From interest contagion to perspective sharing

How social attention affects children's performance in false-belief tasks

Falck, Andreas

2016

Document Version:
Publisher's PDF, also known as Version of record

From interest contagion to perspective sharing

From interest contagion to perspective sharing

How social attention affects children's performance in false-belief tasks

Andreas Falck

LUND
UNIVERSITY

DOCTORAL DISSERTATION
by due permission of the Faculty of Social Science, Lund University, Sweden.
To be defended at Edens Hörsal, Lund, on the 11th of June 2016 at 10:00.

Faculty opponent
Professor Pierre Jacob, Institut Jean Nicod, École Normale Supérieure, Paris

Organization LUND UNIVERSITY	Document name: DOCTORAL DISSERTATION
	Date of issue
Author(s): Andreas Falck	Sponsoring organization

Title and subtitle: From interest contagion to perspective sharing: how social attention affects children's performance in false-belief tasks

Abstract

A large amount of socio-cognitive research has been devoted to questions about different perspectives – how do we understand that other people can have their own perspectives on reality. This is typically studied with so called false-belief-tasks (FBTs), which are experimental tasks designed to tap into an ability to understand that other people may have beliefs differing from our own (perspective-taking). Children show some evidence of such ability between 15 months and 4 years of age, depending on how they are tested, which have led to controversy regarding the nature of this ability. Another branch of research, often called social attention, has been concerned with how we tend to become influenced by other people's attention, so that we attend to what they attend to. Notably, when attending the same information as someone else we are likely to think about the same information and converge on similar beliefs, i.e. share a perspective. In this thesis, I asked whether the explanatory burden of presumptive theories of perspective-taking can be reduced when taking into account how social attention helps us share perspectives. I address this question in three papers, in collaboration with colleagues.

Paper I is a theoretical investigation of what is needed in order to share a perspective. The developmentally earliest claims of children showing sensitivity to others' false beliefs are currently aimed at 7 months of age. In this paper I and colleagues challenge this claim, by arguing that the results are better explained by the low-level mechanism of interest contagion, a human tendency to attend to what others attend to without reflecting on this fact, together with recognition memory. Whereas the experiments with 7 month old infants do not show evidence of perspective-taking, they do provide evidence about how perspectives can be shared, which in turn can help us understand perspective-taking.

Paper II is an empirical study of what happens when 3-4 year old children watch a film in which a boy named Maxi comes to have a false belief about his toy's location. They watch either with an experimenter attending the story together with them, or alone while the experimenter does something else. In the end, they are subject to a FBT – the experimenter asks them about where they expect Maxi to go to fetch his toy. We find that children are more likely to report that Maxi will search where he thinks the toy is, rather than where it is really (thus passing the test), if they have watched the story together with the experimenter. We also tested children's memory of the story, with mixed results. We therefore suggest that whatever helps children succeed on the FBT is specific to reasoning about others' perspectives.

Paper III is a follow-up study using eye-tracking in order to investigate 1) whether children watch the story differently when watching together with the experimenter, and 2) whether they look in advance to where Maxi is expected to go even if they give the wrong answer to the verbal question. We find that children who watch together with an experimenter scan the scene more (more fixations), and that most of the children who fail the FBT question from the experimenter still looks in advance to where Maxi is expected to look, hence showing some sensitivity to the fact that Maxi is mistaken about his toy's location.

Taken together, these studies suggest that sharing a perspective with someone scaffolds understanding of how perspectives may differ. This fits a theory of perspective-taking in which people are contextual cues to situations which was previously shared with them, which I argue explains a lot of the presently existing data parsimoniously.

Key words: attention, social cognition, intersubjectivity, false-belief tasks, contagion

Classification system and/or index terms (if any)

Supplementary bibliographical information	Language: English
ISSN and key title	ISBN: 978-91-7623-852-3
Recipient's notes	Number of pages 165 / Price
	Security classification

From interest contagion to perspective sharing

How social attention affects children's performance in false-belief tasks

Andreas Falck
andreas.falck@psy.lu.se

LUND
UNIVERSITY

Contact: andreas.falck@psy.lu.se

Alternative: falck.andreas@gmail.com

Faculty of Social Science | Department of Psychology

ISBN 978-91-7623-852-3 (tryck)
ISBN 978-91-7623-853-0 (pdf)

