. (2020). Are infants altercentric? The other and the self in early social cognition. *Psychological Review*, *127*(4), 505–523. https://doi.org/10.1037/rev0000182

Southgate, V., & Vernetti, A. (2014). Belief-based action prediction in preverbal infants. *Cognition*, *130*(1), 1-10. https://doi.org/10.1016/j.cognition.2013.08.008

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

Speiger, M. L., Rothmaler, K., Rakoczy, H., Kizilay, F. N. H., Liszkowski, U., & Grosse Wiesmann, C. (2022, January 10-14) *The Sandbox Task-Measuring implicit and explicit Theory of Mind* [Poster presentation]. Budapest CEU Conference on Cognitive Development, Budapest, Hungary.

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

Surian, L., 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

Surtees, A. D., & Apperly, I. A. (2012). Egocentrism and automatic perspective taking in children and adults. *Child Development*, *83*(2), 452-460. https://doi.org/10.1111/j.1467-8624.2011.01730.x

References

Surtees, A. D., 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

Surtees, A., Samson, D., & Apperly, I. (2016). Unintentional perspective-taking calculates whether something is seen, but not how it is seen. *Cognition, 148*, 97–105. https://doi.org/10.1016/j.cognition.2015.12.010

Symeonidou, I., Dumontheil, I., Chow, W. Y., & Breheny, R. (2016). Development of online use of theory of mind during adolescence: An eye-tracking study. *Journal of Experimental Child Psychology*, *149*, 81-97. https://doi.org/10.1016/j.jecp.2015.11.007

Tomasello, M. (1999). *The cultural origins of human cognition*. Harvard University Press.

Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences*, *16*(3), 495-511. https://doi.org/10.1017/S0140525X0003123X

Van der Wel, R. P., Sebanz, N., & Knoblich, G. (2014). Do people automatically track others' beliefs? Evidence from a continuous measure. *Cognition*, *130*(1), 128-133. https://doi.org/10.1016/j.cognition.2013.10.004

Vestner, T., Balsys, E., Over, H., & Cook, R. (2022). The self-consistency effect seen on the Dot Perspective Task is a product of domain-general attention cueing, not automatic perspective taking. *Cognition*, *224*, 105056. https://doi.org/10.1016/j.cognition.2022.105056

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

Warneken, F., & Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy*, *11*(3), 271-294. https://doi.org/10.1111/j.1532-7078.2007.tb00227.x

References

Wellman, H. M. (2011). Developing a theory of mind. In U. Goswami (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development* (2nd ed., pp. 258–284). Blackwell.

Wellman, H. M. (2014). *Making minds: How theory of mind develops*. Oxford University Press.

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

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

Westley, A., Kohút, Z., & Rueschemeyer, S. A. (2017). "I know something you don't know": Discourse and social context effects on the N400 in adolescents. *Journal of Experimental Child Psychology, 164*, 45-54. https://doi.org/10.1016/j.jecp.2017.06.016

Westra, E., Terrizzi, B. F., van Baal, S. T., Beier, J. S., & Michael, J. (2021). Beyond avatars and arrows: Testing the mentalising and submentalising hypotheses with a novel entity paradigm. *Quarterly Journal of Experimental Psychology, 74*(10), 1709-1723. https://doi.org/10.1177/17470218211007388

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

Woodward, A. L. (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

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

References

Yott, J., & Poulin-Dubois, D. (2016). Are infants' theory-of-mind abilities well integrated? Implicit understanding of intentions, desires, and beliefs. *Journal of Cognition and Development, 17*(5), 683-698. https://doi.org/10.1080/15248372.2015.1086771

Zaki, J., Schirmer, J., & Mitchell, J. P. (2011). Social influence modulates the neural computation of value. *Psychological Science, 22*(7), 894-900. https://doi.org/10.1177/0956797611411057

Appendix A: Haskaraca, Proft, Liszkowski, & Rakoczy (2023)

Haskaraca, F.N., Proft, M., Liszkowski, U., & Rakoczy, H. (2023). How robust are egocentric and altercentric interference effects in social cognition? A test with explicit and implicit versions of a continuous False Belief task. *Frontiers in Psychology, 14,* 1142302.

This article has been published in Frontiers in Psychology, published by Frontiers Media S.A.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 10 July 2023
DOI 10.3389/fpsyg.2023.1142302

Check for updates

OPEN ACCESS

EDITED BY
Anja Gampe,
University of Duisburg-Essen, Germany

REVIEWED BY
Igor Bascandziev,
Harvard University, United States
Burcu Arslan,
Educational Testing Service, United States
Lindsay Bowman,
University of California, Davis, United States

*CORRESPONDENCE
Feride Nur Haskaraca
✉ feridenur.haskaracakizilay@uni-goettingen.de

RECEIVED 11 January 2023
ACCEPTED 23 June 2023
PUBLISHED 10 July 2023

CITATION
Haskaraca FN, Proft M, Liszkowski U and Rakoczy H (2023) How robust are egocentric and altercentric interference effects in social cognition? a test with explicit and implicit versions of a continuous false belief task. *Front. Psychol.* 14:1142302.
doi: 10.3389/fpsyg.2023.1142302

How robust are egocentric and altercentric interference effects in social cognition? a test with explicit and implicit versions of a continuous false belief task

Feride Nur Haskaraca[1*], Marina Proft[1], Ulf Liszkowski[2] and Hannes Rakoczy[1]

[1]Department of Developmental Psychology, University of Göttingen, Göttingen, Germany, [2]Department of Developmental Psychology, University of Hamburg, Hamburg, Germany

It has been long assumed that meta-representational theory of mind (ToM) -our ability to ascribe mental states to ourselves and other people- emerges around age four as indicated in performance on explicit verbal false belief tasks. In contrast, newer studies assessing false belief understanding with implicit, non-verbal measures suggest that some form of ToM may be present even in infancy. But these studies now face replication issues, and it remains unclear whether they can provide robust evidence for implicit ToM. One line of research on implicit ToM, however, may remain promising: Studies that tap so-called altercentric biases. Such biases occur when agents in their judgments about the world are influenced (perform slower, more error-prone) in light of another agent's deviating perspective even if that perspective is completely irrelevant to the task; they thus can be seen as indicators of spontaneous and implicit ToM. Altercentric biases are the mirror images of egocentric biases (agents are influenced by their own perspective when evaluating another agent's deviating perspective). In three studies with adults, we aimed to tap both egocentric and altercentric interference effects within the same task format. We used the so-called Sandbox task, a false belief task with continuous locations. In Study 1, we tested an online adaptation of the Sandbox task, which we also used to explore potential cross-cultural differences in these biases. Studies 2 and 3 combined the Sandbox task with mouse-tracking measures. These studies revealed neither egocentric nor altercentric biases. These null results are discussed with regard to the question whether absence of evidence here may present evidence of absence of such spontaneous perspective-taking biases or merely false negatives.

KEYWORDS

social cognition, egocentric bias, altercentric bias, Sandbox task, mouse tracking

1. Introduction

Theory of Mind (ToM), the ability to ascribe mental states to ourselves and other people, is central to our nature as social beings. At the heart of ToM is the capacity to represent people's representational states, i.e., meta-representation. The clearest indicator of this capacity is representing how others represent the world and act accordingly, even when their representation is inaccurate (hence a misrepresentation). The litmus tests for the development of meta-representational ToM have been False Belief (FB) tasks. In the most widely used so-called change-of-location FB task, participants hear a story in which an object is re-located either in the absence (FB condition) or the presence (True Belief control condition) of a protagonist. Then the participants are asked where the protagonist will look for the object (e.g., Wimmer and Perner, 1983). Decades of research using the change-of-location FB task, or other FB tasks such as unexpected contents (e.g., Perner et al., 1987), collectively demonstrated that children come to solve these tasks around 4 years of age (Wellman et al., 2001). Furthermore, children showed similar performance shifts around age four in superficially different ToM tasks, indicating a common cognitive capacity underlying these tasks (e.g., Perner and Roessler, 2012). This body of evidence was the basis for assuming that meta-representational ToM emerges in the preschool years, with the fundamental conceptual transition around age four.

However, newer findings in the last two decades have challenged this assumption. Implicit FB tasks infer FB understanding from participants' spontaneous looking behavior or behavioral interactions -rather than explicit verbal answers. They include violation of expectation (VoE), anticipatory looking (AL), and interactive tasks. VoE studies show that children display longer looking times when an agent acts inconsistently with her (false) belief (Onishi and Baillargeon, 2005). AL paradigms reveal that, when presented with an FB scenario, children anticipate the agent to behave in accordance with her (false) belief, as indicated by their looking patterns (Southgate et al., 2007). Finally, findings from interactive tasks suggest that children can take their partners' (false) beliefs into account when responding to them in communicative and other interactions (Buttelmann et al., 2009; Knudsen and Liszkowski, 2012). These findings suggest that some form of implicit sensitivity to others' beliefs is present earlier than age four, perhaps even in the first year of life (see Scott and Baillargeon, 2017 for a review). Furthermore, a converging line of research with adults suggests that these implicit (and largely automatic) capacities remain intact over the lifespan (e.g., Kovács et al., 2010; Samson et al., 2010; Schneider et al., 2012). Far-reaching theoretical accounts build on these seminal findings, in particular nativist theories (Leslie, 2005; Scott and Baillargeon, 2017) or two-systems approaches (Apperly and Butterfill, 2009; Low et al., 2016).

According to nativist accounts, the findings from the implicit FB tasks suggest that ToM is a domain-specific capacity that is in place early in ontogeny or maybe even innate. Nativist accounts claim that explicit verbal FB tasks cannot uncover these early competencies because of performance factors, such as linguistic and inhibitory demands of the standard FB tasks (e.g., Leslie, 2005; Scott and Baillargeon, 2017). According to two-system accounts, the findings of the implicit FB tasks could tap an early-developing mind-reading system. This system, in contrast to the full-fledged flexible meta-representational system that develops in a more protracted form, is evolutionarily ancient, emerges early in ontogeny, and operates efficiently and broadly automatically, but is crucially limited in its (meta-) representational powers (Apperly and Butterfill, 2009; Low et al., 2016).

However, the empirical basis for the implicit ToM understanding is not solid yet. To date, only a relatively small number of studies provided positive findings for implicit ToM measures. Most of these studies were conducted by relatively few labs, and many had small sample sizes (= < 10 per condition in Onishi and Baillargeon, 2005; Southgate et al., 2007; Senju et al., 2010; or = < 25 per condition in Buttelmann et al., 2009). Moreover, the subsequent replication attempts with infants, children, and adults have produced mixed findings and raised doubts about the replicability and validity of the existing implicit ToM measures (for reviews, see Kulke et al., 2018a; Barone et al., 2019; Rakoczy, 2022). One can identify at least three issues regarding the standard implicit measures of ToM. First, all measures (i.e., VoE, AL, and interaction paradigms) encountered failed replication attempts in the last years (e.g., Yott and Poulin-Dubois, 2016; Burnside et al., 2018; Dörrenberg et al., 2018; Poulin-Dubois and Yott, 2018; Powell et al., 2018; Priewasser et al., 2018; Schuwerk et al., 2018). These replication attempts ranged from conceptual to direct, or even self-replications (e.g., Kampis et al., 2021), which makes it hard to attribute the failed replications to poor implementations of the original procedures. Second, all types of implicit ToM measures have shown poor construct validity so far: the earlier positive findings were replicable in some of the later replication attempts only under certain conditions with confounds (e.g., imbalanced number of cues for one answer). When these confounds were eliminated by introducing appropriate controls, the effects disappeared and the measure in question became non-replicable (e.g., Low and Watts, 2013; Powell et al., 2018; Priewasser et al., 2018; Kulke et al., 2018b). Lastly, implicit tasks have been shown to lack convergent validity. Several recent studies found minimal or no systematic correlations between the three standard measures of implicit ToM, nor even with the different tasks of the same type, which supposedly tap the same ability (e.g., Yott and Poulin-Dubois, 2016; Dörrenberg et al., 2018; Poulin-Dubois and Yott, 2018; Powell et al., 2018; Kulke et al., 2018a,b). These issues point to serious reliability and validity issues regarding the standard implicit measures, i.e., VoE, AL, and interaction paradigms. Thus, whether implicit measures reveal robust evidence for ToM in infancy and automatic and implicit forms of ToM throughout the lifespan remains unclear. Additionally, it remains unclear how early implicit and later explicit ToM performance may be related developmentally, with some studies speaking for continuity (Sodian et al., 2016, 2020) why others fail to replicate longitudinal continuity patterns (Poulin-Dubois et al., 2023; for recent debate see and Sodian, 2023 and Poulin-Dubois et al., 2020).

However, in the last decade, another phenomenon has come into focus as a potentially promising indicator of implicit perspective-taking: Altercentric interference or altercentric bias effects suggest that our own judgments or behaviors are influenced by how other people perceive the world, indicating that we implicitly represent their beliefs and perspectives - even when those are entirely irrelevant or interfere with our own task (Kampis and Southgate, 2020; Southgate, 2020). For example, across different studies, participants were found to be slower and more error-prone in counting objects if another agent was present in the scene but had an incongruent perspective on the object (e.g., only saw a subset of objects; Kovács et al., 2010; Samson et al., 2010). Interestingly, the interference seems to occur spontaneously and automatically, without subjects consciously or intentionally focusing on others' perspectives. Theoretically, this bias could thus reflect more unambiguously implicit ToM processes than the standard implicit FB tasks (i.e., VoL, AL, or interactive tasks). And from a methodological perspective, altercentric bias tasks have several potential advantages over typical implicit FB tasks. For instance, they can provide more fine-grained, continuous measures of implicit ToM (participants can be more or less subject to altercentric interference).

In addition, altercentric bias measures are particularly interesting and promising from a methodological point of view: they allow researchers to construct structurally analogous tasks to tap implicit and explicit ToM within one task format such that the two types of tasks differ merely with regard to the critical test question. On the one hand, in implicit versions employing altercentric bias, participants are asked to make a factual judgment about the world (e.g., How many dots are there? / Where is an object?) in the presence of an irrelevant agent who does or does not share their perspective. If participants are slower or more error-prone in their own factual judgments when the other agent has a deviant perspective, this indicates altercentric bias. On the other hand, the explicit versions exploit the so-called egocentric bias, which refers to the influences of one's own knowledge when judging others' perspectives. In the explicit versions, participants are asked about the other agent's perspective or behavior (e.g., How many dots does the agent see? / Where will the agent look for the object?). If participants become slower or more error-prone in these perspective judgments when their own perspective is different from the agents', this indicates egocentric bias. This bias could then be used to infer the explicit ToM ability of participants: more interference from one's own perspective -even if the task asks to take other's perspective- means poorer ToM.

So far, these two biases have been implemented together in the so-called Dot Perspective Task (Samson et al., 2010). In this task, adult participants were asked to judge the number of dots presented in a scene either from their own perspective (SELF condition) or as seen by an on-screen avatar (OTHER condition). Each condition featured two types of trials: consistent versus inconsistent. In consistent trials, all dots were equally visible for the participant and the avatar, and their perspectives were thus consistent. In inconsistent trials, some of the dots visible for the participant were behind the avatar, therefore, not visible to it. The two perspectives were thus inconsistent. This study revealed that participants were slower and made more errors when detecting the number of dots in inconsistent trials compared to consistent trials. These results were interpreted as providing evidence for both altercentric and egocentric interference effects in SELF and OTHER conditions, respectively.

The dot perspective task has been one of the few measures in which both biases are obtained using the same task format. However, this task is not free from replication issues and validity debates. Some of the later studies using the variations of the dot-perspective task either revealed no bias (e.g., Conway et al., 2017) or the biases were subject to alternative explanations by domain-general mechanisms rather than implicit mentalizing (e.g., Santiesteban et al., 2014; Cole et al., 2016; Conway et al., 2017; O'Grady et al., 2017). As an example of the latter, Santiesteban et al. (2014) tested participants in two different versions of the dot perspective task: some trials featured an avatar as in previous experiments, and in the other trials, the avatar was replaced by an arrow with similar low-level features such as color, size, and orientation. They found comparable altercentric interference effects in avatar and arrow conditions, suggesting that the so-called altercentric bias effects may reflect general cognitive processes such as spatial cueing rather than specifically social-cognitive processes of perspective-taking (but see Michael et al., 2018).

The rationale of the present study is thus to construct alternative tasks to tap implicit and explicit perspective-taking abilities through altercentric and egocentric biases, respectively. To this end, we are capitalizing on an established explicit continuous FB task, the so-called Sandbox task (e.g., Sommerville et al., 2013). In this task, like in standard change-of-location tasks, participants need to track where an agent believes an object to be that was re-located in her absence. But rather than using discrete locations (the object was in box 1 and then moved to box 2), the object is placed and re-located within a continuous space such as a sandbox. Participants can thus track more or less precisely where the object was and now is and where the agent believes it to be. This task has been used with participants of a wide age range (e.g., three- and five-year-olds; young, middle-aged, and senior adults) in the forms of real-object, paper-pencil, or computerized versions (Bernstein et al., 2011; Begeer et al., 2012; Sommerville et al., 2013; Coburn et al., 2015; Mahy et al., 2017; Samuel et al., 2018a,b). In the existing, explicit egocentric bias version, participants witness an object being re-located from Location 1 to Location 2 in the absence of the agent. Then they are asked where the agent would look for the object. Egocentric interference effects suggest that participants' answers would be biased away from Location 1 (i.e., correct answer) in the direction of Location 2 (i.e., object's current location) as they know that the object is now at Location 2. This task lends itself nicely to developing an analogous implicit or altercentric bias version. This new implicit version is just like the explicit version of the task, except for one crucial difference in the test question. In the altercentric bias version, participants are questioned on the object's current location rather than where the agent would look for it. If they are subject to altercentric interference effects, their answers will deviate from Location 2 (i.e., correct answer) in the direction of Location 1 (i.e., agent's belief location). The construction of closely matched explicit and implicit versions of the Sandbox FB task thus allows us to investigate and contrast egocentric and altercentric biases within the same task format in fine-grained ways. Here, we report three studies that explore the viability of such combined task formats with adults. All the studies reported in this paper were conducted online during and due to the Covid-19 pandemic.[1]

2. Study 1

The aim of Study 1 was to investigate whether an online adaptation of the Sandbox task could tap both egocentric and altercentric biases. This study also explored the presence and the magnitude of these biases in two different cultures (German/Western and Turkish/Eastern). Western societies are regarded as independent cultures as they emphasize attention to self and

1 We had started a pilot study that aimed to replicate earlier Sandbox task results (which revealed egocentric interference effects) using the paper-pencil version of the task and the original materials from this version. We had to stop the data collection for this version before reaching the necessary sample size due to the pandemic and we switched to online testing. The data collection for the pilot study has been recently completed. We also conducted a second pilot with more engaging materials. These are reported in the Supplementary Documents.

individualist self-construals. In contrary, Eastern cultures emphasize being attentive to others and harmony between individuals, leading to more interdependent self-construals (Markus and Kitayama, 1991). As a result of this difference, Western people could be subject to egocentric bias more than Eastern people as they operate with a focus on themselves. In contrast, Eastern people could be more prone to altercentric bias as they prioritize others' perspectives above their own. Mixed results have been provided for this potential difference so far. For example, Wu and Keysar (2007) found that Chinese adults showed less egocentric bias than their American counterparts. By contrast, Wang et al. (2019) did not find any difference between Taiwanese and British adults in terms of egocentric and altercentric biases. These studies have either measured only one bias or measured the biases in separate tasks. The current study aims to explore potential differences in both egocentric and altercentric biases within one task format and with samples that have not been compared in this context before (i.e., German and Turkish samples). The study was preregistered.[2]

2.1. Method

2.1.1. Participants

Participants were recruited through social media advertisements and e-mail announcements. All participants were tested online in unmoderated sessions via Qualtrics.[3] We used G*POWER (Faul et al., 2009) to conduct a power analysis and determine the sample size. We aimed to obtain 0.95 power to detect a medium effect size of 0.54 at the standard 0.05 alpha error probability with a more conservative two-tailed paired-samples t-test. The effect size was based on an earlier study by Samuel et al. (2018b). They found a significant difference between experimental and control trials of the egocentric bias condition using a computerized version of the Sandbox task. Since the main aim of the current study was to tap biases revealed as the differences between experimental and control trials, we based the sample size rationale on within-subject comparisons rather than between-subject comparisons. The analysis revealed a required sample size of 47 participants for each bias measured within a group: 94 participants per group and 188 participants in total. Thus, the final sample consisted of 188 participants: 94 German (72 females, $M_{age} = 28.11$, age range: 18 to 62) and 94 Turkish (63 females, $M_{age} = 27.05$, age range: 18 to 57) adults. All participants were tested in their native languages, consented to the study, and, upon completing the study, became eligible for a lottery that distributed vouchers from online bookstores.

In order to have a final sample of 188 participants, we initially tested 356 participants. One hundred fifteen participants were excluded from the final sample as they did not complete all trials ($M_{CompletedTrials} = 3.99$, $SD = 2.06$, range: 1 to 7), 34 of them exceeded the time limit allocated for the study (i.e., 30 min), 16 of them had technical problems, and 3 of them reported that their native languages were different from the desired languages (i.e., German and Turkish).

2.1.2. Materials

2.1.2.1. The sandbox task

The scenarios used in our study were based on those used in Mahy et al. (2017). They always followed the same storyline: Agent A hides an object in Location 1, but then the object is transferred to Location 2 by Agent B either in the absence (False Belief) or in the presence of Agent A (True Belief). After the scenarios were presented, participants worked on a word-search puzzle for 20 s. Puzzles prevented using perceptual cues to answer the question and were created by inserting six randomly generated city names (from participants' respective countries) into a 10 × 10 word-search puzzle using a puzzle maker website.[4]

The images presented to the participants (1,500 × 1,125 pixels) displayed a rectangular container (1,220 × 150 pixels) positioned in the middle of the image and text above the container. The crosses (37.5 × 37.5 pixels) on the container indicated a hidden object's initial and final locations. These locations were always 746 pixels apart, but their relative position changed across trials to prevent participants from learning the locations. In all of our studies, the direction of relocation was counterbalanced: in half of the trials, the object was transferred from left to right, and in the other half, the transfer was from right to left. The objects always crossed the midline of the sandbox during the transposition.

In a mixed design study, participants were presented with either the egocentric or altercentric bias conditions, consisting of experimental and control trials. These two conditions utilized the same task format and scenarios but differed in their test questions (see Figure 1 for examples). In the egocentric bias condition, participants were asked either where Agent A, who has a false belief about the object's location, would look for the object upon return (experimental trials; "Where will X look for the object?") or where s/he hid the object before leaving the scene (control trials; "Where did X hide the object?"). In both of these trials, the correct answer is around Location 1. Participants are expected to deviate in the direction of Location 2 in the experimental trials as they know that the object is actually at Location 2, and this knowledge is expected to interfere with their judgments of others' perspectives and behaviors. In the altercentric bias condition, the test question always inquired where the object currently is, but the preceding scenario differed in terms of Agent A's belief (which was irrelevant to the test question) in experimental and control trials. In the control trials, Agent A witnessed the relocation, therefore, had a true belief about the object's current location. In the experimental trials, the relocation happened in Agent A's absence; therefore, Agent A had a false belief about the object's location. In both of these trials, the correct answer is around Location 2. Participants are expected to deviate in the direction of Location 1 in the experimental trials but not in the control trials, as the

2 https://osf.io/36exv
3 https://www.qualtrics.com
4 https://puzzlemaker.discoveryeducation.com

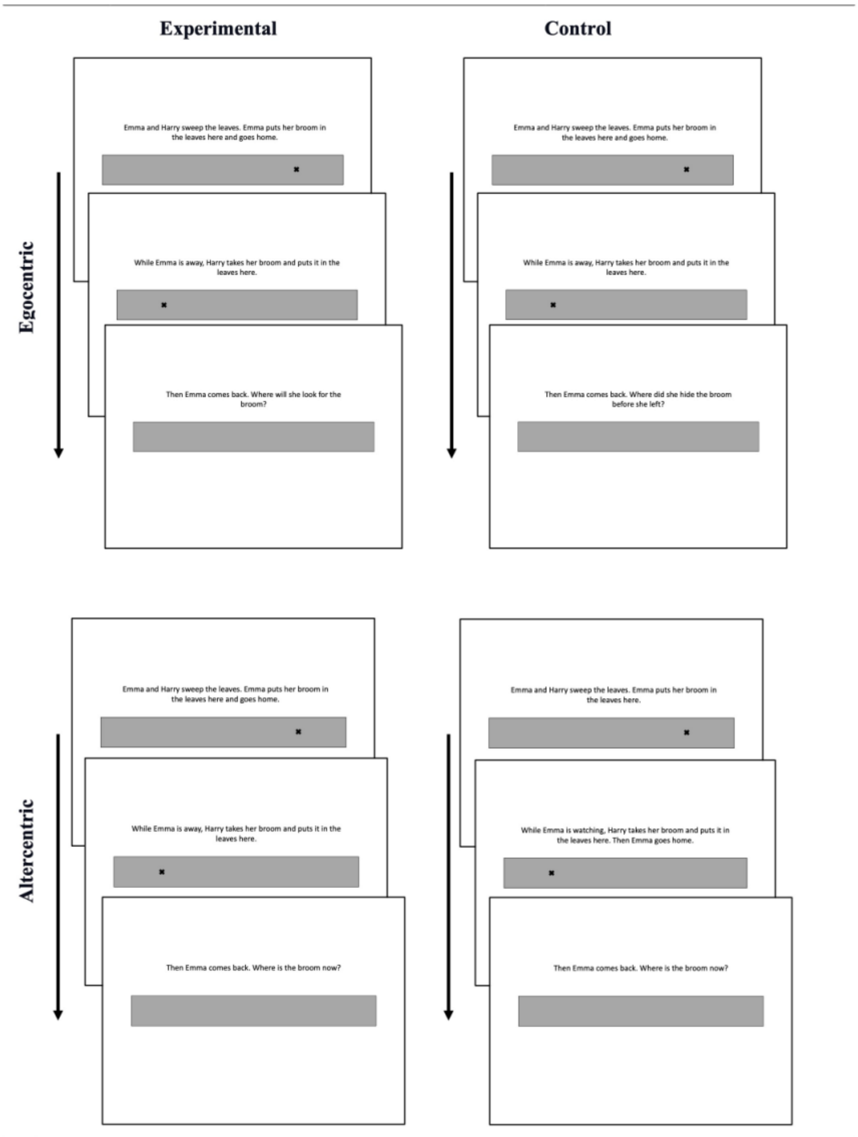


FIGURE 1
Sample materials from Study 1.

agent in the experimental trials thinks that the object is still at Location 1.

2.1.2.2. Self-construal scale

In this 30-item Likert-type scale, participants evaluated the strength of their self-construal in terms of independency and interdependency (15 questions per each type of self-construal, Singelis, 1994; see Supplementary materials for example items). For our study, the original scale was translated and backtranslated by German and Turkish native speakers, who were also fluent in English. Each participant received two independence and interdependence scores ranging between 15 and 105.

10.3389/fpsyg.2023.1142302

2.1.2.3. Demographic questionnaire

This questionnaire consisted of four questions asking the participants' age, gender, highest educational degree achieved, and native language.

2.1.3. Design and procedure

In a mixed design, Turkish and German participants were randomly assigned either to the egocentric or the altercentric bias condition, which included two types of trials: experimental and control. Each participant completed four experimental and four control trials presented in blocks (the order of the blocks counterbalanced) and one filler trial in between.

After consenting to the study, participants started test sessions with a calibration task. Participants were asked to click on the center of six crosses presented on the screen in this task. These crosses represented the endpoints of the Sandbox. Calibration trials provided information about mouse cursors' sensitivity. Then the Sandbox task, the Self-Construal Scale, and the demographic questionnaire were completed in this fixed order. In the end, participants were debriefed about the real aim of the study and registered their contact information if they wanted to participate in the lottery. The study took approximately 20 min.

2.1.4. Bias calculation and analysis

Biases were inferred from the object location measure: the horizontal distance (in pixels) between the correct location (i.e., L1 in egocentric condition, L2 in altercentric condition) and the participant's response (see Figure 2 for an illustration). If the participants' responses were biased toward the wrong location (i.e., between the right and wrong answer, toward to middle of the screen), they received a positive object location value. The responses biased away from the wrong location (i.e., in the direction of the edge of the Sandbox/screen, rather than the middle) received a negative object location value. Once the object location measure was computed for each trial, we calculated the average object location measure in experimental and control trials for each participant. The averages were calculated in two ways: (a) all responses were included in the averages (as done in the original Sandbox task studies), and (b) the completely wrong answers (i.e., responses that were closer to the incorrect location than the correct location) were excluded from the averages. The latter method aimed to exclude the trials to which participants did not pay enough attention. We argue that adults are expected to have full-fledged perspective-taking abilities; therefore, completely wrong answers would reflect participants' failures of attention and could be excluded from the data for explorative purposes (e.g., does a bias exist when only the attended trials are considered?). The average scores were then used to deduce biases: if the average deviation in the experimental trials is bigger than the control trials, this indicates bias. As a result, separate within-subject comparisons were conducted with and without wrong answers to see if a bias exists in different conditions and groups.

In addition to investigating whether a bias exist, we also explored whether German and Turkish participants differed in terms of (a) independency/interdependency and (b) the magnitude of any bias. To compare the biases shown by different groups, we created a pure bias score for each bias type by subtracting the average deviation in control trials from the average deviation in experimental trials. This score enabled us to compare the two groups directly on the difference between experimental and control trials, namely, the deviation expected due to the perspective-taking. Finally, we explored if pure bias scores were related to the level of independency and interdependency.

Following Samuel et al. (2018b), we used non-parametric tests (e.g., matched-pair Wilcoxon signed-rank Test, Mann–Whitney U test) for both within- and between-subject comparisons, as the response data were not normally distributed. For all of our studies, we also ran analyses with parametric tests as we had initially expected a continuous distribution of answers. The pattern of results and significances remained the same across all studies.

2.2. Results

2.2.1. Within-subject comparisons: do the biases exist?

The mean deviations in experimental and control trials can be seen in Figure 3. We started our analysis by comparing these

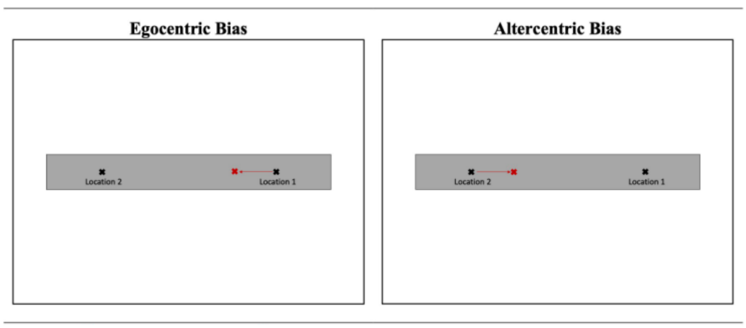


FIGURE 2
Bias calculation in Study 1. X: Participant's answer. →: Object location measure.

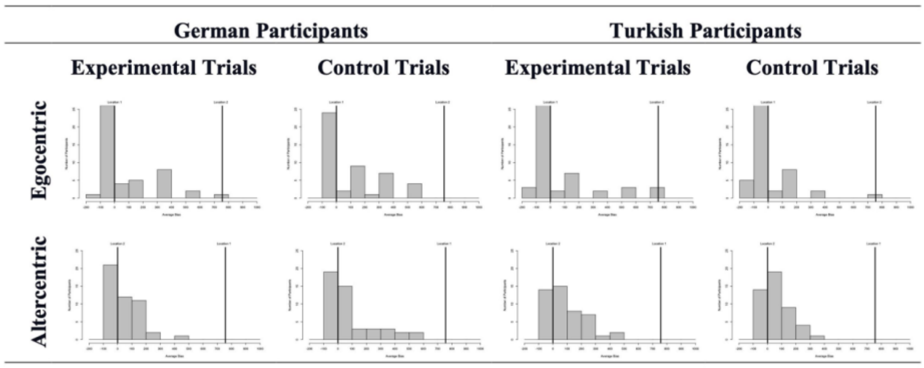


FIGURE 3
Average object location deviations in experimental and control trials across conditions and participants in Study 1. Here (and in the following studies) the binned data is presented for ease of depiction, but the analyses were run with the continous data.

average deviations across biases and groups. Almost none of these analyses revealed a difference between experimental and control trials. Specifically, the experimental vs. control trials differed in neither the egocentric (experimental: *Mdn* = −15.0; control: *Mdn* = −13.0) nor the altercentric bias (experimental: *Mdn* = 5.0; control: *Mdn* = 8.5) conditions for German adults, Z = -0.540, p = 0.589 and Z = -1.360, p = 0.174, respectively. Excluding the wrong answers from the analysis did not change the sample size as none of the participants failed all trials. The general picture drawn by the results remained the same too: experimental and control trials were very similar for both egocentric (experimental: *Mdn* = −39.63; control: *Mdn* = −47.0) and altercentric (experimental: *Mdn* = −15.0; control: *Mdn* = −3.25) bias conditions, Z = -1.830, p = 0.067 and Z = -1.423, p = 0.155, respectively. For Turkish participants, experimental (*Mdn* = −27.75) and control (*Mdn* = −39.5) trials only differed for the egocentric bias when all responses were included in the analysis, Z = -2.349, p = 0.019; but not after the wrong answers were excluded (experimental: *Mdn* = −45.0; control: *Mdn* = −51.375), Z = -1.654, p = 0.098. No altercentric bias was found regardless of calculation, i.e., with (experimental: *Mdn* = 28.75; control: *Mdn* = 6.5) or without (experimental: *Mdn* = −0.25; control: *Mdn* = 0.25) wrong answers, Z = -1.365, p = 0.172 and Z = -0.561, p = 0.575, respectively.

2.2.2. Between-subject comparisons: do the groups differ?

We first compared German and Turkish adults in terms of the independency-interdependency scores to see if the two groups really differed. Results revealed a difference only in one domain: Turkish adults scored higher on the interdependency measure than German adults, U = 3105.50, Z = −3.520, p < 0.001; however, they did not differ in the independency sub-scale, U = 3756.50, Z = −1.775, p = 0.076. Then we explored if the two cultures differed in the pure bias shown in egocentric and altercentric conditions. Mann–Whitney tests revealed a difference neither for the pure egocentric nor the pure altercentric bias of German and Turkish adults, U = 990.00, Z = −0.178, p = 0.859, and U = 920.00, Z = −1.395, p = 0.163, respectively. Finally, we investigated whether any pure bias and a type of self-construal were related. Results revealed no significant correlations (all *r*s = < |10|, all *p*s > = 0.09).

2.2.3. Within-subject comparisons with collapsed datasets

As suggested by one of the anonymous reviewers, in order to increase power, we repeated the within-subject comparisons by collapsing the two groups since they did not differ. The analysis revealed a difference between experimental (*Mdn* = −41.0) and control (*Mdn* = −47.67) trials for egocentric bias (Z = -2.559, p = 0.01) only when the wrong answers were excluded from the analysis; but not for the whole sample (experimental: *Mdn* = −22.38, control: *Mdn* = −27.5; Z = -1.245, p = 0.21). No difference was observed in altercentric bias condition, regardless of the fact that the wrong answers were included (experimental: *Mdn* = −5.08, control: *Mdn* = −2.88; Z = -0.064, p = 0.95) or excluded (experimental: *Mdn* = 19.63, control: *Mdn* = 7.88; Z = -0.601, p = 0.55).

2.3. Discussion

Study 1 revealed almost no differences between experimental and control trials (except the collapsed analyses, where deviations were negative and were not biased toward the second location); hence no evidence for egocentric or altercentric biases and did not reveal any cross-cultural differences in egocentric and altercentric biases either.

The null results found in Study 1 are difficult to interpret, and they should be evaluated with caution for two reasons. First, there was a very high dropout rate (almost 50%). Even though some of these dropouts occurred due to technical issues or timeout, many participants intentionally stopped participating without completing the study simply because they were bored due to the dull materials. We suspect that the not-so-engaging task materials might have caused our remaining participants to fail to pay enough attention to the task, which could have made the task less reliable. Secondly, the altercentric bias measure may have not been spontaneous enough to tap automatic interference effects. Possibly, with too much time, participants begin to reflect on and evaluate their own perspective and correct any potential spontaneous biases. Therefore, although, in theory, we expected the Sandbox task to tap altercentric biases as well as egocentric biases, this task may not be suitable to detect altercentric interferences. These two issues were addressed in the following studies.

3. Study 2

In response to the potential shortcomings of Study 1, Study 2 used more engaging materials in the form of animated videos and presented fewer trials (i.e., two trials per trial type instead of four). Also, to have more spontaneous behavioral measures, the present study combined the Sandbox task with mouse-tracking measures. These measures have previously been used to document altercentric bias (e.g., Van der Wel et al., 2014). When subjects were asked to move their mouse cursors to the target object's location, they took a little detour on their way to their answers when another agent in the scenario had a belief that differs from their own. However, when the participant and the agent shared a belief, participants followed a more direct route while moving their mouse cursors to mark the target location. Therefore, the area between the detour in the direction of the wrong answer and the direct line from the starting point to the target location indicated whether and to what degree participants engaged in altercentric bias. Since mouse-tracking measure is inferred from spontaneous motor responses, it constitutes a more suitable alternative to tap implicit biases, especially altercentric bias, than the object location measure of the Sandbox task. More specifically, this task is less subject to reflections and evaluations as it is not about the content of the final judgment, and it does not respond to an explicit trigger. Rather it occurs spontaneously and is manifested via automatic motor responses. Therefore, these measures are optimally suited to reveal online processes which are more automatic and spontaneous, such as altercentric interference effects, and could tap these biases more reliably (Van der Wel et al., 2014). The study was preregistered.[5]

3.1. Method

3.1.1. Participants

In Study 2, we tested both English- and German-speaking adults online. These two groups were included not because we expected to see a difference but because of practical reasons. Namely, the platform where the study was published had a bigger pool of English-speaking participants, which meant more representative data. English- and German-speaking participants were presented with the same materials, except the language of the materials and instructions. The experimenter translated the materials, which were then double-checked by a native English speaker, who also did the voice recordings for English materials. Participants were recruited through Prolific[6] and they were tested in unmoderated sessions via Labvanced.[7] The sample size reasoning was the same as in the first study. Therefore, we collected 47 participants for each bias measured within a group, 94 participants per group and 188 participants in total. More specifically, we tested 94 German-speaking (34 females, $M_{age} = 30.60$, age range: 18 to 63) and 94 English-speaking (52 females, $M_{age} = 33.14$, age range: 18 to 65) adults. All participants were tested in their native languages, consented to the study, and received compensation upon completing it. At the beginning of the study, participants were questioned on their demographic information, including age, gender, education level, and native language.

In order to have a final sample of 188 participants, we initially tested 198 participants. Five participants were excluded from the final sample as they had technical issues, 3 of them did not complete all trials ($M_{CompletedTrials} = 2.67$, $SD = 0.58$, range: 2 to 3), and 2 of them exceeded the time limit allocated for the study (i.e., 20 min).

3.1.2. Materials

The scenarios used in the second study were similar to the first study. They always featured two agents and an object, and they followed the same storyline: Agent A hides an object in Location 1, but then the object is transferred to Location 2 by Agent B either in the absence (False Belief) or the presence of Agent A (True Belief). After the videos were presented, participants searched and marked dots on the screen for 10 s. This task aimed to prevent participants from using perceptual cues to answer the question and served as a calibration check. This task was created directly on Labvanced. It generated colorful dots (20 × 20 units) and presented them on the screen one by one. Participants were instructed to find and click on these dots. Besides the already mentioned differences, the scenarios and questions used in the control and experimental trials of Study 2 closely resembled Study 1. Figure 4 depicts the important sections of the videos used in Study 2 (and Study 3).

The materials used in the second study were different from the first study in several ways. First, in Study 2, the stories were presented as animated videos instead of still pictures. Second, participants did not see any sandbox drawing with borders. Instead, the whole screen was used as the hiding area, and there was no visual cue that could potentially cause anchoring effects. Third, we opted for a simpler distractor. More specifically, participants were asked to click on dots that appeared on the screen instead of word-search puzzles. Also, in Study 2, the distraction task lasted 10 s instead of 20 s. This distractor task also allowed us to check the calibration of their mouse cursors throughout the experiment and control it if necessary. Fourth, we decreased the number of trials per condition. Participants answered two experimental and two control trials, with one filler trial in between. Finally, adding the mouse-tracking measures called for a small but crucial change in the materials. As in Study 1, participants were asked the test question after completing the distraction task. Unlike Study 1, however, they first needed to press a record button located at the bottom middle of the screen to be eligible for responding to test questions. Only then were they able to do any marking on the screen. This procedure ensured that all participants started to move their mouse cursors at the same point on the screen.

The videos presented to the participants were created in Vyond,[8] then converted into the movie format, and uploaded to Labvanced. The videos (800 × 450 units) were always presented in landscape

5 https://osf.io/rqxs8

6 https://www.prolific.co

7 https://www.labvanced.com

8 https://www.vyond.com

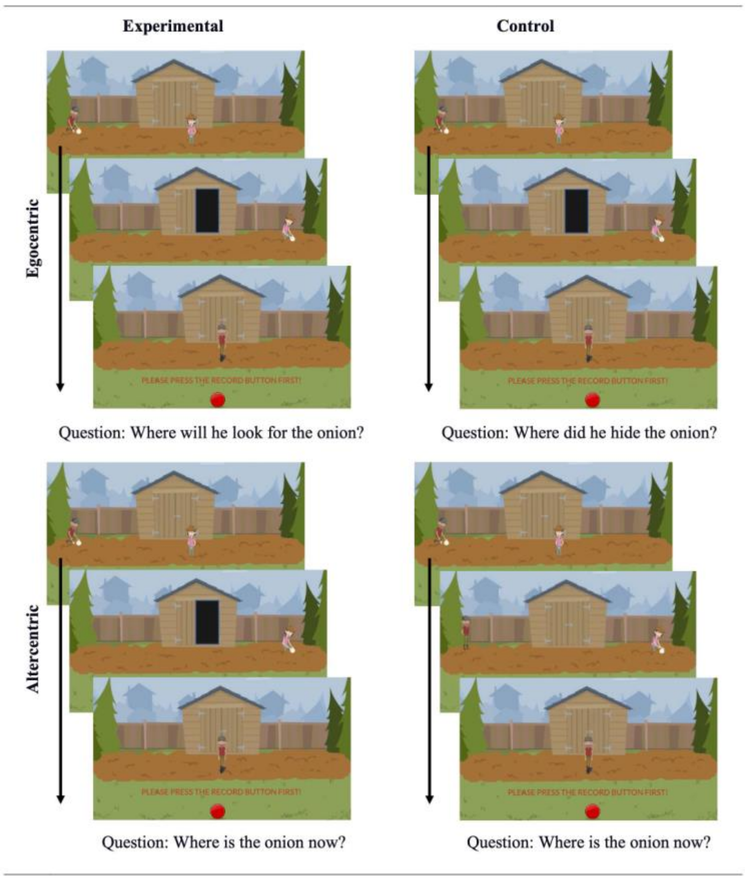


FIGURE 4
Sample materials from Studies 2 and 3.

format. The hiding locations were always 557.5 units apart, which is almost 70% of the screen. Their relative position changed on the same horizontal line across trials to prevent participants from learning the locations.

3.1.3. Design and procedure

In a mixed design, English- and German-speaking participants were randomly assigned either to the egocentric or the altercentric bias condition, which included two types of trials: experimental and control. Each participant completed two experimental and two control trials presented in blocks (the order of the blocks counterbalanced) and one filler trial in between.

After consenting to the study, participants started test sessions with practice trials, where participants were asked to click on the center of two crosses presented on the screen. These crosses appeared at random locations, but their places were consistent across participants. These practice trials provided information about the sensitivity of mouse cursors and taught participants that they should start moving their cursors from a designated point. Namely, to be able to click on a cross, participants were required to first click on a record button, which was always presented at the bottom middle of the screen. This record button corresponded to the starting point of the mouse-tracking measure in later trials. After completing the practice trials, participants indicated which device was used to mark locations: a built-in touchpad or a mouse. Then they started the actual trials of the Sandbox task. In the end, participants were directed back to Prolific and received their payments. The study took approximately 12 min.

3.1.4. Bias calculation and analysis

Bias scores were calculated in two ways in this study (see Figure 5). The first way was the object location measure as

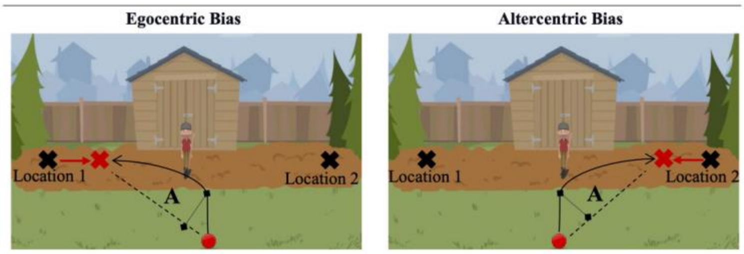


FIGURE 5
Bias calculation in Studies 2 and 3. X: Participant's answer. →: Object location measure, – – –: direct path to the participant's answer. →: Participant's mouse trajectory. A: mouse tracking measure. ♦– – –♦: an example orthogonal line drawn from the mouse coordinates to the direct path.

calculated in Study 1 (i.e., the horizontal distance between the correct location and the participant's response), which was then used to calculate the average biases in experimental and control trials with and without wrong answers. The other way to measure egocentric and altercentric biases was to track the mouse movements of the participants. For this measure, we calculated how many units of area were between the path the participants followed on the way to their answers and the most direct (shortest) path to their answers. More specifically, Labvanced provided us with the time series and the coordinates of participants' mouse movements on a standardized 800×450 unit surface area. These coordinates were then used to infer the actual trajectories followed by the participants, and the start and end points of the trajectories were used to compute the direct path to their answers. Because the trajectories varied in duration, they were standardized into 100 time steps and 100 coordinates using linear interpolation. Then, orthogonal lines were drawn from each coordinate [(e.g., x_1, y_1) to the direct path (e.g., x_2, y_2)] and their lengths were calculated by using the distance formula [$d = \sqrt{((x_2 - x_1)^2 + (y_2 - y_1)^2)}$]. Then the length of every possible line was summed up to get the area between the actual trajectory and direct path for each trial. If participants showed a detour toward the wrong answer (i.e., Location 2 in egocentric bias and Location 1 in altercentric bias), the area between this detour and the direct route is assigned a positive value as this detour indicated that participants were biased. If participants show a detour in the opposite direction (i.e., between the direct path and the edge of the screen) the area is given a negative value. Then the average mouse trajectory measure was calculated for experimental and control trials in each condition. The mouse-tracking measure was only calculated for correct answers; therefore, only one average value was obtained per trial type (experimental versus control).

Finally, similar to the first study, separate within-subject comparisons were conducted for egocentric and altercentric bias conditions and English- and German-speaking participants to see if experimental and control trials differ; hence a bias exists. Following Samuel et al. (2018b) and the first study of this paper, we used non-parametric tests (i.e., matched-pair Wilcoxon signed-rank Test) to conduct within-subject comparisons as the response data were not normally distributed.

3.2. Results

The mean object location deviations in experimental and control trials are depicted in Figure 6. We first separately compared the average biases in experimental and control trials for each condition (i.e., egocentric and altercentric) and group (German- and English-speaking adults). None of these analyses suggested a difference between experimental and control trials; hence no bias was revealed. As object-location measures have shown, the experimental versus control trials differed in neither the egocentric (experimental: $Mdn = -4.17$; control: $Mdn = -5.51$) nor the altercentric (experimental: $Mdn = 1.88$; control: $Mdn = 1.56$) bias conditions for German-speaking adults, $Z = -1.376$, $p = 0.169$ and $Z = -0.317$, $p = 0.751$, respectively. Excluding the wrong answers from the analysis did not change the direction of these results: experimental and control trials were not different from each other, neither for egocentric (experimental: $Mdn = -5.23$; control: $Mdn = -6.2$) nor altercentric (experimental: $Mdn = 1.51$; control: $Mdn = -0.94$) bias conditions, $Z = -0.993$, $p = 0.321$ and $Z = -0.709$, $p = 0.478$, respectively. The mouse-tracking measure revealed similar results: mouse movements in experimental and control trials resembled each other both in egocentric (experimental: $Mdn = 3.23$; control: $Mdn = 3.87$) and altercentric (experimental: $Mdn = 3.75$; control: $Mdn = 3.73$) bias conditions: $Z = -0.672$, $p = 0.502$ and $Z = -0.550$, $p = 0.582$, respectively.

The results obtained from English-speaking participants were in line with those from German-speaking adults. Accordingly, the object-location measure did not reveal any difference between experimental and control trials, neither in the egocentric (experimental: $Mdn = 1.79$; control: $Mdn = 5.52$) nor the altercentric (experimental: $Mdn = -0.71$; control: $Mdn = 3.09$) bias conditions: $Z = -0.492$, $p = 0.622$ and $Z = -1.144$, $p = 0.253$. The results remained almost the same after excluding the wrong answers: the difference between trial types approached significance in the egocentric bias version (experimental: $Mdn = -4.38$; control: $Mdn = 4.53$), $Z = -1.781$, $p = 0.075$; however, no such pattern has been shown for the altercentric bias condition (experimental: $Mdn = -1.24$; control: $Mdn = 3.09$), $Z = -1.453$, $p = 0.146$. Mouse-tracking measures did not reveal any difference, neither for egocentric (experimental: $Mdn = 2.88$; control:

10.3389/fpsyg.2023.1142302

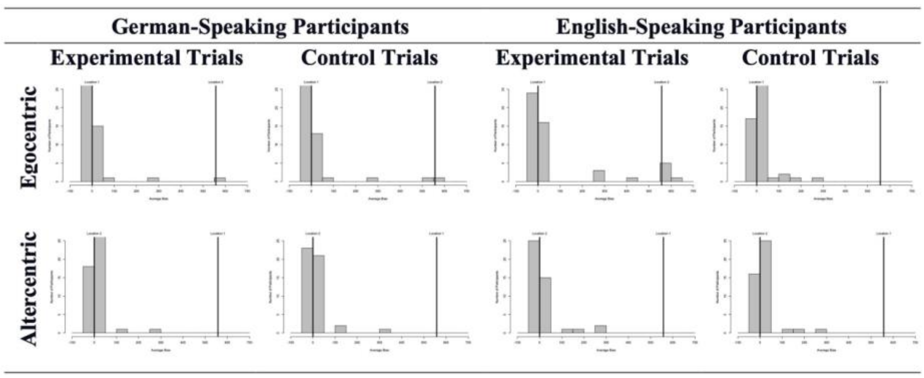


FIGURE 6
Average object location deviations in experimental and control trials across conditions and participants in Study 2.