Printed in Sweden by Media-Tryck, Lund University
Lund 2016

Till Frej, vi tänkte så lika men ändå så annorlunda

Content

Abstract

A large amount of socio-cognitive research has been devoted to questions about different perspectives – how do we understand that other people can have their own perspectives on reality. This is typically studied with so called false-belief-tasks (FBTs), which are experimental tasks designed to tap into an ability to understand that other people may have beliefs differing from our own (perspective-taking). Children show some evidence of such ability between 15 months and 4 years of age, depending on how they are tested, which have led to controversy regarding the nature of this ability. Another branch of research, often called social attention, has been concerned with how we tend to become influenced by other people's attention, so that we attend to what they attend to. Notably, when attending the same information as someone else we are likely to think about the same information and converge on similar beliefs, i.e. share a perspective. In this thesis, I asked whether the explanatory burden of presumptive theories of perspective-taking can be reduced when taking into account how social attention helps us share perspectives. I address this question in three papers, in collaboration with colleagues.

Paper I is a theoretical investigation of what is needed in order to share a perspective. The developmentally earliest claims of children showing sensitivity to others' false beliefs are currently aimed at 7 months of age. In this paper I and colleagues challenge this claim, by arguing that the results are better explained by the low-level mechanism of interest contagion, a human tendency to attend to what others attend to without reflecting on this fact, together with recognition memory. Whereas the experiments with 7 month old infants do not show evidence of perspective-taking, they do provide evidence about how perspectives can be shared, which in turn can help us understand perspective-taking.

Paper II is an empirical study of what happens when 3-4 year old children watch a film in which a boy named Maxi comes to have a false belief about his toy's location. They watch either with an experimenter attending the story together with them, or alone while the experimenter does something else. In the end, they are subject to a FBT – the experimenter asks them about where they expect Maxi to go to fetch his toy. We find that children are more likely to report that Maxi will search where he thinks the toy is, rather than where it is really (thus passing the test), if they have watched the story together with the experimenter. We also tested

children's memory of the story, with mixed results. We therefore suggest that whatever helps children succeed on the FBT is specific to reasoning about others' perspectives.

Paper III is a follow-up study using eye-tracking in order to investigate 1) whether children watch the story differently when watching together with the experimenter, and 2) whether they look in advance to where Maxi is expected to go even if they give the wrong answer to the verbal question. We find that children who watch together with an experimenter scan the scene more (more fixations), and that most of the children who fail the FBT question from the experimenter still looks in advance to where Maxi is expected to look, hence showing some sensitivity to the fact that Maxi is mistaken about his toy's location.

Taken together, these studies suggest that sharing a perspective with someone scaffolds understanding of how perspectives may differ. This fits a theory of perspective-taking in which people are contextual cues to situations which was previously shared with them, which I argue explains a lot of the presently existing data parsimoniously.

Sammanfattning

Denna avhandling syftar till att koppla ihop två forskningsfält inom social utvecklingspsykologi. En stor del av de senaste 30 årens forskning på hur vi fungerar socialt har handlat om hur vi tar andras perspektiv, d.v.s. förstår och hanterar det faktum att andra människor kan ha uppfattningar om världen som skiljer sig från våra egna. Detta undersöks normalt i så kallade false-belief-uppgifter (eng. false-belief task, FBT), i vilka en testdeltagare (oftast ett barn i förskoleåldern) får bevittna en situation där en annan person missar kritisk information. Oftast handlar det om att personen vill ha tag på någon sak som nyligen blivit flyttad, utan att personen vet detta. Ett intressant fynd i detta fält är att barnen klarar sig olika bra på dessa tester beroende på hur de får visa sin förståelse. Barn har ofta svårt att svara på en direkt fråga om var en annan person kommer att försöka hitta något som barnet, men inte den andra personen, vet har blivit flyttat. Treåringar svarar ofta att personen letar där saken faktiskt finns, även om personen inte kan veta detta. Det är först mellan fyra och fem år som barn normalt börjar svara rätt på frågan. Däremot, om barnen blir testade på ett sätt där de inte behöver kommunicera svaret till en annan person, så verkar mycket yngre barn kunna agera som att de förstår konsekvenserna av att ha felaktig information. Det har bland annat visats att 15 månaders bebisar uttrycker förvåning om andra agerar som om de hade information de inte har tillgång till, och att 2-åringar tittar i förväg mot platser där en person förväntas gå givet den information den personen har tillgång till. Dessa motstridiga fynd har lett till en debatt i forskningsfältet där många är oense om vad förmågan att förstå avvikande perspektiv egentligen består i.