Mdn = 4.78) nor altercentric (experimental: *Mdn* = −1.24; control: *Mdn* = 3.09) bias conditions, $Z = -1.319$, $p = 0.187$ and $Z = -0.455$, $p = 0.649$, respectively.

Similar to Study 1, in order to increase power, we repeated the within-subject comparisons by collapsing the two language groups. First, we checked if the groups differed in terms of the pure biases they showed. No differences were observed between groups. Then we repeated the within-subject comparisons with the entire dataset of Study 2 and found no difference between experimental and control trials regardless of the measure, bias condition, and whether the wrong answers were excluded or not. More specifically, object-location measures have shown that the experimental and control trials differed in neither the egocentric (experimental: *Mdn* = −4.17; control: *Mdn* = −1.62) nor the altercentric (experimental: *Mdn* = 1.25; control: *Mdn* = 2.81) bias conditions, $Z = -0.480$, $p = 0.63$ and $Z = -0.564$, $p = 0.57$, respectively. Excluding the wrong answers from the analysis did not change these results: experimental and control trials were not different from each other, neither for egocentric (experimental: *Mdn* = −4.58; control: *Mdn* = −2.24) nor altercentric (experimental: *Mdn* = 0.42; control: *Mdn* = 2.29) bias conditions, $Z = -1.847$, $p = 0.07$ and $Z = -0.544$, $p = 0.59$, respectively. The mouse-tracking measure revealed similar results: mouse movements in experimental and control trials resembled each other both in egocentric (experimental: *Mdn* = 2.88; control: *Mdn* = 3.98) and altercentric (experimental: *Mdn* = 3.86; control: *Mdn* = 3.23) bias conditions: $Z = -1.348$, $p = 0.18$ and $Z = -0.716$, $p = 0.47$, respectively.

3.3. Discussion

Study 2 was successful in the sense that the dropout rate declined considerably (5%) compared to Study 1. However, there was still no evidence for any bias in the previously used (i.e., object location) or in the newly added (i.e., mouse tracking) measure. There is still one possibility that remains open and might provide some potential explanations of the null results: socio-cognitive biases may be sensitive to test designs and procedures and reveal themselves only under specific circumstances. One such design could be within-subject studies where participants are tested on both biases in blocks. Existing evidence is compatible with the possibility that altercentric biases only arise in such mixed-block designs (Furlanetto et al., 2016; Speiger et al., 2022) and not in analogous single-block designs like the one used in Study 2 (e.g., Conway et al., 2017). We investigated this possibility in Study 3.

4. Study 3

Study 3 used the Sandbox task as in Study 2, including both object location and mouse-tracking measures in a within-subjects design with mixed altercentric and egocentric bias blocks. The study was preregistered.[9]

4.1. Method

4.1.1. Participants

Fifty-four German-speaking adults (18 females, $M_{age} = 28.92$, age range: 18 to 66) were tested via Prolific on an online study created through Labvanced. Since the demographics data provided in Study 2 showed that Prolific had access to a representative German-speaking sample, Study 3 tested only German-speaking adults. We used G*POWER (Faul et al., 2009) to conduct a power analysis and determine the sample size. Our goal was to obtain 0.95 power to detect a medium effect size of 0.50 at the standard 0.05 alpha error probability in a MANOVA, in which between- and within-subject comparisons would be conducted with both object location and mouse-tracking measures. This sample size also matched the sample size of Conway et al. (2017) study, where an altercentric bias was found only with a within-subject design. Participants were tested in German, consented

9 https://osf.io/e5a23

to the study, and received compensation upon completing the study. At the beginning of the study, participants were questioned on their demographic information, including age, gender, education level, and native language.

Only one participant was excluded from the initial data set due to incomplete trials (only three trials were completed by this participant).

4.1.2. Materials

The materials used in Study 3 were the same as in Study 2, except that Study 3 included more trials because of the within-subject design and, therefore, more stories than Study 2. Based on the same storyline, four other videos were created for Study 3, again in Vyond. All stories featured two agents, one object, and relocation of the object either in the agent's absence or presence. We did not change anything regarding the practice trials, the distractor task, and the requirement of pressing the record button before answering. We also kept the format of the experimental and control trials the same for both egocentric and altercentric bias conditions.

4.1.3. Design and procedure

In a mixed design, participants were tested on both egocentric and altercentric biases. They were randomly assigned either to the egocentric-first or the altercentric-first condition. Both egocentric and altercentric bias measures included two types of critical trials: experimental and control. Each participant completed two experimental and two control trials presented in blocks per bias (the order of the blocks counterbalanced). This resulted in eight trials in total. Apart from the within-subject testing of the biases, the procedure of Study 3 was the same as in Study 2. The study took approximately 15 min.

4.1.4. Bias calculation and analysis

The biases were calculated in the same way as in Study 2. Experimental and control trials were compared again with non-parametric Mann–Whitney U tests (due to the failure of the normality assumption). We also utilized non-parametric Wilcoxon-Signed Rank tests for between-subject comparisons. Although this analysis is not following the preregistered analysis on which the sample size was based, non-parametric tests were deemed more appropriate as the data were not normally distributed.

4.2. Results and discussion

The mean object location deviations in experimental and control trials are shown in Figure 7. We first compared these averages for each bias separately. For egocentric bias as measured by the Sandbox task, experimental ($Mdn = 2.88$) and control ($Mdn = 3.27$) trials did not differ from each other, $Z = -0.030$, $p = 0.976$; and excluding the wrong answers did not change this result, $Z = -0.598$, $p = 0.550$ (experimental trials: $Mdn = 1.5$; control trials: $Mdn = 3.27$). Mouse-tracking measures provided results along the same lines: no difference was found between experimental ($Mdn = 1.47$) and control ($Mdn = 0.63$) trials, Z = -1.536, $p = 0.125$. The altercentric bias version did not reveal any difference either. When all answers were considered, experimental ($Mdn = 3.31$) and control ($Mdn = 7.9$) trials did not differ from each other, Z = -0.697, $p = 0.486$. When the analysis was repeated with correct answers only, there was still no difference between experimental ($Mdn = 2.71$) and control ($Mdn = 7.06$) trials, Z = -1.54, $p = 0.123$. Mouse-tracking measures did not reveal any difference between trial types, $Z = -1.102$, $p = 0.270$ (experimental trials: $Mdn = 2.56$; control trials: $Mdn = 3.19$). We also compared the experimental and control trials separately in the altercentric-first and egocentric-first conditions. None of these comparisons revealed a difference (all $ps >= 0.06$).

When investigating possible carry-over costs between biases, no difference was found as a function of the presentation order. For pure egocentric bias measured by the Sandbox task, a Mann–Whitney test revealed no difference between egocentric-first vs. altercentric-first conditions, $U = 320.00$, $Z = -0.552$, $p = 0.581$. Mouse-tracking measures revealed similar null results, $U = 278.00$, $Z = -1.299$, $p = 0.194$. As to pure altercentric bias, neither the Sandbox nor the mouse-tracking measure revealed any difference between egocentric-first versus altercentric-first conditions; $U = 278.00$, $Z = -1.299$, $p = 0.194$ and $U = 323.00$, $Z = -498$, $p = 0.618$, respectively.

Overall, the present study thus failed to find any evidence for egocentric or altercentric biases even in a within-subjects block design. There was also no evidence for an effect of the order of blocks presented.

5. General discussion

The present study developed a new task to test egocentric and altercentric biases, as potential indicators of explicit and implicit ToM, within the same task format. To this end, building on existing continuous explicit False Belief (Sandbox) tasks, closely matched altercentric and egocentric versions of an online task were devised. Across three studies and two different measures, we found no evidence for any bias. More formal investigation of the null results via Bayes Factors analyses yielded mostly anecdotal to moderate evidence for the null hypotheses across all studies and conditions (with minor exceptions).[10] Even though the experimental and control trials differed from each other in the collapsed analyses in Study 1, the deviations were not in the expected direction and thus do not reveal a true bias of interest. In addition, there was no evidence for cross-cultural differences (Study 1) or the effect of order of task versions administered (Study 3). In the following,

10 Following the suggestion of one anonymous reviewer, we also conducted separate one-sample Bayesian t-tests for each comparison we did in this study. In these Bayesian t-tests, we investigated if the data supported the null hypothesis (i.e., no difference exists between experimental and control trials). Following Dienes (2014), we accepted the BF_{10} value of 0.33 or smaller as a benchmark of a null result of sufficient sensitivity. BF_{10} values of 0.33 or below suggest that the data are at least three times as likely under the null hypothesis than under the alternative. In all one-sample Bayesian t-tests analyses, we used the difference between the experimental and control trials. This strategy was adopted from Samuel et al. (2018b). Almost all of these analyses revealed anecdotal to moderate evidence for null hypothesis, with Bayes factors ranging between 0.85 and 0.12 (indicating that the data were 1.12 to 8.68 times more likely under the null hypothesis). Only two analyses (Study 1 Turkish participants and Study 2 English-speaking participants, egocentric bias condition, when all answers were included in the compared data) provided anecdotal evidence for a difference between experimental and control trials, with a Bayes Factor of 2.53 in both cases.

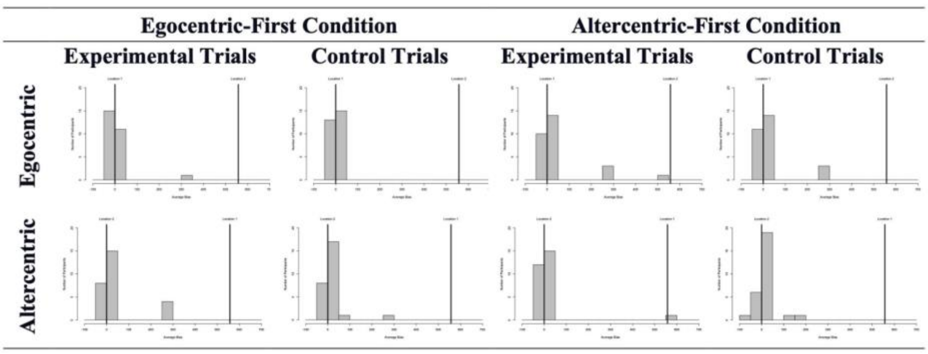


FIGURE 7
Average object location deviations in experimental and control trials across conditions and participants in Study 3.

we discuss whether this absence of evidence may constitute evidence of absence or merely false negatives.

5.1. Absence of evidence or evidence of absence for egocentric bias?

So far, original studies of the Sandbox task have repeatedly revealed significant egocentric interference effects for both children and adults (e.g., Bernstein et al., 2011; Begeer et al., 2012; Sommerville et al., 2013; Coburn et al., 2015; Mahy et al., 2017). These positive findings have been challenged by more recent replication attempts (Samuel et al., 2018a,b), where the egocentric interference effects were either absent or may have occurred due to a general difficulty with reasoning about false representations rather than false beliefs. As an example for the latter, Samuel et al. (2018a) have found equivalent levels of egocentric bias when participants were asked to indicate where a false film would depict an object as when they were asked about a protagonist's false belief regarding the object's location. The results of the current study add to the unsuccessful replication attempts and null results. It should be noted that the current study constitutes a conceptual, rather than a direct, replication attempt. Following Machery (2020), we do not argue that one form of replication is more valuable than the other. We simply emphasize that the current study was different than the original studies in terms of the task format and visual materials (starting from Study 2); and it aimed to extend the original studies to various samples by using additional measures.

But why do some studies succeed in finding evidence for egocentric biases whereas others do not? Are there any deep and systematic differences that can explain this pattern of positive versus null findings? One such potential difference may lie in the format of the studies: These differences between in-person versus online tasks could occur due to various reasons such as video-deficit effect, which has been shown to influence children's performance on FB tasks (e.g., Reiß et al., 2019) or decreased attention and motivation during online testing (see for their possible hindering effects in memory tasks, Finley and Penningroth, 2015). The two pilot studies we have conducted speak against these possibilities and extend the null results to an in-person (paper-pencil) version of the Sandbox task (see Supplementary Documents). However, those pilot studies were not direct and systematic comparisons of the online versions we used, therefore they should be approached with caution. And these possibilities should be systematically tested in the future studies where the live versions are directly compared with the online version of the task.

First systematic comparisons of live vs. online studies have recently been conducted in socio-cognitive developmental research with children, with somewhat mixed findings. For example, Schidelko et al. (2021) found no difference between lab versus online versions of standard False Beliefs tasks, while Sheskin and Keil (2018) found considerable differences between the two versions (with much poorer performance in online FB tasks). This kind of systematic comparisons should be extended to adult samples and measures such as the Sandbox task before we can conclusively interpret null findings like the present one.

Another open question about the Sandbox task is whether it is subject to domain-general reasoning strategies that can explain the biases without any reference to perspectives. For example, it is possible that in the existing egocentric bias version of this task, participants are biased toward the second location merely because they are drawn to the presence of an alternative location, but not because they are influenced by their own perspective. This would then cause participants to be biased to the second location equally in the experimental and control trials; hence, no difference should then be detectable between these trials. These concerns are not relevant for the present work as participants either showed negative biases or binary response patterns in the current studies (i.e., they were not biased toward the second location). However, future studies using the Sandbox task should take preventive precautions for this kind of alternative explanations. For example, adding a nonmental control condition (e.g., objects are moved by the wind rather than agents) would reveal if participants are biased to the incorrect locations just because these locations exist (i.e., participants would be drawn to the second location in both mental and nonmental conditions) or because they are biased by their own perspectives (i.e., participants would be drawn to the second location only in the mental condition).

5.2. Absence of evidence or evidence of absence for altercentric bias?

When it comes to the altercentric interference effects, the null results in our studies are even more difficult to interpret. The adaptation of the implicit altercentric bias version of the Sandbox task used in the current study is completely new and exploratory; hence, there is no existing body of positive or null findings to which we can compare the present results.

In theory, different factors could be at play in terms of the existing null results. One set of factors that could make a difference is the superficial methodological factors such as the online format, boring and easy tasks, and technical limitations. More specifically, the unmoderated online format of the task and not-so-engaging materials might have caused inattention to task, leaving the differences between experimental and control trials undetected by the participants. It is also possible that participants found our altercentric bias version of the Sandbox task extremely easy, leading to ceiling effects in both experimental and control trials and making them indistinguishable from each other. Finally, technical limitations rendered the mouse tracking measures not-so-spontaneous in our studies. To ensure that all participants started moving their mouse cursors from the same point, we asked them to click on a "record button" before moving their mouse cursors. It is possible that this requirement interfered with more spontaneous and automatic altercentric bias effects and gave participants more time to reflect on their answers, leading to null results revealed by mouse movements.

There are also more substantial factors that could result in null altercentric interference effects. For example, it is possible that altercentric biases can be reliably found only in some domains for some types of measures but not in others. This bias is hypothesized to be automatic and spontaneous processing of others' perspectives (e.g., Southgate, 2020). Therefore, more spontaneous temporal measures such as response times integrated into simpler tasks such as Level-1 visual perspective taking could reveal this bias better than the fine-grained spatial deviations in the contents of judgments about an object's location. The latter would require more extended processing due to the preceding scenarios and its answer format whereas the former is more suitable for quick, automatic, spontaneous judgments. These factors should be explored in future research more systematically before the current task format is given up as a potential measure to tap altercentric bias.

5.3. Future directions and conclusion

Overall, the current set of studies thus failed to provide evidence for egocentric or altercentric biases in a novel combined task format. These null findings remain difficult to interpret and raise more questions for future research than they answer. In particular, it remains unclear whether the absence of evidence for egocentric and altercentric biases reflects the fact that these biases may be less robust than previously assumed. Alternatively, it could be that the biases are robust, but the present tasks (due to their online format or the specific content) are not suitable for tapping them. Systematic future investigations are required to answer these questions.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

Author contributions

FH, MP, UL, and HR contributed to conception and design of the study. FH conducted the studies, performed the statistical analysis, and wrote the first draft of the manuscript. MP, UL, and HR gave critical review and commentary on the drafts of the manuscript. MP and HR supervised the planning and execution process. HR provided resources for the data collection. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by Deutsche Forschungsgemeinschaft [Crossing the Borders Project, (LI 1989/3–2)] and by Research Training Group 2070 (Grant/Award Number: 254142454/GRK 2070). We also acknowledge support by the Open Access Publication Funds of the Göttingen University.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1142302/full#supplementary-material

References

Apperly, I. A., and Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychol. Rev.* 116, 953–970. doi: 10.1037/a0016923

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

Begeer, S., Bernstein, D. M., van Wijhe, J., Scheeren, A. M., and Koot, H. M. (2012). A continuous false belief task reveals egocentric biases in children and adolescents with autism spectrum disorders. *Autism* 16, 357–366. doi: 10.1177/1362361311434545

Bernstein, D. M., Thornton, W. L., and Sommerville, J. A. (2011). Theory of mind through the ages: older and middle-aged adults exhibit more errors than do younger adults on a continuous false belief task. *Exp. Aging Res.* 37, 481–502. doi: 10.1080/0361073X.2011.619466

Burnside, K., Ruel, A., Azar, N., and Poulin-Dubois, D. (2018). Implicit false belief across the lifespan: non-replication of an anticipatory looking task. *Cogn. Dev.* 46, 4–11. doi: 10.1016/j.cogdev.2017.08.006

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

Coburn, P. I., Bernstein, D. M., and Begeer, S. (2015). A new paper and pencil task reveals adult false belief reasoning bias. *Psychol. Res.* 79, 739–749. doi: 10.1007/s00426-014-0606-0

Cole, G. G., Atkinson, M., Le, A. T. D., and Smith, D. T. (2016). Do humans spontaneously take the perspective of others? *Acta Psychol.* 164, 165–168. doi: 10.1016/j.actpsy.2016.01.007

Conway, J. R., Lee, D., Ojaghi, M., Catmur, C., and Bird, G. (2017). Submentalizing or mentalizing in a level 1 perspective-taking task: a cloak and goggles test. *J. Exp. Psychol. Hum. Percept. Perform.* 43, 454–465. doi: 10.1037/xhp0000319

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Front. Psychol.* 5:781. doi: 10.3389/fpsyg.2014.00781

Dörrenberg, S., Rakoczy, H., and Liszkowski, U. (2018). How (not) to measure infant theory of mind: testing the replicability and validity of four non-verbal measures. *Cogn. Dev.* 46, 12–30. doi: 10.1016/j.cogdev.2018.01.001

Faul, F., Erdfelder, E., Buchner, A., and Lang, A.-G. (2009). Statistical power analyses using G*power 3.1: tests for correlation and regression analyses. *Behav. Res. Methods* 41, 1149–1160. doi: 10.3758/BRM.41.4.1149

Finley, A. J., and Penningroth, S. L. (2015). "Online versus in-lab: pros and cons of an online prospective memory experiment" in *Advances in psychology research*. ed. A. M. Columbus (New York, NY: Nova), 35–161.

Furlanetto, T., Becchio, C., Samson, D., and Apperly, I. (2016). Altercentric interference in level 1 visual perspective taking reflects the ascription of mental states, not submentalizing. *J. Exp. Psychol. Hum. Percept. Perform.* 42, 158–163. doi: 10.1037/xhp0000138

Kampis, D., Karman, P., Csibra, G., Southgate, V., and Hernik, M. (2021). A two-lab direct replication attempt of Southgate, Senju and Csibra (2007). *R. Soc. Open Sci.* 8:210190. doi: 10.1098/rsos.210190

Kampis, D., and Southgate, V. (2020). Altercentric cognition: how others infuence our cognitive processing. *Trends Cogn. Sci.* 24, 945–959. doi: 10.1016/j.tics.2020.09.003

Knudsen, B., and Liszkowski, U. (2012). 18-month-olds predict specific action mistakes through attribution of false belief, not ignorance, and intervene accordingly. *Infancy* 17, 672–691. doi: 10.1111/j.1532-7078.2011.00105.x

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

Kulke, L., Reiß, M., Krist, H., and Rakoczy, H. (2018a). How robust are anticipatory looking measures of theory of mind? Replication attempts across the life span. *Cogn. Dev.* 46, 97–111. doi: 10.1016/j.cogdev.2017.09.001

Kulke, L., von Duhn, B., Schneider, D., and Rakoczy, H. (2018b). Is implicit theory of mind a real and robust phenomenon? Results from a systematic replication study. *Psychol. Sci.* 29, 888–900. doi: 10.1177/0956797617747090

Leslie, A. M. (2005). Developmental parallels in understanding minds and bodies. *Trends Cogn. Sci.* 9, 459–462. doi: 10.1016/j.tics.2005.08.006

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

Low, J., and Watts, J. (2013). Attributing false beliefs about object identity reveals a signature blind spot in humans' efficient mind-reading system. *Psychol. Sci.* 24, 305–311. doi: 10.1177/0956797612451469

Machery, E. (2020). What is a replication? *Philos. Sci.* 87, 545–567. doi: 10.1086/709701

Mahy, C. E., Bernstein, D. M., Gerrard, L. D., and Atance, C. M. (2017). Testing the validity of a continuous false belief task in 3-to 7-year-old children. *J. Exp. Child Psychol.* 160, 50–66. doi: 10.1016/j.jecp.2017.03.010

Markus, H. R., and Kitayama, S. (1991). Culture and the self: implications for cognition, emotion, and motivation. *Psychol. Rev.* 98, 224–253. doi: 10.1037/0033-295X.98.2.224

Michael, J., Wolf, T., Letesson, C., Butterfill, S., Skewes, J., and Hohwy, J. (2018). Seeing it both ways: using a double-cuing task to investigate the role of spatial cuing in Level-1 visual perspective-taking. *J. Exp. Psychol. Hum. Percept. Perform.* 44, 693–702. doi: 10.1037/xhp0000486

O'Grady, C. J., Scott-Phillips, T., Lavelle, S., and Smith, K. (2017). "The dot perspective task revisited: evidence for directional effects" in *Cogsci 2017: Proceedings of the 39th annual meeting of the cognitive science society*, Eds. G. Gunzelmann, A. Howes, T. Tenbrink and E. Davelaar 2821–2826.

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

Perner, J., Leekam, S. R., and Wimmer, H. (1987). Three-year-olds' difficulty with false belief: the case for a conceptual deficit. *Br. J. Dev. Psychol.* 5, 125–137. doi: 10.1111/j.2044-835x.1987.tb01048.x

Perner, J., and Roessler, J. (2012). From infants' to children's appreciation of belief. *Trends Cogn. Sci.* 16, 519–525. doi: 10.1016/j.tics.2012.08.004

Poulin-Dubois, D., Azar, N., Elkaim, B., and Burnside, K. (2020). Testing the stability of theory of mind: A longitudinal approach. *Plos One*, 15:e0241721. doi: 10.1371/journal.pone.0241721

Poulin-Dubois, D., Goldman, E. J., Meltzer, A., and Psaradellis, E. (2023). Discontinuity from implicit to explicit theory of mind from infancy to preschool age. *Cogn. Dev.* 65:101273. doi: 10.1016/j.cogdev.2022.101273

Poulin-Dubois, D., and Yott, J. (2018). Probing the depth of infants' theory of mind: disunity in performance across paradigms. *Dev. Sci.* 21:e12600. doi: 10.1111/desc.12600

Powell, L. J., Hobbs, K., Bardis, A., Carey, S., and Saxe, R. (2018). Replications of implicit theory of mind tasks with varying representational demands. *Cogn. Dev.* 46, 40–50. doi: 10.1016/j.cogdev.2017.10.004

Priewasser, B., Rafetseder, E., Gargitter, C., and Perner, J. (2018). Helping as an early indicator of a theory of mind: mentalism or teleology? *Cogn. Dev.* 46, 69–78. doi: 10.1016/j.cogdev.2017.08.002

Rakoczy, H. (2022). Foundations of theory of mind and its development in early childhood. *Nat. Rev. Psychol.* 1, 223–235. doi: 10.1038/s44159-022-00037-z

Reiß, M., Krüger, M., and Krist, H. (2019). Theory of mind and the video deficit effect: video presentation impairs children's encoding and understanding of false belief. *Media Psychol.* 22, 23–38. doi: 10.1080/15213269.2017.1412321

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., and Bodley Scott, S. E. (2010). Seeing it their way: evidence for rapid and involuntary computation of what other people see. *J. Exp. Psychol. Hum. Percept. Perform.* 36, 1255–1266. doi: 10.1037/a0018729

Samuel, S., Legg, E. W., Lurz, R., and Clayton, N. S. (2018a). Egocentric bias across mental and non-mental representations in the sandbox task. *Q. J. Exp. Psychol. Hum. Exp. Psychol.* 71, 2395–2410. doi: 10.1177/1747021817742367

Samuel, S., Legg, E. W., Lurz, R., and Clayton, N. S. (2018b). The unreliability of egocentric bias across self-other and memory-belief distinctions in the sandbox task. *R. Soc. Open Sci.* 5:181355. doi: 10.1098/rsos.181355

Santiesteban, I., Catmur, C., Hopkins, S. C., Bird, G., and Heyes, C. (2014). Avatars and arrows: implicit mentalizing or domain-general processing? Journal of experimental psychology. *Hum. Perception Performance* 40, 929–937. doi: 10.1037/a0035175

Schidelko, L. P., Schünemann, B., Rakoczy, H., and Proft, M. (2021). Online testing yields the same results as lab testing: a validation study with the false belief task. *Front. Psychol.* 12:4573. doi: 10.3389/fpsyg.2021.703238

Schneider, D., Bayliss, A. P., Becker, S. I., and Dux, P. E. (2012). Eye movements reveal sustained implicit processing of others' mental states. *J. Exp. Psychol. Gen.* 141, 433–438. doi: 10.1037/a0025458

Schuwerk, T., Priewasser, B., Sodian, B., and Perner, J. (2018). The robustness and generalizability of findings on spontaneous false belief sensitivity: a replication attempt. *R. Soc. Open Sci.* 5:172273. doi: 10.1098/rsos.172273

Scott, R. M., and Baillargeon, R. (2017). Early false-belief understanding. *Trends Cogn. Sci.* 21, 237–249. doi: 10.1016/j.tics.2017.01.012

Senju, A., Southgate, V., Miura, Y., Matsui, T., Hasegawa, T., Tojo, Y., et al. (2010). Absence of spontaneous action anticipation by false belief attribution in children with autism spectrum disorder. *Dev. Psychopathol.* 22, 353–360. doi: 10.1017/S0954579410000106

Sheskin, M., and Keil, F. (2018). *TheChildLab.com a video chat platform for developmental Research PsyArXiv*. 30:2018. doi: 10.31234/osf.io/rn7w5.

Singelis, T. M. (1994). The measurement of independent and interdependent self-construals. *Personal. Soc. Psychol. Bull.* 20, 580–591. doi: 10.1177/0146167294205014

Sodian, B. (2023). Reply to Poulin-Dubois et al. (2023): Replication problems concerning both implicit and explicit false belief reasoning greatly reduced the chance of finding longitudinal correlations. *Cognitive Development* 65:101294. doi: 10.1016/j.cogdev.2022.101294

Sodian, B., Kristen-Antonow, S., and Kloo, D. (2020). How does children's theory of mind become explicit? A review of longitudinal findings. *Child Dev. Perspect.* 14, 171–177. doi: 10.1111/cdep.12381

Sodian, B., Licata, M., Kristen-Antonow, S., Paulus, M., Killen, M., and Woodward, A. (2016). Understanding of goals, beliefs, and desires predicts morally relevant theory of mind: a longitudinal investigation. *Child Dev.* 87, 1221–1232. doi: 10.1111/cdev.12533

Sommerville, J. A., Bernstein, D. M., and Meltzoff, A. N. (2013). Measuring beliefs in centimeters: private knowledge biases preschoolers' and adults' representation of others' beliefs. *Child Dev.* 84, 1846–1854. doi: 10.1111/cdev.12110

Southgate, V. (2020). Are infants altercentric? The other and the self in early social cognition. *Psychol. Rev.* 127, 505–523. doi: 10.1037/rev0000182

Southgate, V., Senju, A., and Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychol. Sci.* 18, 587–592. doi: 10.1111/j.1467-9280.2007.01944.x

Speiger, M. L., Rothmaler, K., Rakoczy, H., Kizilay, F. N. H., Liszkowski, U., and Grosse Wiesmann, C. (2022). *The Sandbox Task-Measuring implicit and explicit Theory of Mind* [Conference presentation abstract]. Budapest CEU Conference on Cognitive Development. Virtual. Available at: https://cdc.ceu.edu/sites/cdc.ceu.edu/files/attachment/basicpage/260/bcccdfinalprogram2022.pdf.

Van der Wel, R. P., Sebanz, N., and Knoblich, G. (2014). Do people automatically track others' beliefs? Evidence from a continuous measure. *Cognition* 130:128. doi: 10.1016/j.cognition.2013.10.004

Wang, J. J., Tseng, P., Juan, C. H., Frisson, S., and Apperly, I. A. (2019). Perspective-taking across cultures: shared biases in Taiwanese and British adults. *R. Soc. Open Sci.* 6:190540. doi: 10.1098/rsos.190540

Wellman, H. M., Cross, D., and Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Dev.* 72, 655–684. doi: 10.1111/1467-8624.00304

Wimmer, H., and Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13, 103–128. doi: 10.1016/0010-0277(83)90004-5

Wu, S., and Keysar, B. (2007). The effect of culture on perspective taking. *Psychol. Sci.* 18, 600–606. doi: 10.1111/j.1467-9280.2007.01946.x

Yott, J., and Poulin-Dubois, D. (2016). Are infants' theory-of-mind abilities well integrated? Implicit understanding of intentions, desires, and beliefs. *J. Cogn. Dev.* 17, 683–698. doi: 10.1080/15248372.2015.1086771

SUPPLEMENTARY DOCUMENTS

Pilot Studies (Paper-Pencil Version of the Sandbox Task)

These studies aimed to investigate if the previously found egocentric interference effects could be replicated with German- (and Turkish) speaking adults using the paper-pencil version of the Sandbox task (e.g., Coburn et al., 2015). The first pilot study utilized simple materials displaying just a sandbox drawing on the paper and the location markers (as done by the previous studies, e.g., Coburn et al., 2015; Mahy et al., 2017). The second pilot study was very similar to the first one, except it used more engaging materials which displayed elaborate scenes. We conducted this second pilot study to control for our criticism of Study 1 (i.e., "We suspect that the not-so-engaging task materials might have caused our remaining participants to fail to pay enough attention to the task, which could have made the task less reliable.") The other difference between the first and the second pilot study was the task language (and participants' native language). As we did not find any difference between German- and Turkish-speaking participants in Study 1 and since Turkish-speaking participants were easier to reach at the time of the second pilot study, Pilot Study 2 tested Turkish-speaking participants.

Pilot Study 1

Method

Participants

Participants were recruited through personal communication channels and print ads. All participants were tested in person in a quiet room without any distractions. We used G*POWER (Faul et al., 2009) to conduct a power analysis and determine the sample size. We aimed to obtain .95 power to detect a medium effect size of .50 at the standard .05 alpha error probability with a more conservative two-tailed paired-samples t-test. The analysis revealed a required sample size of 54 participants. We tested 55 participants to achieve a sample size of 54 (one participant was excluded from the dataset as she was not a native German speaker and showed poor understanding of the task). Thus, the final sample consisted of 54 German-speaking adults: (24 females, 20 males, 10 unknown, M_{age} = 26.8, age range: 18 to 65). All participants provided written consent for the study and received a candy bar after the test session.

Appendix A: Haskaraca, Proft, Liszkowski, & Rakoczy (2023)

Materials (The Sandbox Task)

The scenarios used in our study were based on those used by Coburn et al. (2015) and Mahy et al. (2017). They always followed the same storyline: Agent A hides an object in Location 1, but then the object is transferred to Location 2 by Agent B in the absence of Agent A. The stories were always presented with accompanying images. The images (29.5 x 21 cm) displayed a rectangular container (21.9 x 2.7 cm) positioned in the middle of the image and text above the container. The crosses (0.5 x 0.5 cm) on the container indicated a hidden object's initial and final locations. These locations were always 13.4 cm apart, but their relative position changed across trials to prevent participants from learning the locations. In all of our studies, the direction of relocation was counterbalanced: in half of the trials, the object was transferred from left to right, and in the other half, the transfer was from right to left. The objects always crossed the midline of the sandbox during the transposition.

Once the scenarios were presented, participants worked on a word-search puzzle for 20 seconds. Puzzles prevented using perceptual cues to answer the question and were created by inserting family-related words into a 21 x 21 word-search puzzle using a puzzle maker website (https://puzzlemaker.discoveryeducation.com).

After the distraction task, participants were asked either where Agent A, who had a false belief about the object's location, would look for the object upon return (experimental trials; "Where will X look for the object?") or where s/he hid the object before leaving the scene (control trials; "Where did X hide the object?"). In both of these trials, the correct answer was around Location 1. Participants were expected to deviate in the direction of Location 2 in the experimental trials as they knew that the object was actually at Location 2, and this knowledge was expected to interfere with their judgments of others' perspectives and behaviors.

Design & Procedure

All participants completed the paper-pencil version of the Sandbox task, which aimed to tap egocentric biases. Each participant completed four experimental and four control trials presented in blocks (the order of the blocks counterbalanced) and one filler trial in between.

After consenting to the study, participants were seated at a table along with the experimenter. The experimenter was responsible for reading the stories and question prompts out loud, and she moderated the test session (i.e., proceeded across trials, made sure that participants saw only one location at a time, and managed the timing for distraction task).

After each scenario, there was a 20-seconds of puzzle solving, and then participants were presented with the question and asked to mark their answers on an empty sandbox drawing. The study took approximately 10 minutes.

Bias Calculation & Analyses

Biases were inferred from the object location measure: the horizontal distance (in cm) between the correct location (i.e., L1) and the participant's response. If the participants' responses were biased toward the wrong location (i.e., between the right and wrong answer, toward to middle of the paper), they received a positive object location value. The responses biased away from the wrong location (i.e., toward the edge of the Sandbox/paper rather than the middle) received a negative object location value. Once the object location measure was computed for each trial, we calculated the average object location measure in experimental and control trials for each participant. The averages were calculated in two ways: a) all responses were included in the averages (as done in the original Sandbox task studies), and b) the completely wrong answers (i.e., responses that were closer to the incorrect location than the correct location) were excluded from the averages. The latter method aimed to exclude the trials to which participants did not pay enough attention. We argue that adults are expected to have full-fledged perspective-taking abilities; therefore, completely wrong answers would reflect participants' failures of attention and could be excluded from the data for explorative purposes (e.g., does a bias exist when only the attended trials are considered?). The average scores were then used to deduce biases: if the average deviation in the experimental trials is bigger than the control trials, this indicates bias. As a result, different within-subject comparisons (paired-sample t-tests) were conducted with and without wrong answers to see if a bias exists in different conditions and groups. We used non-parametric tests (e.g., matched-pair Wilcoxon signed-rank Test) when the response data were not normally distributed. Even when non-parametric tests were more appropriate due to the non-normal distribution of the data, we also ran parametric tests as we had initially expected a continuous distribution of answers. The pattern of results and significance remained the same across all analyses.

Results

We first compared the average biases in control versus experimental trials without the wrong answers. No difference was detected between experimental (M=-1.40 SD=.89) and control (M=-1.47 SD=.83) trials, $t(53)$=.906, p=.369, d=.12. Then we included the wrong answers in the data and repeated the comparisons. Again, no difference was revealed between experimental (*M*=.49 *SD*=2.82, *Mdn*=-.59) and control (*M*=-.27 *SD*=2.24, *Mdn*=-1.11) trials, $Z = -1.520$, $p = .129$.

Pilot Study 2

Method

Participants

Participants were recruited through personal communication channels and e-mail announcements. All participants were tested in person in a quiet room without any distractions. The sample size was based on the same rationale as the first pilot study; 54 participants were needed. We tested 56 participants to achieve a sample size of 54 (one participant was excluded from the dataset as she was not a native Turkish speaker, and one participant could not answer all trials as the test session was interrupted). The final sample consisted of 54 Turkish-speaking adults: (23 females, 31 males, $M_{age} = 25.6$, age range: 21 to 67). All participants provided written consent for the study and received a candy bar after the test session.

Materials (The Sandbox Task)

The materials used in the second pilot study were different from the first one in terms of the visual materials. Instead of showing dull materials with only a box drawing and the location markers, more engaging and elaborative materials have been displayed to the participants in the current study. For example, story-compatible background images were added; the object locations were not marked with Xs, but drawings of objects were shown; agents were displayed on the materials; and colorful materials were used. The images (29.5 x 21 cm) used in this study displayed a continuous rectangular area (width: 28.8 cm) positioned in the middle of the image and text at the bottom of the image. A different object was displayed for each trial; however, their surface area was kept constant across trials (2.5 cm^2). Objects' initial and final locations were always 20.3 cm apart. Besides these differences, the materials used in the current study were the same as the first pilot study.

Design & Procedure, and Bias Calculation & Analyses: same as Pilot Study 1.

Results

We first compared the average biases in control versus experimental trials without the wrong answers. No difference was detected between experimental (M=1.19 SD=.74) and control (M=1.2 SD=.77) trials, $t(53)$=-.070, p=.945, d=-.009. Then we included the wrong answers in the data and repeated the comparisons. Again, no difference was revealed between experimental (M=1.56 SD=1.81, Mdn=1.23) and control (M=1.45 SD=1.29, Mdn=1.27) trials, Z = -.065, p = .949.

Appendix A: Haskaraca, Proft, Liszkowski, & Rakoczy (2023)

Self-Construal Scale (Singelis, 1994)

Example Items:

Item	1	2	3	4	5	6	7
I enjoy being unique and different from others in many respects.	Strongly Disagree	Disagree	Somewhat Disagree	Don't Agree or Disagree	Somewhat Agree	Agree	Strongly Agree
I can talk openly with a person who I meet for the first time, even when this person is much older than I am.	Strongly Disagree	Disagree	Somewhat Disagree	Don't Agree or Disagree	Somewhat Agree	Agree	Strongly Agree
Even when I strongly disagree with group members, I avoid an argument.	Strongly Disagree	Disagree	Somewhat Disagree	Don't Agree or Disagree	Somewhat Agree	Agree	Strongly Agree
I have respect for the authority figures with whom I interact.	Strongly Disagree	Disagree	Somewhat Disagree	Don't Agree or Disagree	Somewhat Agree	Agree	Strongly Agree

Additional Analyses

Response Time Measure

Following a reviewer's suggestion, we conducted some additional analyses with the response time data. Using this measure, we identified and excluded the too quick and too slow outliers and repeated the within-subject analyses reported in the paper. Furthermore, we also used this data to compare the control and experimental trials in terms of response times.

Calculation of Response Time Measure

Response time measure was not one of the measures in which we were interested from the beginning. Therefore, we do not have direct measures for response times in our studies. In Study 1, we only have the information about the total duration of the study, which is not very informative for the analyses in this section. Thus, Study 1 will not be included in the investigations done with response times.

In Studies 2 & 3, we did not measure the response time directly. As an indirect measure, we can use the time series data from the mouse-tracking measure. It is, however, important to point out that this data would give us only the time that was taken to move the mouse around but not the overall time spent on a trial. Therefore, its validity as a response time measure is unclear.

Within-subject comparisons without the response time outliers

The response time data in Study 2 revealed four very slow outliers in the altercentric bias condition (one from the English-speaking sample and three from the German-speaking sample) and three very slow outliers in the egocentric bias condition (all from the English-speaking sample). We removed these outliers and repeated the within-subject comparisons with the object location and mouse-tracking measures. The overall result pattern did not change: neither an altercentric nor an egocentric bias was found in any group. More specifically, no difference between experimental and control trials was observed for English-speaking participants with object-location measure, neither in egocentric (Z = -1.723, p = .085) nor altercentric (Z = -1.069, p = .285) bias condition. No difference was revealed by mouse-tracking measures, neither in egocentric (Z = -1.425, p = .154) nor altercentric (Z = -.628, p = .530) bias condition.

Similarly, no difference between experimental and control trials was observed for German-speaking participants in the altercentric bias condition, neither with object-location

(Z = -.467, p = .641) nor with mouse-tracking (Z = -.537, p = .591) measure. As the response time data did not reveal any outliers for German-speakers in the egocentric bias condition, the analyses were not repeated for egocentric bias in this group.

Study 3 revealed only two slow outliers. When those were eliminated from the data, no change in the result pattern was observed: no difference between experimental and control trials was observed with object location measures, neither for egocentric (Z = -.403, p = .687) nor altercentric (Z = -1.336, p = .181) bias. Also, no difference between experimental and control trials was observed with mouse-tracking measures, neither for egocentric (Z = -1.678, p = .093) nor altercentric (Z = -.778, p = .437) bias.

Mixed-Models

Following the suggestion of one reviewer, we used a mixed-effects model approach to see if the individual bias scores are influenced by the trial type (i.e., experimental and control). More specifically, we conducted mixed-effect models for individual bias scores with the trial type as the fixed effect and the participant and item as the random effects.

In Study 1, the trial type had no effect on the bias score for German (egocentric bias: F=.87, p = .81, 95% CI [-81.34, 98.96]; altercentric bias: F=.46, p = .524, 95% CI [-91.24, 160.98]) and Turkish participants (egocentric bias: F=4.49, p = .08, 95% CI [-163.33, 11.72]; altercentric bias: F=.37, p = .565, 95% CI [-169.87, 102.11]).

In Study 2, we again found no effect of trial type on bias scores as measured by the object location measure, neither for German- (egocentric bias: F=.49, p = .513, 95% CI [-20.65, 11.64]; altercentric bias: F=.04, p = .852, 95% CI [-34.68, 29.53]) nor English-speaking (egocentric bias: F=1.60, p = .209, 95% CI [-12.71, 57.43]; altercentric bias: F=.16, p = .704, 95% CI [-31.04, 42.26]) participants. Mouse-tracking measures revealed similar results for both German- (egocentric bias: F=.07, p = .792, 95% CI [-5.99, 4.57]; altercentric bias: F=.74, p = .390, 95% CI [-6.41, 2.51]) and English-speaking (egocentric bias: F=.33, p = .568, 95% CI [-3.69, 6.70]; altercentric bias: F=.00, p = .990, 95% CI [-6.90, 6.83]) participants.

Study 3 provided similar line of results. Namely, no effect of trial type was observed regardless of the measure, i.e., object location (egocentric bias: F=.00, p = .955, 95% CI [-17.215, 17.98]; altercentric bias: F=1.73, p = .190, 95% CI [-2.53, 12.68]) or mouse-tracking (egocentric bias: F=1.06, p = .374, 95% CI [-10.13, 5.04]; altercentric bias: F=1.33, p = .321, 95% CI [-5.57, 12.75]).

Appendix B: Haskaraca, Proft, Liszkowski, & Rakoczy (2023)

Haskaraca, F.N., Proft, M., Liszkowski, U., & Rakoczy, H. (2023). *Testing Egocentric and Altercentric Biases in a Referential Communication Task* [Unpublished manuscript]. Department of Developmental Psychology. University of Göttingen.

Appendix B: Haskaraca, Proft, Liszkowski, & Rakoczy (2023)

Testing Egocentric and Altercentric Biases in a Referential Communication Task

Appendix B: Haskaraca, Proft, Liszkowski, & Rakoczy (2023)

Abstract

Theory of Mind (ToM), the ability to attribute mental states to oneself and others, has been long assumed to develop around age four. Research from a lifespan perspective has revealed that ToM remains somewhat error-prone even in adulthood, with individuals often displaying egocentric interferences (i.e., interferences from their own perspective) when ascribing perspectives to others. More recent work on ToM also showed that this ability is potentially preceded by implicit, automatic forms of representing others' perspectives. One such form is altercentric bias, i.e., interferences from others' differing perspectives on our own judgments. The current study aims to implement a task format that can analogously measure egocentric and altercentric biases. To this end, in two online studies with adults, we used the so-called Director Task (tapping egocentric interference effects) and a newly adapted variation (tapping altercentric interference effects). We aimed to replicate egocentric bias effects observed in previous research using this task and examine whether the modified Director Task can measure altercentric interference effects. The results revealed robust egocentric biases across two studies. However, the findings from the altercentric bias version provided mixed results for the two measures used in the study (accuracy and reaction times). Also, the non-social condition of the altercentric bias version revealed an identical result pattern as the social version, indicating that the interferences found in this version were not due to other person's differing perspectives. The results are discussed in terms of their reliability (for the egocentric bias version) and validity (for the altercentric bias version).

Keywords: Theory of Mind, altercentric bias, egocentric bias, Director Task

Testing Egocentric and Altercentric Biases in a Referential Communication Task

Theory of Mind (ToM) is a fundamental aspect of our social nature that allows us to attribute mental states to ourselves and others. At the core of this capacity lies the ability to meta-represent, i.e., to represent one's own and others' representational states. This capacity is most clearly indicated by the ability to represent misrepresentations, i.e., representing how others represent the world and behave accordingly, even when these representations are incorrect. In light of this idea, False Belief (FB) tasks have become the litmus tests of the meta-representational ToM abilities. The so-called change-of-location task (e.g., Wimmer & Perner, 1983) is one of the most frequently used FB tasks. In this task, participants are presented with a story in which an object is transferred from one location to another, either in the absence (false belief condition) or the presence (true belief condition) of a protagonist. Then participants are asked where the protagonist will search for the object. Years of research have shown that children come to solve this task (they start answering according to the protagonist's belief rather than reality) around four years of age (e.g., Wellman et al., 2001). Similar performance shifts have also been observed for the other FB tasks (e.g., the unexpected-content task; Perner et al., 1987) and superficially different ToM tasks (e.g., the appearance-reality task; e.g., Perner & Roessler, 2012). These findings formed the base for the assumption that ToM emerges and develops in preschool years based on a cognitive capacity that allows the conceptual transition around age four -namely, the 4-year-revolution- in different ToM tasks.

ToM has been mainly studied from a developmental point of view, focusing on when and how this capacity emerges. Lately, though, a growing number of studies have also investigated what proceeds the 4-year revolution and how ToM functions over the lifespan (e.g., Henry et al., 2013). These studies have revealed that ToM remains fragile and somewhat error-prone even in adulthood (e.g., Samson et al., 2010; Sommerville et al., 2013). In particular, even adults suffer from egocentric interferences or biases: when ascribing mental states to others, they perform more poorly in situations where the other's and their own perspective diverge. In such situations, subjects are slower in their judgments, and they sometimes tend to judge that other agent sees the world more like they themselves than is actually the case.

More recent research in ToM literature has also investigated what precedes the fully-fledged ToM abilities measured by explicit FB tasks. These newer studies revealed that the 4-year-revolution is ontogenetically (and perhaps also phylogenetically) anteceded by implicit, automatic forms of representing others' perspectives, i.e., implicit ToM. These studies capitalized on implicit FB tasks instead of explicit verbal measures. In those tasks, FB understanding has been inferred from participants' spontaneous looking behavior or behavioral interactions, such as Violation of expectation (VoE) or anticipatory looking (AL) paradigms, and interactive helping tasks. In VoE studies, children look longer at events that violate their expectations, for instance, when an agent acts inconsistently with her (false) belief (e.g., Onishi & Baillargeon, 2005). In AL paradigms, children's looking patterns indicate that they anticipate the agents to act consistently with their (false) beliefs (e.g., Southgate et al., 2007). In interactive tasks, children take others' (false) beliefs into account during their communicative interactions (e.g., Buttelmann et al., 2009). Collectively, the findings of these implicit measures suggested that children have an implicit sensitivity to others' beliefs even in the first year of life (Scott & Baillargeon, 2017).

The effects revealed by these measures are not robust, though. More recent research using the implicit ToM measures has revealed serious replication issues. More specifically, a relatively small number of positive findings were reported by only a few labs based on small sample sizes (e.g., Onishi & Baillargeon, 2005; Senju et al., 2010; Southgate et al., 2007). Moreover, many subsequent replication studies revealed mixed findings: some failed to replicate the original findings, while others found fragile effects that disappeared under more stringent test conditions (e.g., Dörrenberg et al., 2018; Poulin-Dubois & Yott, 2018; Powell et al., 2018; Priewasser et al., 2018; Schuwerk et al., 2018; Yott & Poulin-Dubois, 2016). These contradicting findings raised doubts about the reliability and validity of these measures.

One type of implicit measure is still a promising indicator of implicit perspective-taking abilities. This measure capitalizes on the so-called altercentric bias, which is somewhat a mirror image of the egocentric bias. More specifically, altercentric interference effects correspond to the interferences from others' perspectives on one's own judgments and behaviors (Kampis & Southgate, 2020; Southgate, 2020). This bias is a potential indicator of implicit perspective-taking abilities as it suggests that humans automatically and spontaneously take others' perspectives into account, even though those are irrelevant or detrimental to our own judgments. For example, individuals become slower or more error-prone in their judgments about a state of the world (e.g., how many dots are visible to them or if they could detect an object in a given context) when another agent is present but has an incongruent perspective, compared to situations where the agent has a congruent perspective (Dot-perspective task, e.g., Samson et al., 2010; Object-detection task, e.g., Kovács et al., 2010). Or, when indicating the place of an object on the screen, adults' mouse cursor trajectories detour in the direction of where an irrelevant agent mistakenly thinks the object is (Mouse-tracking task, e.g., Van der Wel et al., 2014).

Interestingly, some tasks can be designed in such a way that egocentric and altercentric types of biases can be tapped in analogous ways within the same (or closely matched) task format. For example, the dot-perspective task (Samson et al., 2010) has been one of the few measures in which egocentric and altercentric biases have been obtained using the same task. In this task, adult participants were asked to judge the number of dots presented in a scene either from the perspective of an on-screen avatar (OTHER condition) or from their own perspective (SELF condition). Each condition employed two types of trials: consistent and inconsistent. In consistent trials, all items were equally visible to the participant and the avatar, so the perspectives were consistent. In inconsistent trials, some dots that were visible to the participant were behind, therefore not visible to, the avatar. Hence, the participant's and the agent's perspectives were inconsistent. The study revealed that participants were slower and made more errors while detecting the number of dots in the inconsistent trials compared to the consistent trials in both SELF and OTHER trials. These results were interpreted as evidence for egocentric and altercentric interference effects, respectively. However, more recent studies with the dot perspective method have raised doubts about both the replicability and the validity of the task: Some studies using variations of the dot-perspective task have either found no biases (e.g., Conway et al., 2017), and others suggest that potential biases could be explained away by domain-general mechanisms (sub-mentalizing) rather than by mentalizing (e.g., Cole et al., 2016; Conway et al., 2017; O'Grady et al., 2017; Santiesteban et al., 2014).

The main aim of the current study, therefore, is to design an alternative method that can analogously and reliably tap egocentric and altercentric biases within the same task format. To this end, we capitalize on the so-called *Director Task* (e.g., Dumontheil et al., 2010; Keysar et al., 2000; Keysar et al., 2003; Samuel et al., 2016; Samuel et al., 2019). This task has been initially used in an early eye-tracking study investigating egocentric heuristics in communication (Keysar et al., 2000). Then it has also been used in social-cognition research to tap the egocentric interferences on (visual) perspective-taking abilities (e.g., Legg et al., 2017; Samuel et al., 2019). One advantage of the Director Task over the previous egocentric bias measures is that it is a more natural and interactive task, where no explicit question is needed to prompt the perspective-taking. Instead, in this task, the bias is triggered in a conversational setting, which would increase the ecological validity of the measure.

In the original version of the task, participants are instructed to select and move items on a grid shelf consisting of multiple compartments. The grid is placed between the participant and another agent, the director, who gives the instructions (typically a confederate in non-computerized versions and a computer-generated avatar in the computerized versions). Due to several occlusions in the grid, the director has a restricted view of the objects. Hence, an instruction from the director to select an object (e.g., “hand me the biggest box” or “select the biggest box”) in the grid is ambiguous if the instruction can refer to two items: one that the director can see and another one he cannot see due to an occluder. Across both real-life and computerized versions of the Director Task, participants revealed robust egocentric biases: they made more errors and were slower in the ambiguous trials compared to the unambiguous trials where the instructions of the director could describe only one object that was seen from both perspectives (e.g., Apperly et al., 2010; Dumontheil et al., 2010; Dumontheil et al., 2012; Keysar et al., 2000; Keysar et al., 2003; Legg et al., 2017; Samuel et al., 2019).

The first aim of our study was to replicate these egocentric bias findings using accuracy rate and reaction time measures in an online version of the Director Task. The second aim of the current study was to extend the Director Task methodology so that both egocentric and altercentric biases can be tapped in the same basic task format. To these ends, we built on one of the most recent implementations of the Director’s Task by Samuel and colleagues (2019) in order to implement a novel altercentric condition. In this condition, the director looked at the shelf from the participant’s side while a passive agent stood on the other side of the shelf (i.e., where the director stood in the egocentric bias version). No attention was given to this agent, and she had no apparent role in the task. The task itself remained the same in the altercentric bias version: the director asked the participant to find specific objects on the shelf. This time the director always had the same visual access to the objects as the participant. However, some of the objects referred to by the director were not visible from the irrelevant agent’s perspective due to occluders that disabled the view of some compartments. We investigated if participants were influenced by the irrelevant agent’s perspective and revealed altercentric bias such that they were slower and made more errors detecting the objects that were not visible to the agent compared to those that were visible to all parties.

Appendix B: Haskaraca, Proft, Liszkowski, & Rakoczy (2023)

Here we report two online studies. The first study tested adults on either the original egocentric or the adapted altercentric bias version of the Director Task. The second study aimed to replicate the first study with a more controlled dataset. Additionally, it included a non-social control version of the altercentric bias condition, in which the irrelevant agent was removed from the task. This condition was designed in order to rule out the possibility that the altercentric biases occur due to low-level domain-general situational factors that are related to the task design or materials but not implicit mentalizing. If altercentric interferences are found only in the social version but not in the non-social version of the task (or if they are stronger in the social version), this would imply that the divergent perspective of the (irrelevant) agent plays a role in the interference process. If, however, similar altercentric interferences are found in social and non-social conditions, this would imply that these effects are likely to occur due to situational demands instead of perspective calculation.

Study 1

Method

Participants

Participants were recruited through online advertisements and e-mail announcements. Initially, 224 participants were tested: 72 participants were excluded due to technical issues, 19 participants were excluded due to incomplete trials and visual impairments, and finally, 7 participants were excluded as they consistently answered the familiarization trials and all the critical test trials wrong, indicating failed comprehension of the task instructions. The final sample consisted of 126 German-speaking adults: 67 in the egocentric bias version (M_{age} = 35 years; 44 females and 23 males), 59 in the altercentric bias version (M_{age} = 35 years; 37 females, 21 males, and one non-binary individual). All participants gave informed consent before participating. This study has been preregistered (https://osf.io/nbxcd/?view_only=07cbb28bb5664017a37e089266b3a046).

Materials (The Director Task)

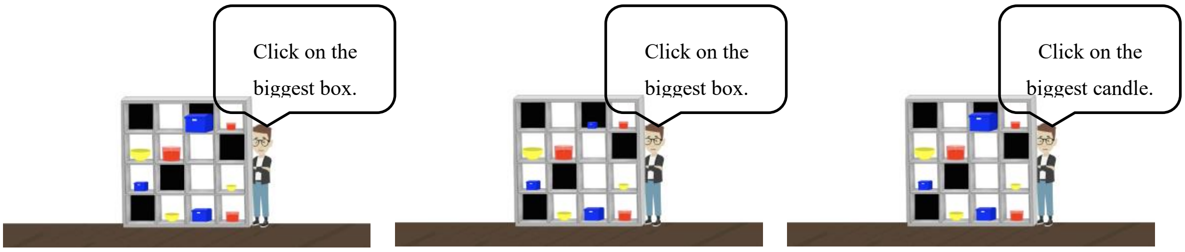
Based on the earlier studies using the real-life or computerized versions of the Director Task (e.g., Keysar et al., 2000; Samuel et al., 2019), an online version was created using Vyond (for material creation; https://www.vyond.com) and Labvanced (for designing and running the experiment; https://www.labvanced.com). The material consisted of a virtual room on a computer screen with a 3-D model of a shelf with 16 (4x4) compartments and an avatar. Five compartments of the shelf were always covered by opaque black occluders from the back. We used two different occluder constellations throughout the study. The contents of the occluded compartments were not visible from behind the shelf, where a computer-generated avatar was placed. This avatar depicted a young male in the egocentric bias condition, who corresponded to the director himself. In the altercentric bias condition, however, the avatar behind the shelf was an old female avatar who was irrelevant to the task. The director in the altercentric bias condition was the same young male avatar as in the egocentric bias condition; however, this time, he stood on the near side of the shelf and shared the same perspective as the participant.

Nine objects were placed randomly on the shelf. Only one object was placed per compartment. In each trial, there were three groups of three objects (triplets). The objects of a triplet were identical in shape and color but differed in size (see Figures 1 & 2 for sample occluder and object constellations).

In the study, participants were presented with either the egocentric or altercentric bias condition, consisting of comparable ambiguous and unambiguous trials and filler trials. In the ambiguous trials of the egocentric bias version (see Figure 1.a), the director instructed the participant to click on an object with a particular feature, e.g., “click on the biggest box”; however, the actual biggest box was occluded from the director’s side. Therefore, it was invisible to the director and could not be referred to by him. Hence, in these trials, participants should have gone for the medium-sized box, which was visible to both parties and corresponded to the biggest box from the director’s point of view. In the comparable unambiguous version of this sample trial (see Figure 1.b), the director’s instruction remained the same, e.g., “click on the biggest box.” However, this time the previously biggest box was replaced by a tiny one; therefore, the previously medium-sized box was now the biggest, and it was visible from both sides. In the filler trials (see Figure 1.c), the instruction referred to an object of a completely visible triplet (i.e., none of the three objects in this group was occluded, such as a candle in this case).

Figure 1

Sample Egocentric Bias Task Materials

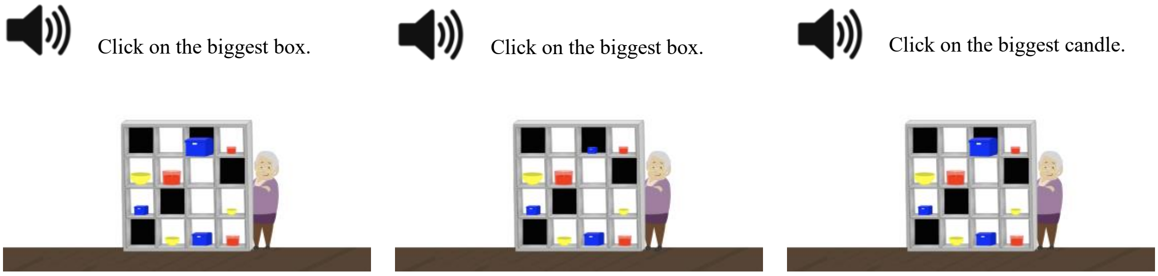


(a) ambiguous trial (b) unambiguous trial (c) filler trial

In this altercentric bias version, the director was standing on the same side of the shelf as the participant, therefore sharing his/her perspective. An irrelevant agent stood on the other side of the shelf and had limited visual access to the objects. The ambiguous trials in this condition referred to the trials where the searched object was not visible to the irrelevant agent, whereas, in the unambiguous trials, the irrelevant agent could see the searched object (see Figure 2).

Figure 2

Sample Altercentric Bias Task Materials



(a) ambiguous trial (b) unambiguous trial (c) filler trial

Design & Procedure

The study was conducted following a 2 (between-subjects factor: egocentric bias condition or altercentric bias condition) x 2 (within-subject factor: ambiguous trials and unambiguous trials) mixed design. More specifically, participants were randomly assigned either to the egocentric or the altercentric bias condition. In each condition, participants answered 12 sets of trials. Each set was composed of two filler trials and one critical trial (either ambiguous or unambiguous) in between. In total, participants answered 36 trials: 24 filler, six ambiguous, and six unambiguous. The order of the trial sets was pseudo-randomized so that comparable ambiguous and unambiguous trials did not directly follow each other.

After consenting to the study, participants watched a short video where the director's avatar introduced himself and took his position relative to the shelf based on the bias condition: at the other side of the shelf in the egocentric bias condition and the near side of the shelf in the altercentric bias condition. In the latter, the director disappeared from the screen once he took his position. After the introduction, participants were presented with three familiarization trials: one ambiguous, one unambiguous, and one filler trial. A manipulation check followed the familiarization trials. In the manipulation check, participants were asked about the director and his perspective to ensure their comprehension of the task. Participants who failed these questions were presented with the task instructions again. Then the main part of the experiment started.

The general procedure was identical for all types of trials. First, the shelf was presented with an avatar standing behind it (i.e., the director in the egocentric bias and the irrelevant agent in the altercentric group) for 2000 ms. This was followed by the director's verbal instruction lasting approximately 4500 ms. Then the participants were allowed up to 4000 ms to click on the object. As soon as an answer was given, or if no answer was given in 4000 ms, the procedure progressed with a new trial. In the end, participants were debriefed about the real aim of the study. The study took approximately 10-12 minutes.

Bias Calculation & Plan of Analysis

Biases were computed from the accuracy rates and the reaction times (milliseconds) of the participants in ambiguous versus unambiguous trials. The accuracy rate was defined as the proportion of correct answers out of all answers the participant gave in the ambiguous and unambiguous trials. In order to see if participants made fewer errors in the unambiguous trials than in the ambiguous trials (i.e., showed bias), separate within-subject comparisons were conducted for the egocentric and altercentric bias conditions. When the data were not normally distributed, the non-parametric alternative of the paired-sample t-test, i.e., the Wilcoxon test, was used to conduct within-subject comparisons.

The reaction times were measured as the time between the end of the instruction and the mouse click of the participant in the correctly answered trials. Then separate aggregate scores were created for both ambiguous and unambiguous trials by averaging the reaction times for these conditions. The average scores were then used to deduce biases: if the mean reaction times in the ambiguous trials were more extended than in the unambiguous trials, this indicated bias. As a result, separate within-subject comparisons were conducted to see if a bias existed in egocentric and altercentric bias groups. In addition to the analysis with the mean reaction time values, participants were also compared using median values, which are less prone to distortions from the outlier values. This strategy was based on Dumontheil et al. (2010), where the median reaction time values were tested in addition to the mean due to the small number of critical trials per participant, which is also the case in the current study.

Results

The means of the accuracy rates are depicted in Figure 3. The means and medians of reaction times are depicted in Figure 4. In the study, accuracy rates provided non-normally distributed data, whereas reaction times were normally distributed. Therefore, accuracy rates in ambiguous versus unambiguous trials were compared using Wilcoxon tests, whereas reaction times were compared via paired-sample t-tests. First, potential egocentric biases were investigated (i.e., if participants made more errors and became slower in the ambiguous trials than in unambiguous ones). Then, the same comparisons were conducted to reveal potential altercentric interferences.

Figure 3

Appendix B: Haskaraca, Proft, Liszkowski, & Rakoczy (2023)

Mean accuracy rates in Study 1

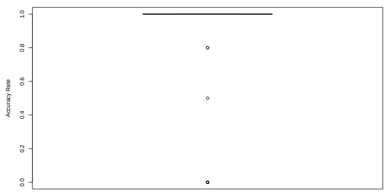
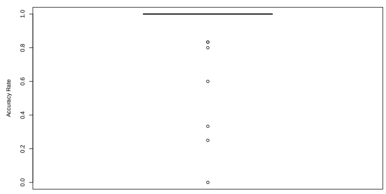
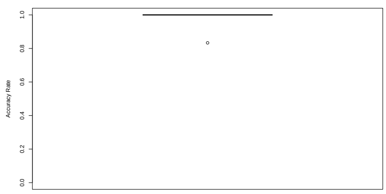
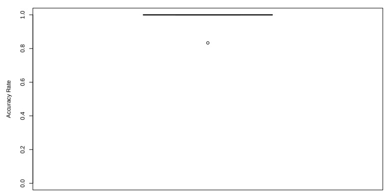
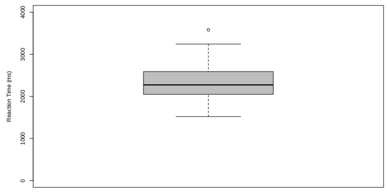
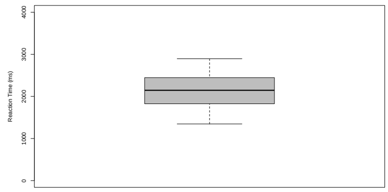
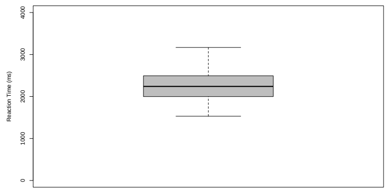
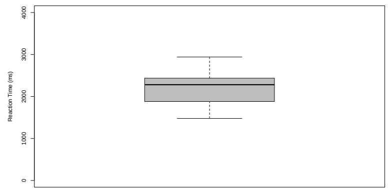
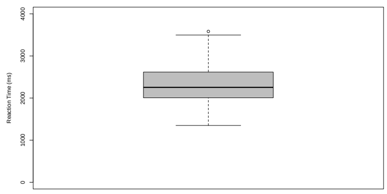
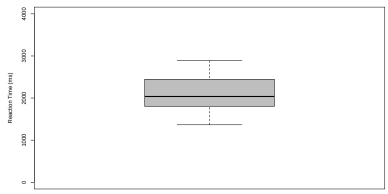
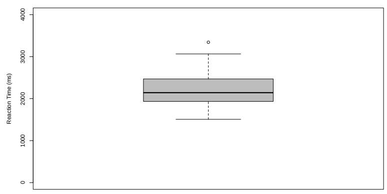
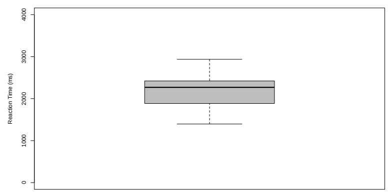
	Egocentric Bias	Altercentric Bias
Ambiguous Trials		
Unambiguous Trials		

Figure 4

Mean & median reaction times in Study 1

		Egocentric Bias	Altercentric Bias
Mean Reaction Times	Ambiguous Trials		
	Unambiguous Trials		
Median Reaction Times	Ambiguous Trials		
	Unambiguous Trials		

Does the Director Task reveal egocentric bias?

Within-subject comparisons revealed that participants' accuracy rates in the egocentric bias condition were lower in the ambiguous trials (M = .98; SD = .08) compared to the unambiguous trials (M = 1.0; SD = .0), Z=-3.571, p < .001. The reaction time measure with the mean values did not reveal any difference between ambiguous (M = 2330.30 ms; SD = 422.20) versus unambiguous (M = 2271.67 ms; SD = 368.70) trials, t(54) = -1.27, p = .21, *Cohen's d* = -.17; whereas the median reaction time values differed between ambiguous (M = 2309.10 ms; SD = 466.90) and unambiguous (M = 2209.52 ms; SD = 383.75) trials, t(54) = -2.12, p = .04, *Cohen's d* = -.29, suggesting slower reaction times in ambiguous trials.

Does the Director Task reveal altercentric bias?

Within-subject comparisons revealed that participants' accuracy rates in the altercentric bias condition were lower in the ambiguous trials (M = .96; SD = .14) compared to the unambiguous trials (M = 1.0; SD = .02), Z=-2.252, p = .02. The reaction time measures with mean values, however, revealed a difference favoring the ambiguous trials. Namely, participants responded to the ambiguous trials (M = 2136.56 ms; SD = 387.18) faster than the unambiguous (M = 2210.09 ms; SD = 382.13) trials, t(57) = 2.10, p = .04, *Cohen's d* = 0.28. The reaction time measure with median values revealed similar results, i.e., faster reaction times in the ambiguous trials (M = 2088.58 ms; SD = 405.53) compared to the unambiguous ones (M = 2174.85 ms; SD = 375.14) trials, t(57) = 2.34, p = .02, *Cohen's d* = 0.31.

Discussion

The main findings of the present study will be discussed separately for egocentric and altercentric bias versions of the Director Task. First, regarding the egocentric bias, the accuracy rate measure revealed that participants made more mistakes when moving objects according to someone else's incongruent perspective. The findings from the two reaction time measures, i.e., mean and median, were not as straightforward: mean reaction time measured did not reveal any difference between the ambiguous versus unambiguous trials, whereas median reaction times revealed slower reactions for ambiguous trials compared to unambiguous ones. These results provide partial support for the previously found egocentric interference effects in this task: the effects revealed by accuracy rates and median reaction times, but not mean reaction times, were in line with previous Director Task studies. Dumontheil and colleagues (2010) argued that median values are less affected by extreme data, which can distort the mean values to a greater extent, especially when the number of trials is not very high. However, most of the earlier Director Task studies found differences between trials using the mean reaction times (e.g., Apperly et al., 2010; Samuel et al., 2019). Therefore, the current study does not fully support the earlier findings and raises the question of whether the egocentric interference effects revealed by the present study are robust.

Regarding the altercentric bias, the accuracy rate measure revealed that participants made more mistakes in a simple search task when the scene involved a person with an incongruent perspective on the searched objects. For the altercentric bias version, the two reaction time measures, i.e., mean and median, provided consistent but intriguing findings: participants responded faster in ambiguous trials than in unambiguous ones. These results again provide only partial evidence for an altercentric bias as measured by the Director Task: the accuracy of participants' judgments decreased in the presence of an irrelevant agent with a different perspective, indicating an interference effect, whereas the same condition facilitated the speed of reaction. This result raises the possibility of a speed-accuracy trade-off: it is possible that participants made more errors in the ambiguous trials because they responded faster. In order to check this possibility, we explored if the accuracy was (negatively) related to reaction times in the given data. We found no correlation between the overall accuracy rate and reaction time measures. A closer look into the data has revealed that reaction times were negatively correlated with accuracy only for two of the six ambiguous trials ($r = -.384$, $p = .003$ and $r = -.371$, $p = .004$). Thus, the contradictory results revealed by the newly adapted altercentric bias version are still in need of further explanation.

Before arriving at any conclusion regarding the reliability of the egocentric interference effects and before further investigating the intriguing altercentric bias results, the findings of the current study should be replicated with better-quality data. The data in the present study were collected via convenience sampling, which resulted in a high number of dropouts due to lack of incentive and in many technical problems due to lack of rigorous equipment checks. Therefore, in Study 2, an exact replication of Study 1 was conducted on an established data collection platform, i.e., Prolific (https://www.prolific.co), where participants received incentive for their participation and a proper equipment check was possible. Also, with the increasing opportunities provided by the study platform (i.e., Labvanced), technical issues were minimized.

Study 2 also included a non-social condition in the altercentric bias task in which the irrelevant agent was removed from the scene. This condition aimed to rule out the alternative submentalizing explanations for the potential altercentric bias effects. Submentalizing accounts postulate that the altercentric interferences occur due to domain-general processes triggered by the task designs, not due to implicit mentalizing (e.g., Heyes, 2014a; 2014b). For example, Santiesteban and colleagues (2014) argued that the interferences found in the dot-perspective task are due to attentional/directional cues caused by the avatar's gaze rather than the incongruent perspective of the avatar. Following a similar logic, we eliminated the avatar in a non-social control condition in the current study but kept the rest of the task unchanged. If altercentric biases are found to be stronger in the social version compared to the non-social version, this would indicate that the presence of an irrelevant agent caused interference that cannot be explained otherwise. However, if similar interferences appear in social and non-social versions, this would imply that the interferences are not really altercentric; but they result from low-level situational factors.

Study 2

Method

Participants

Participants were recruited through Prolific. We used G*POWER (Faul et al., 2009) to conduct a power analysis and determine the sample size. We aimed to obtain .95 power to detect a medium effect size of .50 at the standard .05 alpha error probability with a more conservative two-tailed paired-samples t-test. Since the main aim of the current study was to tap biases revealed as the differences between experimental and control trials, we based the sample size rationale on within-subject comparisons. We computed the sample size separately for egocentric and altercentric bias versions. The analysis revealed a required sample size of 50 for each condition. Therefore, the final sample consisted of 100 English-speaking adult participants: 50 in the egocentric bias condition (M_{age} = 31 years; 16 females, 37 males, and one non-binary individual) and 50 in the altercentric bias condition (M_{age} = 30 years; 20 females and 34 males). For the non-social version of the altercentric bias task, 56 participants were tested. Two of them were excluded from the dataset due to technical problems, resulting in 54 participants in the non-social altercentric bias condition (M_{age} = 30 years; 15 females and 39 males). All participants gave informed consent before participating.

Materials, Design & Procedure, and Bias Calculation & Analysis

All methods of Study 2 were identical to Study 1, except for the additional non-social condition of the altercentric bias task. In the non-social version, no agent was displayed on the other side of the shelf, and the back of the director's head was displayed at the near end of the shelf (so that both altercentric bias conditions display an agent on the scene). Besides this difference, the non-social version was identical to the previously used altercentric bias task.

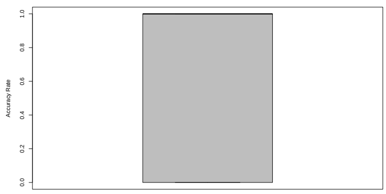
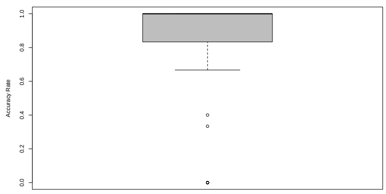
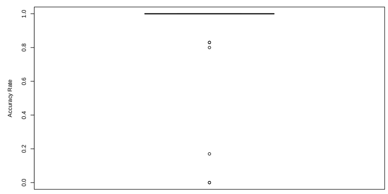
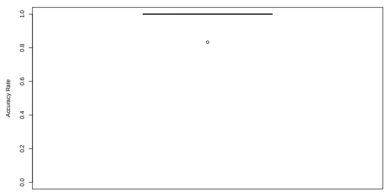
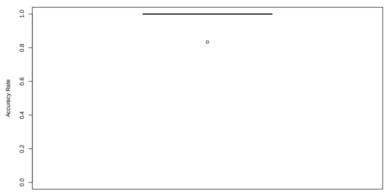
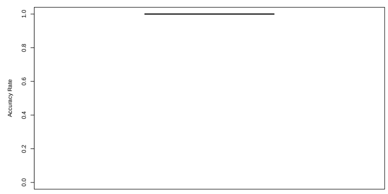
The non-social condition was introduced to the study design later than the other conditions. More specifically, as in Study 1, Study 2 was initially conducted following a mixed design of a 2 (between-subjects factor: egocentric or altercentric bias condition) x 2 (within-subject factor: ambiguous and unambiguous trials). Then, the data from the second round of egocentric and altercentric bias versions were analyzed. The results hinted at (partial) altercentric interference effects. Then, the nature and reliability of these effects were tested in the non-social condition with additional data. This condition had the same trial structure as the social altercentric bias task.

Results

The means of the accuracy rates are depicted in Figure 5. The means and medians of reaction times are depicted in Figure 6. In the study, accuracy rates always provided non-normally distributed data. Reaction times were non-normally distributed only for altercentric bias. Therefore, accuracy rates in all conditions and the reaction times in the altercentric bias condition were analyzed by using Wilcoxon tests. The remaining reaction time analyses were conducted by using paired-sample t-tests. First, potential egocentric biases were investigated (i.e., if participants made more errors and became slower in the ambiguous trials than in unambiguous ones). Then, the same comparisons were conducted to reveal potential altercentric interferences. Finally, the same comparisons were conducted to reveal if any interference was present in the non-social condition. Moreover, for the social and non-social altercentric bias tasks, the effect of condition (social vs. non-social) and trial type (ambiguous vs. unambiguous) on the reaction was investigated via a mixed ANOVA, and the effect of condition on accuracy rates was investigated via a Mann-Whitney Test.

Figure 5

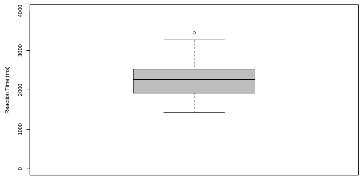
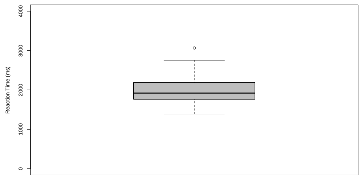
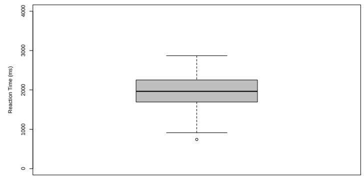
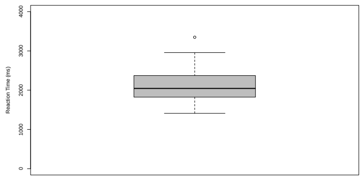
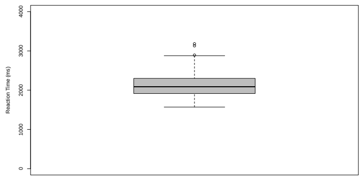
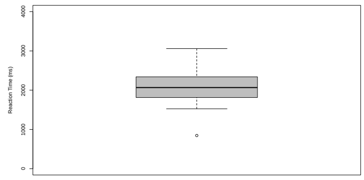
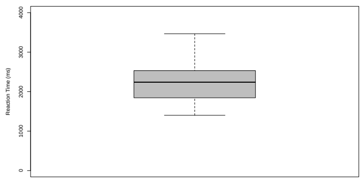
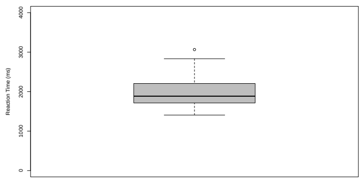
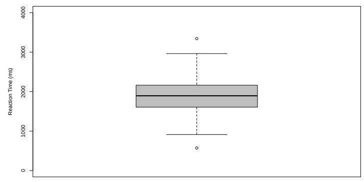
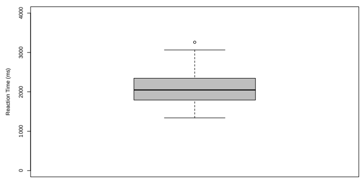
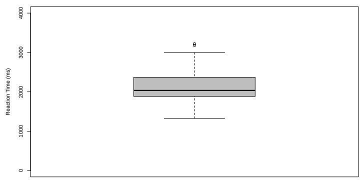
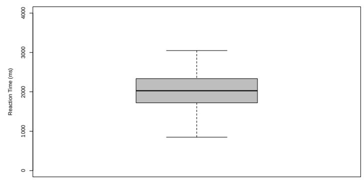
Mean accuracy rates in Study 2

	Egocentric Bias	Altercentric Bias	Altercentric Bias (Non-Social)
Ambiguous Trials			
Unambiguous Trials			

Appendix B: Haskaraca, Proft, Liszkowski, & Rakoczy (2023)

Figure 6

Mean & median reaction times in Study 2

		Egocentric Bias	Altercentric Bias	Altercentric Bias (Non-Social)
Mean Reaction Times	Ambiguous Trials			
	Unambiguous Trials			
Median Reaction Times	Ambiguous Trials			
	Unambiguous Trials			

Does the Director Task reveal egocentric bias?

Within-subject comparisons have revealed that participants' accuracy rates were lower in the ambiguous trials (M = .96; SD = .13) compared to the unambiguous trials (M = 1.0; SD = .0), Z=-4.013, p < .001. The reaction time measure with the mean values revealed similar results: participants were slower in the ambiguous trials (M = 2303.68 ms; SD = 492.40) compared to unambiguous (M = 2139.71 ms; SD = 433.10) trials, t(37) = -2.71, p = .01, *Cohen's d* = -.44 . Median reaction time values also differed between ambiguous (M = 2274.33 ms; SD = 537.65) and unambiguous (M = 2103.17 ms; SD = 439.30) trials, t(37) = -2.64, p = .01, *Cohen's d* = -.43.

Does the Director Task reveal altercentric bias?

Within-subject comparisons have revealed that participants' accuracy rates were lower in the ambiguous trials (M = .78; SD = .40) compared to the unambiguous trials (M = 1.0; SD = .02), Z=-3.440, p < .001. The reaction time measures with mean values, however, revealed a difference favoring the ambiguous trials. Namely, participants responded to the ambiguous trials (M = 1993.50 ms; SD = 381.34) faster than the unambiguous (M = 2146.11 ms; SD = 380.43) trials, Z=-2.557, p = .01. The reaction time measure with median values revealed similar results, i.e., faster reaction times in the ambiguous trials (M = 1954.76 ms; SD = 385.65) compared to the unambiguous ones (M = 2113.99 ms; SD = 390.92) trials, Z=-2.472, p = .01.

Altercentric bias in social versus non-social conditions

Within-subject comparisons were conducted to see if ambiguous and unambiguous trials in the non-social condition differed. These comparisons revealed that participants' accuracy rates were lower in the ambiguous trials (M = .94; SD = .22) compared to the unambiguous trials (M = 1.0; SD = .0), Z=-2.214, p = .027. The reaction time measures with mean values again revealed a difference favoring the ambiguous trials. Namely, participants responded to the ambiguous trials (M = 1955.27 ms; SD = 450.37) faster than the unambiguous (M = 2125.29 ms; SD = 425.02) trials, t(53) = 5.21, p < .001, *Cohen's d* = .71. The reaction time measure with median values revealed similar results, i.e., faster reaction times in the ambiguous trials (M = 1886.37 ms; SD = 500.18) compared to the unambiguous ones (M = 2081.55 ms; SD = 453.10) trials, t(53) = 4.74, p < .001, *Cohen's d* = .65.

Then a 2 (condition: social versus non-social) x 2 (trial type: ambiguous versus unambiguous) Mixed MANOVA was conducted in order to investigate the effect of condition and trial type on RT measures. Only a main effect of trial type has been revealed (Mean RT: $F(1,97)=24.73$, $p<.001$ & Median RT: $F(1,97)=24.76$, $p<.001$). No main effect of condition and no interaction effect has been found. For the non-normally distributed accuracy measures, a Mann-Whitney test has been used to see the effect of the condition on experimental and control trials. The condition did not show an effect on the accuracy rates of control trials, whereas the accuracy in the experimental trials of the social condition ($M = .78$; $SD = .40$) was lower than those in the non-social condition ($M = .94$; $SD = .22$), $U=1225.00$, $Z=-2.116$, $p=.03$.

Discussion

First, the findings pertaining to egocentric and altercentric bias versions of the Director Task will be discussed. Then the interferences found in the non-social condition of the altercentric bias and the (lack of) condition effect will be debated. In the egocentric bias condition of the current study, the accuracy rate measure revealed decreased accuracy when participants were asked to move objects according to someone else's incongruent perspective compared to congruent. The results from the two reaction time measures supported the interferences revealed in the accuracy measure. Namely, they revealed slower reactions for ambiguous trials compared to unambiguous ones. It can be speculated that the mean reaction time measure revealed a difference in Study 2 but not in Study 1 due to the improvements in data quality and technical retrofits, which reduced extreme values in the data. Overall, these results aligned with the findings of the previous Director Task studies and replicated the egocentric bias effects in a completely online version of the task.

The findings from the altercentric bias version of Study 2 were in line with the first study. Participants made less accurate but faster judgments in a search task when the scene involved a person with an incongruent perspective on the searched objects. Then, similar to Study 1, the possibility of speed-accuracy trade-off was investigated for Study 2 too. A closer look at the data revealed no correlation between overall accuracy rate and reaction time measures, and none of the trials revealed a negative correlation between accuracy and speed.

Finally, the non-social condition of the altercentric bias task (i.e., without any agent present in the scene) showed the same interference pattern as the social version: participants were more error-prone but also faster in ambiguous trials compared to unambiguous trials. In this condition, there is no agent on the other side of the shelf; hence, no incongruent perspective that can interfere with participants' own judgments. Nevertheless, participants were still affected in their judgments if the searched object was not visible from the other side of the shelf. In fact, presenting the altercentric bias task with or without an irrelevant agent did not have any effect on the reaction times of the participants. Condition only revealed an effect on the accuracy rates of ambiguous trials, which was very low in the social condition.

General Discussion

The present study capitalized on the so-called Director Task in order to tap egocentric and altercentric biases. Across two online studies, we found evidence for robust egocentric interferences. The findings from the altercentric bias version drew a more complex picture. More specifically, the accuracy rates and reaction times favored different types of trials in the altercentric bias version: participants were more accurate but slower when an irrelevant agent's perspective on the task material aligned with their own, whereas they were less accurate but faster when the situation involved an irrelevant agent whose perspective differed from the agents. Also, a non-social modification of the altercentric bias version revealed similar interference effects as the social version. In the following, we first discuss the reliability of the egocentric interference effects found in the Director Task. Then we discuss the validity and nature of the interferences found in this task's newly adapted altercentric bias version.

Egocentric Interferences Revealed by the Director Task

Director Task has been repeatedly used to reveal egocentric bias effects in communication and social cognition (e.g., Keysar et al., 2000; Samuel et al., 2019). Different versions of this task (e.g., in-person version with confederates or computerized version with computer-generated avatars) have shown that people are subject to egocentric biases such that their own perspective or knowledge interfere with their judgments about other people's perspectives, even though they are explicitly and repeatedly asked to take someone else's perspective (e.g., Apperly et al., 2010; Dumontheil et al., 2010; Keysar et al., 2000; Legg et al., 2017; Samuel et al., 2019). However, no earlier studies tested the Director Task in a completely online, unmoderated format. The current study filled this gap with two studies in which both participant recruitment and data collection were handled 100% online and unmoderated. Recently, systematic comparisons of in-presence versus online studies conducted in sociocognitive development research have revealed mixed findings. For example, Sheskin and Keil (2018) have found that children show poorer performance in the online FB tasks compared to the live versions. More recently, Schidelko and colleagues (2021) found no difference between the two versions. The current study provided support for the latter line of findings. However, before declaring the suitability of the Director Task for online testing, systematic and direct comparisons with the live versions should be conducted.

One point that needs further attention in the current study, and also in potential future studies, is the reliability of the reaction time measures in online tasks. So far, the reaction times measures in web-based studies provided comparable results to lab-based studies and have been shown to detect even very brief time intervals (e.g., Hilbig, 2015). However, the discrepancy between the mean and median reaction times in the first study of this paper has revoked the possibility that reaction times tend to get extreme values in less-controlled task designs. Therefore, preventive measures are needed. In this sense, using more than one reaction time measure (e.g., mean and median) and improving the data quality (e.g., via participant recruitment strategies and technical enhancements) can be useful tactics to get more reliable results.

Altercentric Interferences Revealed by the Director Task

The findings from the altercentric bias version of the Director Task are challenging to interpret for two reasons. First, the adaptation of the altercentric bias version of the Director Task is original, meaning there is no existing literature with which we can compare the current results. Secondly, the two measures used in the current study (i.e., accuracy rate and reaction times) pointed in opposite directions in terms of the altercentric interferences: participants were more error-prone but faster when an irrelevant agent had an incongruent perspective on the materials, compared to when the irrelevant agent had a congruent perspective. This inconsistency connoted a potential speed-accuracy trade-off; however, the data were not supporting this possibility. Then what is the reason for the inconsistency between accuracy and reaction time measures?

One answer to this question can be found in the potential differentiation between the mechanisms tapped by accuracy and reaction times. This kind of differentiation has been revealed in different areas of psychology. For example, the research on attention has shown that accuracy and reaction times are affected differentially by voluntary and involuntary mechanisms (e.g., Prinzmetal et al., 2005). A similar distinction may have been at play in the current study, making the accuracy and reaction time measures reveal differential interference effects. The presence and nature of these mechanisms remain to be studied. Considering the concerns mentioned above regarding the reaction times measures in unmoderated online studies, future research should first ensure that the reaction time is a reliable measure, even in web-based studies.

Finally, the current study showed that the presence or absence of an irrelevant agent had no effect on the interferences revealed by the newly adapted altercentric bias version of the Director Task. This finding speaks against the validity of the altercentric bias version of this task as the interferences occurred even in the absence of an irrelevant perspective: if the interferences were truly *altercentric*, namely other-centric, they would have been specific to the social condition. In light of this result, it is possible that the newly adapted version of the Director Task, which supposedly targeted altercentric interferences, is instead revealing low-level domain-general perceptual or attentional interferences triggered by the task design. For example, it is possible that participants chose the objects that were visible for both sides in the ambiguous trials more readily than the one standing in front of a black screen (which was the correct answer) due to the color contrast between the white background and colorful object. For the social condition of the altercentric bias version, it is also possible that participants, during the task, failed to remember that the agent was irrelevant and instead started to think of this person as the director and responded from her perspective. This possibility is supported by the exceptionally low accuracy found in the ambiguous trials of social condition (which eventually caused the accuracy rates to differ between the two conditions). Future research should eliminate any potential domain-general interferences before using the current task format as a potential measure to tap altercentric interferences.

Conclusion

The current study capitalized on the so-called Director Task to tap implicit and explicit perspective-taking abilities via altercentric and egocentric biases, respectively. The results replicated the previously found egocentric bias effects in an online version of the Director Task. The findings pertaining to the altercentric bias version were likely due to domain-general attentional cues. It is crucial for future studies to carefully consider this possibility before making use of the Director Task as an altercentric bias measure.

References

Apperly, I. A., Carroll, D. J., Samson, D., Humphreys, G. W., Qureshi, A., & Moffitt, G. (2010). Why are there limits on theory of mind use? Evidence from adults' ability to follow instructions from an ignorant speaker. *Quarterly Journal of Experimental Psychology, 63*(6), 1201-1217. https://doi.org/10.1080/17470210903281582

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

Cane, J. E., Ferguson, H. J., & Apperly, I. A. (2017). Using perspective to resolve reference: The impact of cognitive load and motivation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(4), 591–610. https://doi.org/10.1037/xlm0000345

Cole, G. G., Atkinson, M., Le, A. T. D., & Smith, D. T. (2016). Do humans spontaneously take the perspective of others? *Acta Psychologica, 164*, 165–168. https://doi.org/10.1016/j.actpsy.2016.01.007

Conway, J. R., Lee, D., Ojaghi, M., Catmur, C., & Bird, G. (2017). Submentalizing or mentalizing in a Level 1 perspective-taking task: A cloak and goggles test. *Journal of Experimental Psychology: Human Perception and Performance, 43*(3), 454–465. https://doi.org/10.1037/xhp0000319

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

Dumontheil, I., Apperly, I. A., & Blakemore, S. J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science, 13*(2), 331-338. https://doi.org/10.1111/j.1467-7687.2009.00888.x

Dumontheil, I., Hillebrandt, H., Apperly, I. A., & Blakemore, S. J. (2012). Developmental differences in the control of action selection by social information. *Journal of Cognitive Neuroscience, 24*(10), 2080-2095. https://doi.org/10.1162/jocn_a_00268

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Flavell, J. H. (1974). The development of inferences about others. In T. Mischel (Ed.), *Understanding other persons* (pp. 66–116). Blackwell.

Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the Level I-Level 2 distinction. *Developmental Psychology, 17*, 99–103. https://doi.org/10.1037/0012-1649.17.1.99

Henry, J. D., Phillips, L. H., Ruffman, T., & Bailey, P. E. (2013). A meta-analytic review of age differences in theory of mind. *Psychology and Aging, 28*(3), 826–839. https://doi.org/10.1037/a0030677

Hilbig, B. E. (2016). Reaction time effects in lab-versus Web-based research: Experimental evidence. *Behavior Research Methods, 48*, 1718-1724. https://doi.org/10.3758/s13428-015-0678-9

Kampis, D., & Kovács, Á. M. (2021). Seeing the world from others' perspective: 14-month-olds show altercentric modulation effects by others' beliefs. *Open Mind, 5*, 189–207. https://doi.org/10.1162/opmi_a_00050

Kampis, D., & Southgate, V. (2020). Altercentric cognition: How others influence our cognitive processing. *Trends in Cognitive Sciences, 24*(11), 945-959. https://doi.org/10.1016/j.tics.2020.09.003

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science, 11*(1), 32-38. https://doi.org/10.1111/1467-9280.00211

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

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

Legg, E. W., Olivier, L., Samuel, S., Lurz, R., & Clayton, N. S. (2017). Error rate on the director's task is influenced by the need to take another's perspective but not the type of perspective. *Royal Society open science, 4*(8), 170284. https://doi.org/10.1098/rsos.170284

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

O'Grady, C. J., Scott-Phillips, T., Lavelle, S., & Smith, K. (2017). The dot perspective task revisited: Evidence for directional effects. In *Cogsci 2017: Proceedings of the 39th annual meeting of the cognitivie science society* (pp. 2821–2826). https://pdfs.semanticscholar.org/90a2/0cdfbc0bd26c0b977 d6294d7f5ade726f7db.pdf

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

Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology, 5*(2), 125–137. https://doi.org/10.1111/j.2044- 835x.1987.tb01048.x

Perner, J., & 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

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

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

Priewasser, B., Rafetseder, E., Gargitter, C., & Perner, J. (2018). Helping as an early indicator of a theory of mind: Mentalism or Teleology?. *Cognitive Development, 46*, 69-78. https://doi.org/10.1016/j.cogdev.2017.08.002

Prinzmetal, W., McCool, C., & Park, S. (2005). Attention: Reaction Time and Accuracy Reveal Different Mechanisms. *Journal of Experimental Psychology: General, 134*(1), 73–92. https://doi.org/10.1037/0096-3445.134.1.73

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

Samuel, S., Roehr-Brackin, K., Jelbert, S., & Clayton, N. S. (2019). Flexible egocentricity: Asymmetric switch costs on a perspective-taking task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(2), 213-218. https://doi.org/10.1037/xlm0000582

Samuel, S., Roehr-Brackin, K., & Roberson, D. (2016). 'She says, he says': Does the sex of an instructor interact with the grammatical gender of targets in a perspective-taking task? *International Journal of Bilingualism, 20*(1), 40-61. https://doi.org/10.1177/1367006915576831

Santiesteban, I., Catmur, C., Hopkins, S. C., Bird, G., & Heyes, C. (2014). Avatars and arrows: implicit mentalizing or domain-general processing*? Journal of Experimental Psychology. Human Perception and Performance, 40*(3), 929–937. https://doi.org/10.1037/a0035175

Schidelko, L. P., Schünemann, B., Rakoczy, H., & Proft, M. (2021). Online testing yields the same results as lab testing: a validation study with the false belief task. *Frontiers in Psychology, 12*, 4573. https://doi.org/10.3389/fpsyg.2021.703238

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

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

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

Sheskin, M., & Keil, F. (2018, December 30). TheChildLab.com A Video Chat Platform for Developmental Research. https://doi.org/10.31234/osf.io/rn7w5

Sommerville, J. A., Bernstein, D. M., & Meltzoff, A. N. (2013). Measuring beliefs in centimeters: Private knowledge biases preschoolers' and adults' representation of others' beliefs. *Child Development, 84*(6), 1846-1854. https://doi.org/10.1111/cdev.12110

Southgate, V. (2020). Are infants altercentric? The other and the self in early social cognition. *Psychological Review, 127*(4), 505-523. https://doi.org/10.1037/rev0000182

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

Van der Wel, R. P., Sebanz, N., & Knoblich, G. (2014). Do people automatically track others' beliefs? Evidence from a continuous measure. *Cognition, 130*(1), 128-133. https://doi.org/10.1016/j.cognition.2013.10.004

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72*(3), 655–684. https://doi.org/10.1111/1467-8624.00304

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

Yott, J., & Poulin-Dubois, D. (2016). Are infants' theory-of-mind abilities well integrated? Implicit understanding of intentions, desires, and beliefs. *Journal of Cognition and Development, 17*(5), 683-698. https://doi.org/10.1080/15248372.2015.1086771

Appendix C: Haskaraca, Proft, Liszkowski, & Rakoczy (2023)

Haskaraca, F.N., Proft, M., Liszkowski, U., & Rakoczy, H. (2023). *Measuring Altercentric Biases in Level-1 & Level-2 Perspective-Taking Tasks by the Means of Mouse-Tracking* [Unpublished manuscript]. Department of Developmental Psychology. University of Göttingen

Appendix C: Haskaraca, Proft, Liszkowski, & Rakoczy (2023)

Measuring Altercentric Biases in Level-1 & Level-2 Perspective-Taking Tasks by the Means of Mouse-Tracking

Appendix C: Haskaraca, Proft, Liszkowski, & Rakoczy (2023)

Abstract

Altercentric bias refers to an implicit and automatic form of representing others' perspectives: our own judgments and behavior are modulated by how we think other agents perceive the world, indicating that we implicitly represent their perspectives, even when those are irrelevant or interfere with our own task (e.g., Southgate, 2020). In our upcoming study, we investigate this bias by capitalizing on previously used mouse-tracking measures (e.g., Van der Wel et al., 2014): when subjects are asked to move the mouse cursor to the location of a target object, they take a little detour on their way to the correct answer when another agent in the scenario has a deviant belief. The area under this detour thus indicates whether/to what degree participants engage in altercentric bias. The novelty of the current study is that it employs mouse-tracking measures to explore altercentric biases in Level-I versus Level-II perspective-taking (PT) tasks. Whether and to what degree these widely automatic biases are found in different levels of PT tasks has theoretical implications. On the one hand, two-system accounts expect biases only in Level-I tasks capturing automatic System-I but not in Level-II tasks tapping effortful System-II. On the other hand, nativism accounts assume unified PT abilities and predict no differences between the two levels of PT. In two online studies, English-speaking adult participants were tested on Level-I and Level-II PT tasks. In Study 1, they were tested on either altercentric bias or egocentric bias, i.e., a mirror image of altercentric biases in the sense that they correspond to the interferences from one's own perspective on their judgments about others' perspectives. Results revealed altercentric biases for both levels, whereas egocentric biases were evident only for Level-II. In Study 2, participants were tested either in the social or non-social version of the altercentric bias task in order to test the effect of an irrelevant agent on altercentric interference effects. Results revealed interferences in the social version (similar to Study 1), but no interference effect was observed in the non-social version. The results will be discussed in terms of their methodological bona fides and theoretical implications.

Keywords: Theory of Mind, altercentric bias, mouse-tracking, level-I perspective-taking, level-II perspective-taking

Appendix C: Haskaraca, Proft, Liszkowski, & Rakoczy (2023)

Measuring Altercentric Biases in Level-1 & Level-2 Perspective-Taking Tasks by the Means of Mouse-Tracking

Theory of Mind (ToM) -the ability to ascribe mental states to ourselves and others- is an integral aspect of our social nature. At its core, ToM involves the ability to meta-represent, i.e., representing our own and others' representational states. A key indicator of this capacity is the ability to represent misrepresentations or false beliefs of others, namely understanding how others represent the world, even when their representations are incorrect. Therefore, False Belief (FB) tasks have emerged as crucial assessments of meta-representational ToM abilities. One widely used FB task is the change-of-location task (e.g., Wimmer & Perner, 1983), where participants are presented with a story involving the transfer of an object from one location to another in either presence or absence of a protagonist. Participants are then asked where the protagonist will look for the object. Extensive research over the years has demonstrated that children typically develop the ability to solve this task (i.e., they start responding according to the protagonist's mistaken belief rather than reality) around the age of four (Wellman et al., 2001). Similar shifts in performance have been observed in other FB tasks and superficially different ToM tasks (e.g., Perner et al., 1987; Perner & Roessler, 2012), leading to the assumption that ToM emerges and goes through fundamental changes in the preschool years, specifically around age four.

In the past 15 years, new findings have emerged, challenging the traditional understanding of ToM development and introducing the concept of implicit ToM. Implicit ToM measures assess FB attribution without relying on verbal measures. Surprisingly, these studies have demonstrated that infants as young as one possess an understanding of mistaken beliefs, which was much earlier than the previously assumed age of four (see for a review, Scott & Baillargeon, 2017). Various types of tasks have been utilized to investigate this early FB understanding, including violation of expectation (VoE), anticipatory looking (AL), and interactive paradigms (e.g., Buttelmann et al., 2017; Onishi & Baillargeon, 2005; Southgate et al., 2007). VoE paradigms showed that children exhibited prolonged looking when events deviated from their expectations. For instance, when an agent's actions were inconsistent with her false belief, children displayed increased looking times (e.g., Onishi & Baillargeon, 2005). In AL paradigms, children’s gaze patterns indicated that they expected agents to behave in accordance with their false beliefs (e.g., Southgate et al., 2007). Finally, in interactive tasks, children were able to incorporate others' false beliefs into their communicative interactions (e.g., Buttelmann et al., 2009). Overall, these implicit measures showed that children possess an implicit sensitivity to others' (false) beliefs much earlier than age four.

Various theoretical accounts have emerged to explain the discrepancy between the findings of explicit and implicit ToM tasks. On the one hand, nativist accounts (e.g., Leslie, 2005; Scott & Baillargeon, 2017) suggest that infants possess innate ToM abilities that enable them to understand and attribute mental states. According to nativist accounts, the ToM abilities remain mostly unified throughout the lifespan, but they become more sophisticated and refined over time. These accounts claim that traditional FB tasks cannot reveal the true performance of infants because of their high cognitive and linguistic demands. On the other hand, two-system views (e.g., Apperly & Butterfill, 2009; Butterfill & Apperly, 2013; Low et al., 2016) interpret the early capacity revealed by the implicit tasks as a mere sensitivity to beliefs, not as the subjective, full-blown understanding measured by traditional (verbal) FB tasks. These accounts integrate the findings of implicit and explicit tasks by positing two distinct systems: System I, which is limited but efficient and is evolutionarily and ontogenetically more ancient, explaining the results of implicit tasks, and System II, which is a later-developing, complex, and more flexible system, accounting for full-blown performance in explicit tasks.

Dual-system accounts also make distinctions in terms of the types of mental state representations that are handled by the two systems. System I is characterized by signature limits, specifically regarding processing propositional attitudes. This system is unable to track an agent's perception of reality or determine how the agent represents something. Instead, this simpler system operates based on relational belief-like states, which can only track an agent's ability to perceive certain things (e.g., Low & Watts, 2013). For example, System I is limited to level-I perspective-taking, which corresponds to understanding that different viewers can see different things in a given situation (Flavell et al., 1981). On the other hand, System II is more flexible and capable of working with proper propositional attitudes. It enables reasoning about an agent's perception and whether the agent represents reality differently. For instance, System II is necessary for successful level-II perspective-taking, which refers to understanding that different viewers can see the same thing differently depending on their perspectives (Flavell et al., 1981).

This distinction between level-I and level-II perspective-taking can be useful when testing the two mainstream theories on implicit ToM findings against each other, i.e., nativism accounts and two-system theories. More specifically, according to the two-system approach, level-II perspective-taking serves as a signature limit for implicit ToM such that automatic perspective-taking abilities occur for level-I perspective-taking but not in level-II. In contrast, nativist accounts propose a unified ToM ability and do not expect a difference between the two levels. According to these accounts, foundational cognitive processes that are necessary for ToM are present from early infancy, and they remain consistent across different levels of perspective-taking. Based on these assumptions, examining implicit ToM in level-I versus level-II perspective-taking tasks can help distinguish between these theories. More specifically, if implicit perspective-taking manifests itself in both level-I and level-II tasks, this would support the nativism accounts. If, however, implicit perspective-taking is observed in level-I tasks but not in level-II tasks, then this would provide support for the dual-system accounts.

Before subjecting this idea to empirical investigation, a reliable implicit ToM task is warranted since the standard measures of implicit ToM (i.e., VoE, AL, and interaction tasks) have been shown to be questionable in terms of their reproducibility and reliability (e.g., Dörrenberg et al., 2018; Poulin-Dubois & Yott, 2018; Powell et al., 2018; Schuwerk et al., 2018). Another type of implicit task, the altercentric bias, still shows promise as an alternative measure for assessing implicit perspective-taking abilities. This bias refers to the interferences from others' perspectives on our first-order judgments (Southgate, 2020). For instance, people tend to get slower and more error-prone while judging a state of the world if there is another agent in the scene who holds a diverging perspective, even if this perspective is irrelevant or detrimental to the task (e.g., Kovács et al., 2010; Samson et al., 2010).

So far, a few studies have investigated the implicit perspective-taking abilities in level-I versus Level-II perspective-taking tasks by utilizing altercentric interference effects (e.g., Surtees et al., 2012; Surtees et al., 2016). For example, in a study by Surtees and colleagues (2012), altercentric bias effects were revealed for level-I perspective-taking only but not for level-II. However, the findings of these studies may not be conclusive for several reasons. First, the limited research on this issue has almost exclusively depended on one task format, namely the so-called dot-perspective task (Samson et al., 2010), and its variations. However, recently, the dot-perspective task has also been shown to be subject to reliability and validity issues (e.g., Cole et al., 2016; Conway et al., 2017; Santiesteban et al., 2014). Also, in the earlier studies, the materials used to test level-I versus level-II perspective-taking were slightly different, making it harder to compare the effects found in these two different tasks. Finally, so far, the biases have been searched in response times and accuracy measures, which have been criticized for being susceptible to situational factors and for providing binary -not continuous- outcomes (e.g., Marshall et al., 2018).

The current studies propose a solution to these methodological challenges with the primary objective of investigating the presence of altercentric biases in the two levels of visual perspective-taking. This investigation will provide empirical support for one of the two competing theories on implicit ToM. To this end, the current study capitalizes on a different type of measure: mouse-tracking. Mouse trajectories have been used to reveal altercentric biases before (Van der Wel et al., 2014). In this task, participants were asked to move their mouse cursors to mark a target location on the computer screen. On the one hand, when another agent in the scenario had a belief that differed from their own, participants' mouse cursors followed a convex route deviating in the direction of the incorrect location endorsed by the agent's belief. On the other hand, when the participant and the agent had the same belief about the location of the target object, participants' mouse cursors followed a more direct route. The extent of this detour, measured as the area between the convex deviation towards the incorrect answer and the direct line connecting the starting point to the target location, provided insights into the presence and magnitude of the altercentric bias.

In the present studies, mouse-tracking measures were incorporated into the typical level-I and level-II perspective-taking tasks (e.g., Flavell et al., 1981). This allowed us to tap altercentric biases in both levels via analogous tasks while providing an extra measure to reveal the potential biases, i.e., mouse trajectories, in addition to response times and accuracy. Study 1 explored the presence of altercentric bias by using this novel task format in level-I versus level-II perspective-taking tasks. This study also investigated egocentric biases in order to determine if the current task format can be used to reveal both implicit and explicit perspective-taking abilities. Egocentric biases refer to the interferences from our own perspective or knowledge on our judgments about others' perspectives, and they occur even when we are explicitly asked to focus on the perspectives of others (e.g., Samson et al., 2010; Sommerville et al., 2013). This bias can then serve as an indicator of (the lack of) explicit ToM abilities. In Study 2, the focus shifts to further investigations of the robustness of the altercentric biases revealed by the current task format. These investigations examined the altercentric interferences in social versus non-social contexts. This distinction aimed to ensure that the altercentric biases were specifically due to the irrelevant agent's perspective but not due to low-level domain-general situational factors, as would be predicted by non-mentalizing accounts (e.g., submentalizing by Heyes, 2014). These two studies also provided insights into which measure, whether it be mouse-tracking, reaction time, or accuracy, offers a more reliable means to assess altercentric bias within this task format.

Appendix C: Haskaraca, Proft, Liszkowski, & Rakoczy (2023)

Study 1

Method

Participants

A total of 100 English-speaking participants were recruited from an online data collection platform (Prolific, https://www.prolific.co): 50 in the egocentric bias condition (M_{age} = 28 years; 25 females and 25 males) and 50 in the altercentric bias condition (M_{age} = 27 years; 25 females and 25 males). All participants gave informed consent before participating and received an incentive upon their participation. The study has been preregistered (https://osf.io/2ej78/?view_only=1b7e75cf61a24d19a000e7e4bda83b35).

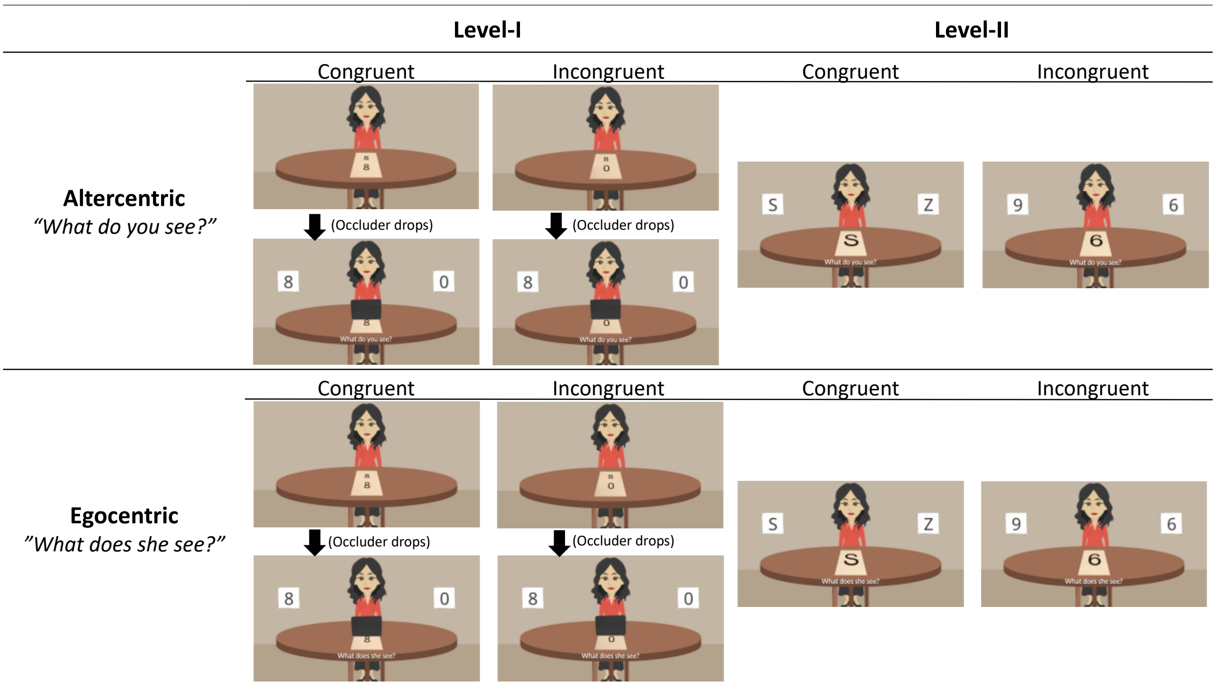
Materials

The logic behind the materials used in this study was adapted from the classic level-I, and level-II perspective-taking tasks used in developmental psychology research (e.g., Flavell et al., 1981), and the stimuli were inspired by a study by Surtees and colleagues (2012). The materials were constructed by integrating real-life pictures into videos created on Vyond (https://www.vyond.com). The study was designed and run on Labvanced (https://www.labvanced.com). The materials consisted of a virtual room with a computer-generated avatar looking at the table's surface whose sex matched the participant. The table on the screen was always located between the avatar and the participant. A sheet of paper displaying numbers or letters was placed on the table between the avatar and the participant.

In the level-I task, the paper initially showed two identical numbers/letters in congruent trials and two different numbers/letters in incongruent trials (see Figure 1). After a short display of the stimuli, an occluder entered the scene and blocked the view of one number/letter from each perspective. In the level-II task, only one number/letter was presented on the paper. This number/letter looked the same from both perspectives in congruent trials and looked different in incongruent trials (see Figure 1). Then participants were asked to identify the stimulus presented on the screen either from their own perspective (altercentric bias condition) or from the avatar’s perspective (egocentric bias condition). Although it had no apparent role in the task, the avatar remained in the scene in the altercentric bias condition. Participants answered by clicking on one of the two options presented on the screen. One option was always the correct answer. The other corresponded to the right answer from the opposite perspective in incongruent trials or to a similarly shaped distractor number/letter in congruent trials. The relative location of the correct answer (right or left) was counterbalanced across trials.

Figure 1

Sample Materials



Design & Procedure

The study was conducted following a 2 (between-subjects factor: egocentric or altercentric bias condition) x 2 (within-subject factor: level-I and level-II PT tasks) x 2 (within-subject factor: congruent and incongruent trials) mixed design. More specifically, participants were randomly assigned either to the egocentric or the altercentric bias condition. In each condition, participants answered 32 trials in total: 16 level-I and 16 level-II trials. Level-I and level-II trials were presented in separate blocks, and the order of the blocks was counterbalanced between participants. For both levels, eight trials were congruent, and eight were incongruent. The order of congruent and incongruent trials within a level was randomized.

After consenting to the study, participants were presented with a brief explanation, which instructed them to indicate either what they saw on the screen (altercentric bias) or what the agent saw (egocentric bias). Then they were presented with two (one for the level-I task and one for the level-II task) familiarization trials. If participants answered the familiarization trials incorrectly, they were given feedback and had to repeat the incorrect trial until they gave the correct answer. This strategy ensured that participants understood from which perspective they were answering the question. Another aim of the familiarization trials was to train participants in the use of the "record button." This button was always located on the bottom middle of the screen and functioned as the standardized starting point for mouse trajectories. Participants had to click on this button before answering each trial; otherwise, they could not proceed with the study.

After the familiarization trials, participants were asked which of these devices they were using to move their mouse cursors: a mouse or the built-in touchpad. If they chose the mouse, they were also asked on which side their mouse was located: right or left. No participant chose left. Then participants were reminded of the instructions, and the main part of the experiment started.

In level-I trials, first, a computer-generated avatar was presented looking at a table. On the table was a sheet of paper showing two numbers/letters, which were visible to both the avatar and the participant. After one second of display, an occluder dropped from above and separated the perspectives of the avatar and the participants. The paper showed only one number/letter in level-II trials, and no occluder was dropped. The display time of the materials was compatible with the level-I trials. For both levels, after the presentation of the stimuli, the prompt question emerged on the screen with a warning for participants to click on the record button first. As soon as they clicked the record button, answer options appeared on the screen, and the recording of reaction times and mouse trajectories began. Once an answer was given, all recordings for the trial were finalized, and the procedure progressed with a new trial. In the end, participants were presented with a brief demographic questionary asking about their age, identified gender, native language, and the highest level of education completed. The study took approximately 5 minutes.

Bias Calculation & Plan of Analysis

Biases were computed using three measures: accuracy, reaction times (milliseconds), and mouse trajectories (in standardized units on 800 x 450 units of area). Accuracy was defined as the total number of correct answers. In order to see if participants made fewer errors in the congruent trials than in the incongruent trials (i.e., showed bias), separate within-subject comparisons were conducted for the two levels of perspective-taking and for egocentric and altercentric bias conditions. When the data were not normally distributed, the non-parametric alternative of the paired-sample t-test, i.e., the Wilcoxon test, was used to conduct within-subject comparisons with the accuracy measure.

The reaction times were calculated as the time between the presentation of answer options and the participant's response. After calculating the reaction time for each correctly answered individual trial, separate aggregate scores were created for congruent and incongruent trials by averaging the reaction times for these trial types. The average scores were then used to deduce biases: if the mean reaction times in the incongruent trials were more extended than in the congruent trials, this indicated bias. As a result, separate within-subject comparisons (if normally distributed, paired-sample t-tests; otherwise, Wilcoxon tests) were conducted to see if a bias existed in level-I and level-II tasks and in egocentric and altercentric bias groups.

Finally, the mouse-tracking measure corresponded to the area between participants' mouse trajectories and the direct (shortest) path to their answers (hence sometimes called the area under the curve or AUC). Similar to the reaction times, the mouse-tracking measure was also calculated for correct answers only. We capitalized on the time series and the coordinates of participants' mouse movements on a standardized 800x450 unit surface area while calculating this measure. First, the coordinates obtained from the participants' trajectories were utilized to infer their paths. These paths' start and end points were then used to calculate the direct path leading to their answers. To account for variations in trajectory duration, a standardization process was conducted by dividing the trajectories into 100 time steps and 100 coordinates using linear interpolation. Orthogonal lines were subsequently drawn from each coordinate (e.g., x1, y1) to the corresponding point on the direct path (e.g., x2, y2). The lengths of these lines were calculated using the distance formula ($d=\sqrt{((x2 - x1)^2 + (y2 - y1)^2)}$). The area between the actual trajectory and the direct path for each trial was determined by summing the lengths of all possible lines. If participants exhibited a detour towards the incorrect answer, the area between this detour and the direct route was assigned a positive value. This detour indicated a bias on the part of the participants. On the other hand, if participants deviated in the opposite direction (i.e., towards the edge of the screen), the corresponding area was assigned a negative value. Then aggregate mouse-tracking measures were calculated for congruent and incongruent trials of each level and bias condition. Consequently, separate within-subject comparisons (if normally distributed, paired-sample t-tests; otherwise, Wilcoxon tests) were performed to examine biases in both level-I and level-II tasks of egocentric and altercentric bias conditions.

Appendix C: Haskaraca, Proft, Liszkowski, & Rakoczy (2023)

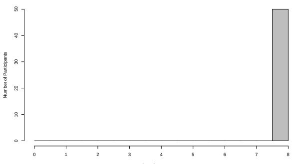
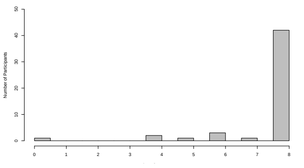
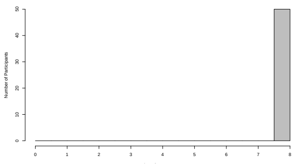
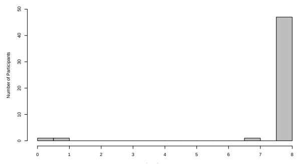
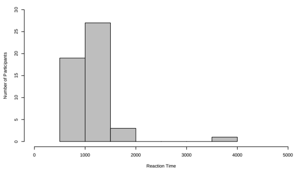
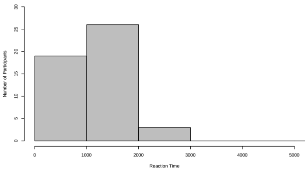
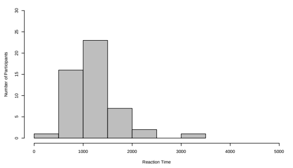
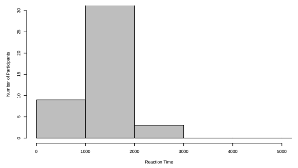
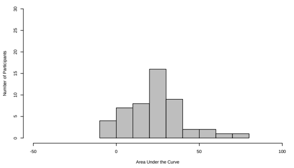
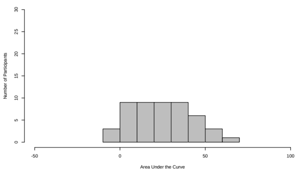
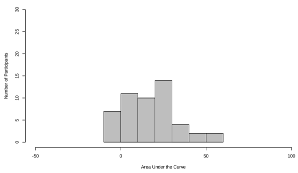
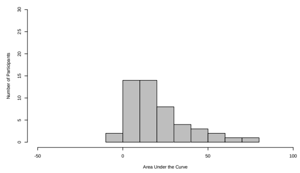
Results

Altercentric Condition: Biases in Level-I versus Level-II Perspective-Taking Tasks

The bias-revealing within-subject comparisons (of congruent versus incongruent trials) will be reported separately for each measure and for Level-I and Level-II tasks. The distributions of average accuracy, reaction time, and mouse-tracking measures are displayed in Figure 2.

Figure 2

Distribution of average measures in altercentric bias condition

	Level-I		Level-II	
	Congruent	Incongruent	Congruent	Incongruent
Accuracy				
Reaction Times				
Mouse-Tracking				

Regarding the Level-I task, the accuracy measure revealed that participants were more accurate in the congruent trials (*M*=8.0 correct answers) than incongruent trials (*M*=7.48 correct answers), *Z*=-2.536, *p*=.011. They also responded faster in the congruent trials (*M*=1205.06 ms) compared to incongruent trials (*M*=1337.95 ms) trials, *Z*=-2.606, *p*=.009. The mouse-tracking measures did not reveal a difference between the trajectories followed in the congruent (*M*=24.34 units) versus incongruent (*M*=24.55 units) trials, *Z*=-.214, *p*= .831. As to the Level-II perspective-taking, accuracy and mouse-tracking measures did not reveal a difference between congruent ($M_{Accuracy}$= 8.0 correct answers; $M_{MouseTracking}$=17.42 units) versus incongruent ($M_{Accuracy}$= 7.68 correct answers; $M_{MouseTracking}$=20.67 units) trials, *Z*=-

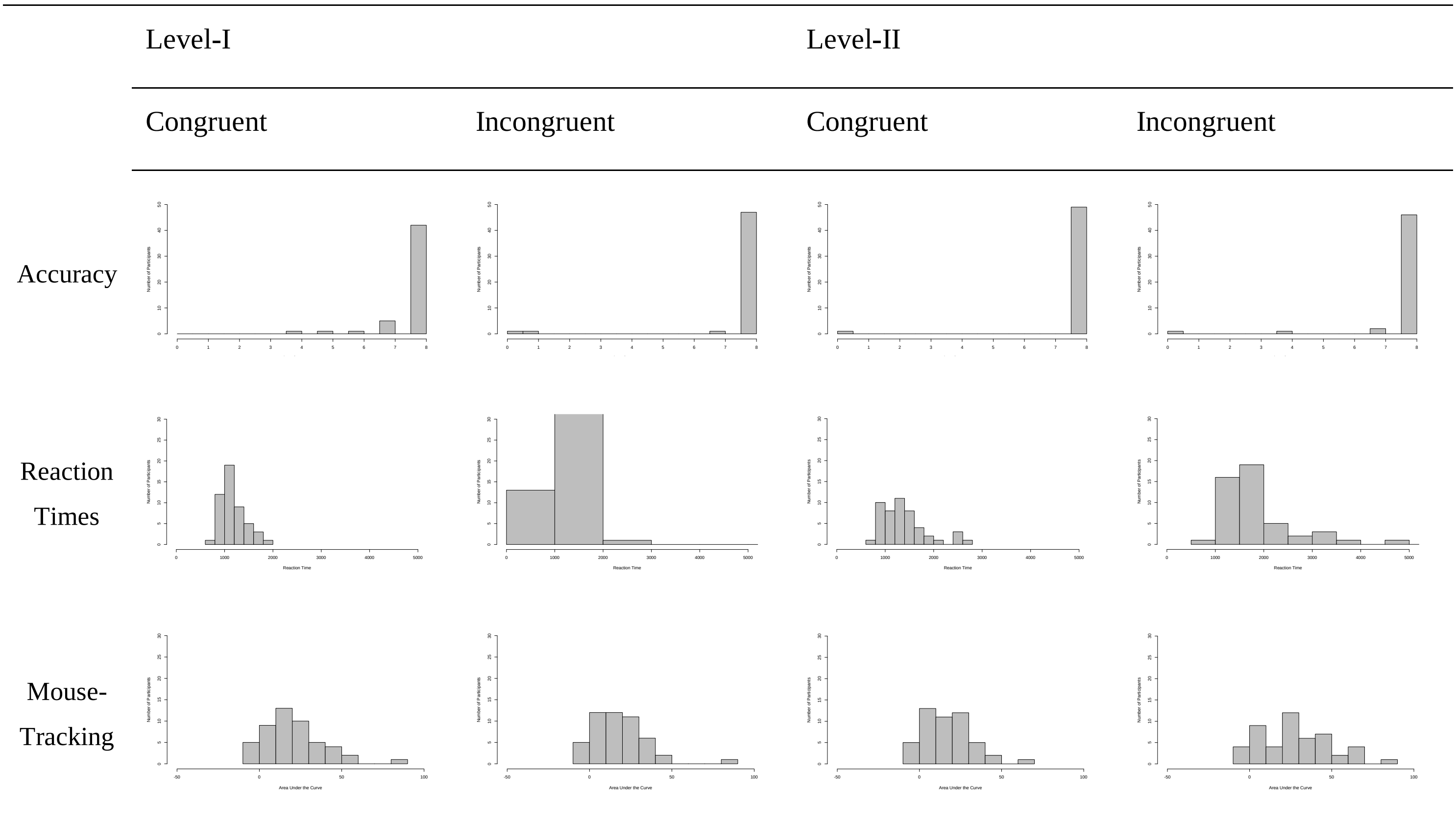
1.604, *p*=.109 and *Z*=-1.766, *p*=.077, respectively. The reaction time measure showed that participants were slower in incongruent trials (*M*=1488.59 ms) than the congruent ones (*M*=1256.98 ms), *Z*=-3.288, *p*=.001.

Egocentric Condition: Biases in Level-I versus Level-II Perspective-Taking Tasks

The bias-revealing within-subject comparisons (of congruent versus incongruent trials) will be reported separately for each measure and for Level-I and Level-II tasks. The distributions of average accuracy, reaction time, and mouse-tracking measures are displayed in Figure 3.

Figure 3

Distribution of average measures in egocentric bias condition



For the Level-I task, no difference was revealed between congruent ($M_{Accuracy}$= 7.72 correct answers; $M_{ReactionTime}$=1164.88 ms; $M_{MouseTracking}$=21.16 units) versus incongruent ($M_{Accuracy}$= 7.68 correct answers; $M_{ReactionTime}$=1279.96 ms; $M_{MouseTracking}$=18.55 units) trials by any of the measures (accuracy: *Z*=-.769, *p*=.442; reaction times: *Z*=-.124, *p*=.901; mouse trajectories: *Z*=-1.020, *p*=.308). For the Level-II task, the accuracy measure did not reveal any difference between congruent (*M*=7.68 correct answers) and incongruent trials (*M*=7.72 correct answers), $Z=-.816$, $p=.414$. However, reaction times showed that participants were slower in the incongruent trials (*M*=1813.05 ms) compared to congruent trials (*M*=1388.68 ms), $Z=-4.174$, $p<.001$. The mouse-tracking measures also revealed a difference: participants showed a bigger detour in incongruent trials (*M*=28.53 units) than in congruent trials (*M*=16.97 units), $t(47)=-4.374$, $p<.001$.

Discussion

The findings of Study 1 revealed that participants were slower and more error-prone in their own first-order judgments when the level-I task perspective-taking task involved an irrelevant agent with a divergent perspective in the scene. In the level-II trials, participants showed a similar altercentric interference effect, but only in their response times. When participants were required to take the agent's perspective, i.e., in the egocentric bias version of the task, they were not influenced by their own knowledge in the level-I trials. However, their reaction times and mouse trajectories extended in level-II trials if their own and the agent's perspective differed from each other. The findings pertaining to egocentric bias were especially challenging to interpret as the interferences occurred only in the level-II task. This result might have been caused by the demands of the level-II perspective-taking in the egocentric bias condition. More specifically, the level-II task in this condition requires participants to mentally rotate the stimulus seen on the table, whereas no mental rotation is needed in the level-I trials. It is possible that this process made the participants slower and caused them to approach the wrong location until they made their final decisions, revealing egocentric-bias-like effects in the level-II trials. In this case, the observed findings can be attributed to difficulties posed by the mental rotation requirements of the task, rather than to egocentric interferences. Considering these difficulties regarding the interpretation of the egocentric bias effects in the current task format, and since this bias has been included with explorative purposes only, the egocentric bias version was dropped in the second study.

The findings from the altercentric bias version pointed out the possibility that altercentric interferences could be found in both level-I and level-II perspective-taking. However, these findings were inconclusive as it was still unknown if the interferences revealed in Study 1 were attributable to the altercentric perception of the agent's incongruent perspective or did they come about because of low-level domain-general processes, such as attentional cues. Study 2 aimed to provide answers to these questions by testing two versions of the altercentric bias task: social versus non-social. The expectation was that if altercentric interferences were found only in the social condition, but not in the non-social condition, this would have shown that these interferences occurred due to mentalizing the agent's incongruent perspective. If similar interferences were to be found in both conditions, then those would have been attributable to the other non-mentalistic factors. The second study also investigated if the altercentric biases revealed in the first study could be replicated in the social condition of the second study and which measure(s) would continue providing consistent results in terms of the biases it revealed.

Study 2

Method

Participants

Initially, 100 English-speaking participants were recruited via Prolific for Study 2: 50 in the social condition (M_{age} = 29 years; 25 females and 25 males) and 50 in the non-social condition (M_{age} = 28 years; 25 females and 25 males). Then, in order to have enough power for mixed ANOVA analyses, additional data were collected from 110 participants: 55 in the social condition (M_{age} = 28 years; 28 females and 27 males) & 55 in the non-social condition (M_{age} = 28 years; 28 females and 27 males). All participants gave informed consent before participating and received an incentive upon their participation.

Materials, Design & Procedure, and Bias Calculation & Plan of Analysis

The methodology of Study 2 closely resembled the first study. For example, the materials used in the current study were directly taken from the altercentric bias version of Study 1. In the social condition, they were used in their existing format. In the non-social condition, only a small but crucial alteration was made to the materials: the agent was removed from the scene. Besides this difference, the non-social condition was identical to the social condition (e.g., in terms of procedure and trial structure). The design of the second study was also the same as the first study, except participants were randomly assigned to social and non-social conditions of the altercentric bias version, and egocentric bias was excluded from the groups. Procedure and bias calculations were identical in two conditions.

In terms of the analyses, first, within-subject comparisons were conducted (in the same way as the first study) in order to see if the bias patterns were similar in the social versus non-social conditions. As Study 2 also aimed to replicate the first study regarding the biases found in the social version, the sample size for the within-subject comparisons was limited to 50 participants (i.e., the sample size used in Study 1). Then with the additional data (total *N*=210), mixed ANOVA analyses were conducted separately for the three measures (i.e., accuracy, reaction times, and mouse-tracking). In these analyses, the effects of condition (social or non-social), perspective-taking level (Level-I or Level-II), and trial type (congruent versus incongruent) and their interactions were investigated.

Results

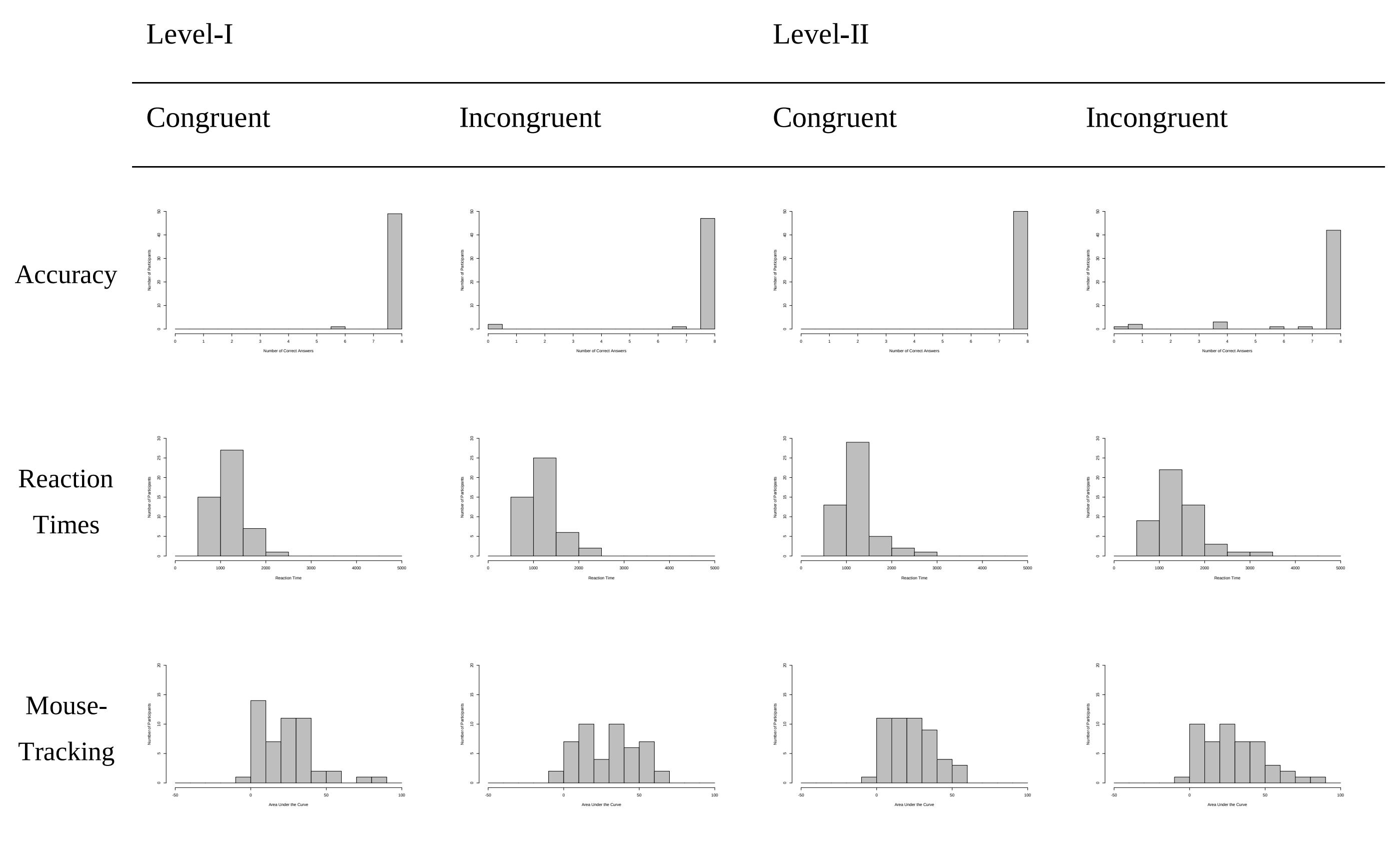
Social Condition: Biases in Level-I versus Level-II Perspective-Taking Tasks

The bias-revealing within-subject comparisons (of congruent versus incongruent trials) will be reported separately for each measure and for Level-I and Level-II tasks. The distributions of average accuracy, reaction time, and mouse-tracking measures are displayed in Figure 4.

Regarding the Level-I task, accuracy and mouse-tracking measures did not reveal a difference between congruent ($M_{Accuracy}$= 7.96 correct answers; $M_{MouseTracking}$=24.51 units) versus incongruent ($M_{Accuracy}$= 7.66 correct answers; $M_{MouseTracking}$=28.85 units) trials, Z=-1.604, p=.109 and Z=-1.795, p=.07, respectively. The reaction time measure showed that participants were slower in incongruent (M=1214.80 ms) trials than in congruent (M=1155.67 ms) trials, Z=-2.476, p=.013. In the Level-II perspective-taking task, the accuracy measure revealed that participants were more accurate in the congruent trials (M=8.0 correct answers) than the incongruent trials (M=7.26 correct answers), Z=-2.536, p=.011. They also responded faster in the congruent trials (M=1230.02 ms) compared to incongruent trials (M=1415.31 ms) trials, Z=-3.775, p<.001. The mouse-tracking measures did not reveal a difference between the trajectories followed in the congruent (M=23.00 units) versus incongruent (M=29.15 units) trials, Z=-1.885, p= .06.

Figure 4

Distribution of average measures in the social condition



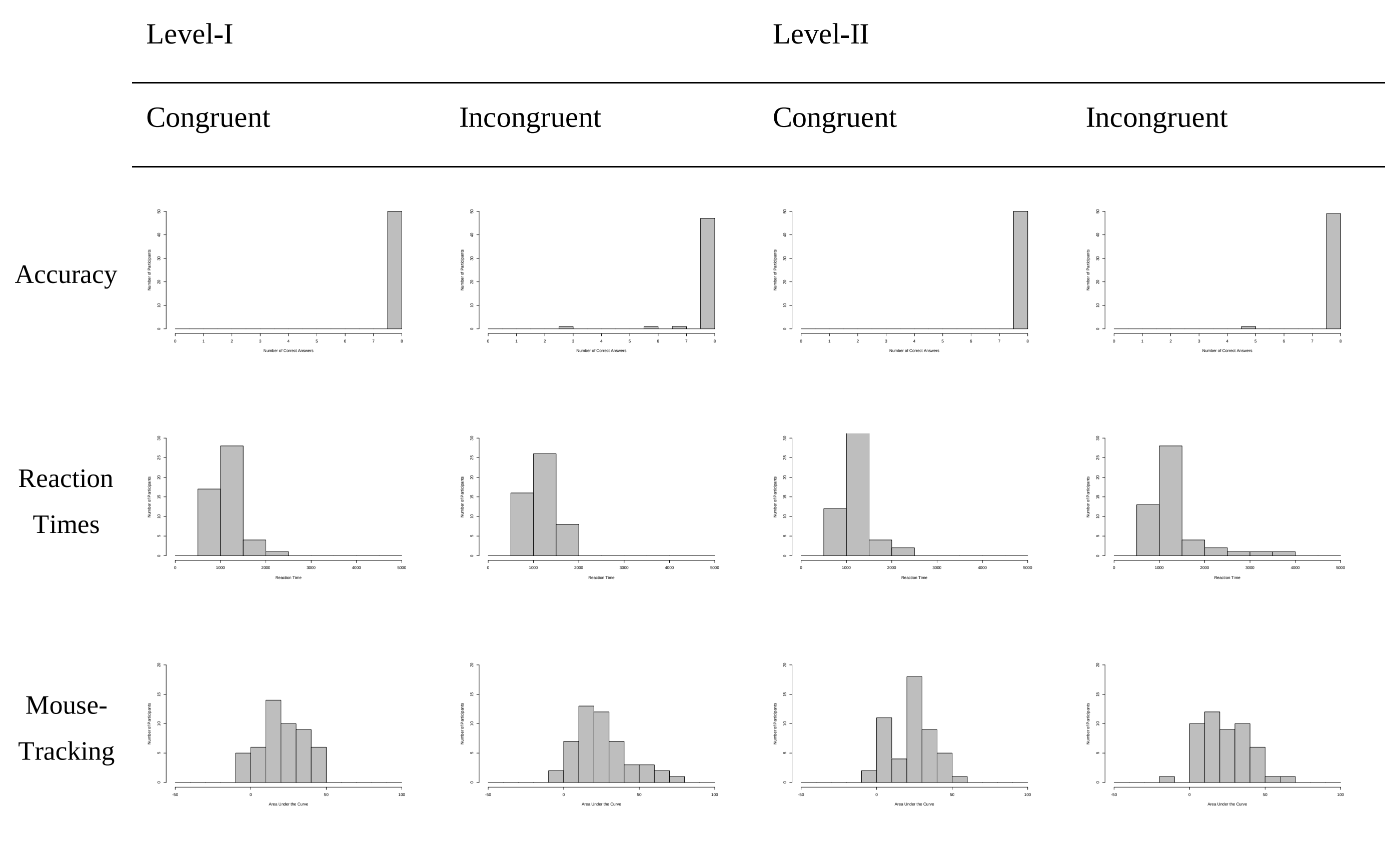
Non-Social Condition: Biases in Level-I versus Level-II Perspective-Taking Tasks

The bias-revealing within-subject comparisons (of congruent versus incongruent trials) will be reported separately for each measure and for Level-I and Level-II tasks. The distributions of average accuracy, reaction time, and mouse-tracking measures are displayed in Figure 5.

For the Level-I task, no difference was revealed between congruent ($M_{Accuracy}$= 8 correct answers; $M_{ReactionTime}$=1183.64 ms ms; $M_{MouseTracking}$=21.39 units) versus incongruent ($M_{Accuracy}$= 7.84 correct answers; $M_{ReactionTime}$=1185.93 ms; $M_{MouseTracking}$=25.02 units) trials by any of the measures (accuracy: Z=-1.604, p=.109; reaction times: Z=-.275, p=.783; mouse trajectories: Z=-1.569, p=.117). Result patterns remained the same for the Level-II task: no difference was revealed between congruent ($M_{Accuracy}$= 8 correct answers; $M_{ReactionTime}$=1215.46 ms; $M_{MouseTracking}$=22.82 units) versus incongruent ($M_{Accuracy}$= 7.94 correct answers; $M_{ReactionTime}$=1333.86 ms; $M_{MouseTracking}$=24.09 units) trials by any of the measures (accuracy: Z=-1.000, p=.317; reaction times: Z=-1.366, p=.172; mouse trajectories: Z=-.401, p=.689).

Figure 5

Distribution of average measures in the non-social condition



Social versus Non-Social Conditions

Separate Mixed 2 (group: social vs. non-social; between-subject) x 2 (level: level-I vs. level-II) x 2 (trial type: congruent vs. incongruent; within-subject) ANOVAs were conducted for the three measures used in the study. This analysis included additional data from 110 participants, constituting a sample of 210 participants. The type of device was controlled in the ANOVAs conducted for reaction times and mouse trajectories as the device had an effect on these measures ($F(1,202)=15.114$, $p<.001$ and $F(1,202)=29.228$, $p<.001$, respectively). More specifically, mouse users were faster and showed bigger detours than touchpad users.

For the accuracy measure, the analysis revealed a main effect of trial type and a main effect of group: participants were less accurate in incongruent trials than the congruent trials ($F(1,208)=16.895$, $p<.001$) and in the social condition compared to the non-social condition ($F(1,208)=6.229$, $p=.013$). No main effect of level was found ($F(1,208)=.059$, $p=.808$). There was an interaction effect of group and trial type: the difference between the number of correct answers in congruent versus incongruent trials was bigger in the social version compared to the non-social version ($F(1,208)=6.358$, $p=.012$). There was no interaction effect of level and group ($F(1,208)=.686$, $p=.409$), no interaction effect of level and trial type ($F(1,208)=.002$, $p=.961$), and no interaction effect of group, level, and trial type ($F(1,208)=1.608$, $p=.303$).

In terms of the reaction time measure, the analysis revealed a main effect of trial type and a main effect of level: participants were slower in incongruent trials than the congruent trials ($F(1,202)=4.070$, $p=.045$) and in Level-II tasks compared to the Level-I tasks ($F(1,202)=13.080$, $p<.001$). No main effect of group was found ($F(1,202)=.669$, $p=.414$). There was an interaction effect of level and trial type: the difference between the response times in congruent versus incongruent trials was bigger in Level-II tasks compared to the Level-I tasks ($F(1,202)=9.978$, $p=.002$). There was no interaction effect of level and group ($F(1,202)=2.457$, $p=.119$), no interaction effect of group and trial type ($F(1,202)=2.916$, $p=.089$), and no interaction effect of group, level, and trial type ($F(1,202)=.079$, $p=.779$).

Finally, for the mouse-tracking measure, there was a main effect of level ($F(1,202)=4.590$, $p=.033$): participants showed bigger detours in Level-I trials than in Level-II trials. There was no main effect of group ($F(1,202)=.078$, $p=.780$) and no main effect of trial type ($F(1,202)=.236$, $p=.627$). There was an interaction effect of level and trial type: the difference between the response times in congruent versus incongruent trials was bigger in Level-II tasks compared to the Level-I tasks ($F(1,202)=3.894$, $p=.05$). There was no interaction effect of level and group ($F(1,202)=.553$, $p=.458$), no interaction effect of group and trial type ($F(1,202)=.480$, $p=.489$), and no interaction effect of group, level, and trial type ($F(1,202)=.096$, $p=.758$).

Discussion

The results of Study 2 revealed that participants were slower in their own first-order judgments when the level-I task perspective-taking task involved an irrelevant agent with a divergent perspective in the scene. In the level-II trials of this social condition, participants were both slower and more-error prone when there was an agent in the scene who had a differential perspective on the stimulus. In the non-social condition, in which no agent was displayed on the screen, no interferences were revealed in level-I and level-II tasks. These results revealed that participants showed differential bias patterns in the existence of an (irrelevant) agent, indicating that the interferences were due to the presence of a divergent perspective, not due to more domain-general processes triggered by the task.

The results from the separate mixed-effect analyses with different measures as the dependent variable supported the effect of the condition to some extent. These analyses revealed main effects of trial type and condition on accuracy in the expected direction: participants made more errors in incongruent trials compared to the congruent trials, and in the social condition compared to the non-social condition. The trial type and condition also interacted in their prediction of the accuracy, with the difference between the accuracy in the congruent versus incongruent trials being bigger for the social condition compared to the non-social condition. For the reaction times, main effects of trial type and level were observed: participants were slower in incongruent trials compared to the congruent trials and in the level-II trials compared to the level-II trials. These two factors also interacted while predicting reaction times, with the difference between the reaction times in the congruent versus incongruent trials being bigger for the level-II trials compared to the level-I trials. These measures will be discussed in more detail in the *General Discussion* section.

For the mouse-tracking measure, only an unexpected main effect of level has been observed: participants showed bigger detours in Level-I trials than in Level-II trials. This finding is surprising as it hints at the possibility that participants were influenced by the existence of a second number on the screen in the level-I trials, rather than the agent's perspective on this number: regardless of the trial type and condition, participants started to move their mouse cursors in the wrong direction in level-I trials, but then they ended up at the right location (as no effect of level has been shown for accuracy). These results, together with its insensitivity in the within-subject comparisons, raised doubts about the reliability and validity of the mouse-tracking measures in this task format. Therefore, the findings from this measure will not be taken into consideration in further discussion.

General Discussion

This study investigated altercentric biases in level-I and level-II perspective-taking tasks by means of mouse-tracking measures. No altercentric bias has been revealed by mouse-tracking data in the current study; however, reaction times (in both studies) and accuracy (in different studies) measures revealed altercentric interference effects in both level-I and level-II trials. But which measure used in the current study provides a more reliable means to reveal altercentric biases: accuracy or reaction times? It is important to explore this question further before any theoretical deductions can be made from the current results, as these measures yielded differential bias patterns for level-I and level-II trials across two studies. For example, the accuracy measure revealed an altercentric bias for level-I perspective-taking trials in Study 1, but not for level-II. Instead, in Study 2, this measure revealed a bias for only level-II perspective-taking trials. Conversely, reaction times provided consistent results for both levels in the two studies. However, reaction time measures were highly affected by the device type, whereas accuracy was immune to the effect of the device. Accuracy also captured the effect of the condition better in the mixed-effect analyses, whereas reaction time measures were shown to be insensitive to the effect of the condition, although different bias patterns were revealed by within-subject comparisons in social versus non-social conditions. Also, reaction time measures might not be sensitive enough in the current study as no time limitation was introduced for participants to click on the answer. This could have caused participants to linger while answering, making the reaction times less sensitive in terms of detecting spontaneous biases. Overall, the question of which measure provides a more reliable means to deduce altercentric biases in the current study is still an open question.

Although the current studies did not provide a clear answer in terms of the optimality of the employed measures, the results still have theoretical implications as both accuracy and reaction time revealed altercentric biases for both levels at some point in the study (whether it be across studies or within studies). But what are these implications? The findings of the present studies are contrary to what is expected by dual-system accounts and what has been shown by earlier studies, namely signature limits for the implicit perspective-taking abilities (e.g., Surtees et al., 2012; Surtees et al., 2016). However, this contradiction cannot lead us directly to the possibility of unified ToM abilities as predicted by nativist accounts: although social and non-social versions of the tasks have revealed differential bias patterns, it is still possible that the interferences occurring in the current task are due to attentional/directional cues caused by the agent. For example, in the current setup, the agent's gaze possibly attracts participants' attention to the materials that are seen from the agent's perspective (or to how things are seen from that perspective), making participants slower and more error-prone in the mere presence of the agent's distracting gaze, without being interfered by this agent's perspective. This possibility then implies that the effects found in the current studies are not due to mentalizing, but because of more general attentional interferences as suggested by submentalizing accounts (Heyes, 2014). Future studies should eliminate this confounding variable from the stimuli and also establish the reliability of different measures in order to draw conclusive implications from the current task format.

Conclusion

The current study investigated the presence of altercentric biases in typical level-I versus level-II perspective-taking tasks by means of mouse-tracking measures, as well as accuracy and reaction times. Altercentric biases were revealed for both levels. However, the findings of the current study do not allow for conclusive arguments due to methodological confounds and the diminished reliability of the measures. Further empirical investigations are needed in order to establish the suitability of the current task format for making theoretical inferences about implicit ToM.

References

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review, 116*(4), 953-970. https://doi.org/10.1037/a0016923

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., & 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

Cole, G. G., Atkinson, M., Le, A. T., & Smith, D. T. (2016). Do humans spontaneously take the perspective of others?. *Acta Psychologica, 164*, 165-168. https://doi.org/10.1016/j.actpsy.2016.01.007

Conway, J. R., Lee, D., Ojaghi, M., Catmur, C., & Bird, G. (2017). Submentalizing or mentalizing in a Level 1 perspective-taking task: A cloak and goggles test. *Journal of Experimental Psychology: Human Perception and Performance, 43*(3), 454–465. https://doi.org/10.1037/xhp0000319

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

Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the Level 1–Level 2 distinction. *Developmental Psychology*, *17*(1), 99–103. https://doi.org/10.1037/0012-1649.17.1.99

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

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

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

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

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

Marshall, J., Gollwitzer, A., & Santos, L. R. (2018). Does altercentric interference rely on mentalizing?: Results from two level-1 perspective-taking tasks. *PloS One, 13*(3), e0194101. https://doi.org/10.1371/journal.pone.0194101

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

Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology, 5*(2), 125–137. https://doi.org/10.1111/j.2044- 835x.1987.tb01048.x

Perner, J., & 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

Poulin-Dubois, D., Rakoczy, H., Burnside, K., Crivello, C., Dörrenberg, S., Edwards, K., ... & Ruffman, T. (2018). Do infants understand false beliefs? We don't know yet–A commentary on Baillargeon, Buttelmann and Southgate's commentary. *Cognitive Development, 48*, 302-315. https://doi.org/10.1016/j.cogdev.2018.09.005

Powell, L. J., Hobbs, K., Bardis, A., Carey, S., & Saxe, R. (2018). Replications of implicit theory of mind tasks with varying representational demands. *Cognitive Development, 46*, 40-50. https://doi.org/10.1016/j.cogdev.2017.10.004

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

Santiesteban, I., Catmur, C., Hopkins, S. C., Bird, G., & Heyes, C. (2014). Avatars and arrows: Implicit mentalizing or domain-general processing? *Journal of Experimental Psychology: Human Perception and Performance, 40*, 929–937. https://doi.org/10.1037/a0035175

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

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

Sommerville, J. A., Bernstein, D. M., & Meltzoff, A. N. (2013). Measuring beliefs in centimeters: Private knowledge biases preschoolers' and adults' representation of others' beliefs. *Child Development*, *84*(6), 1846-1854. https://doi.org/10.1111/cdev.12110

Southgate, V. (2020). Are infants altercentric? The other and the self in early social cognition. *Psychological Review, 127*(4), 505–523. https://doi.org/10.1037/rev0000182

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

Surtees, A. D., 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

Surtees, A., Samson, D., & Apperly, I. (2016). Unintentional perspective-taking calculates whether something is seen, but not how it is seen. *Cognition*, *148*, 97–105. https://doi.org/10.1016/j.cognition.2015.12.010

Van der Wel, R. P., Sebanz, N., & Knoblich, G. (2014). Do people automatically track others' beliefs? Evidence from a continuous measure. *Cognition*, *130*(1), 128-133. https://doi.org/10.1016/j.cognition.2013.10.004

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

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

Curriculum Vitae

Curriculum Vitae can be found in the printed version of this dissertation.