I avhandlingen tar jag avstamp i en annan del av forskningen om sociala förmågor, nämligen det som kallas social uppmärksamhet. I detta fält studeras bland annat hur vi följer andras blickriktning, och på så vis tenderar uppmärksamma det som andra uppmärksammar. En intressant konsekvens av detta, som jag argumenterar är negligerat i forskningen om hur vi förstår andras perspektiv, är att om vi tittar på samma saker som andra så tenderar vi dela deras perspektiv. Vi kan alltså, med relativt självgående psykologiska färdigheter (blickföljande är så gott som automatiskt), komma att få samma perspektiv som andra. Detta kräver till skillnad från många FBT ingen medveten reflektion. Frågan som ställs i denna avhandling är om detta faktum om hur perspektiv kan delas kan utnyttjas för att förklara hur vi

relaterar till andras avvikande perspektiv. Jag argumenterar för att frågan om hur vi förstår andras perspektiv delvis kan reduceras till frågan om hur vi skiljer på perspektiv som är delade, och perspektiv som inte är delade. Detta berör jag i tre studier, vilka inbördes är ganska olika i termer av mått, frågeställningar och metoder.

Studie I är en teoretisk redogörelse för hur social uppmärksamhet leder till delning av perspektiv. Jag med kollegor utgår från publicerade experimentella resultat som anses visa att mycket små barn (7 månader) förstår andras avvikande perspektiv. Vi argumenterar för att dessa resultat bättre förklaras av att *intresse smittar*, så att vi blir intresserade av det som intresserar andra utan att tänka på det. Tillsammans med igenkänningsminne, som alla har från födseln, så kan intressesmitta förklara småttingarnas beteenden utan att det betyder att de hanterar andras avvikande perspektiv. Detta innebär dock inte att vi förklarar bort småbarnens beteende som ointressant – då intressesmitta är en social funktion i sig, så utgör den i sig en bas för både barns och vuxnas förmåga att fungera socialt. Vi argumenterar därför för att intressesmitta kan användas som avstamp för att förklara senare utveckling av förmågor att förstå andras avvikande perspektiv.

Studie II är en empirisk undersökning bestående av två experiment. Här tittar vi på hur lite större barn (3-4 år) förstår andras perspektiv, med en typ av FBT som är lämplig för denna ålder. Barnen tittar på en film om lille Maxi, som får sin favoritleksak (ett flygplan) flyttad medan han är ute och leker. När Maxi ska hämta sitt flygplan frågar experimentatorn vart barnet tror att han kommer att gå – dit där den finns eller dit där han lämnade den. Vi ser att barnen som tittat på filmen tillsammans med experimentatorn (och följaktligen delar ett perspektiv) oftare svarar att Maxi går dit där han lämnade flygplanet, d.v.s. dessa barn klarar oftare testet. För att undersöka varför de klarade sig bättre testade vi också hur mycket barnen kom ihåg från historien, och huruvida deras minne skilde sig mellan de barn som tittade tillsammans med experimentatorn och de som tittade själva. Vi fann att minnet inte skilde sig generellt mellan villkoren, men vissa saker var barnen som tittade tillsammans mer benägna att prata om efter filmen. Det var framför allt det faktum att Maxis flygplan hade blivit flyttad när han var ute och lekte som barnen som tittat tillsammans med experimentledaren oftare kom ihåg, eller åtminstone oftare fann värt att berätta om. En viktig detalj som skiljer vårt FBT från övriga är att vi låtit filmen förmedla berättelsen, medan i de flesta andra verbala FBT är det berättaren som ställer frågan om var Maxi förväntas leta. Vi argumenterar för att vår variant är en mer naturlig situation, då den som ställer frågan inte uppenbart vet slutet på berättelsen, vilket skulle vara fallet om berättaren frågar. Barnen i vårt experiment har därför ingen anledning att tro att de blir testade, vilket de har i de vanliga FBT-experimenten. Detta är en potentiell förbättring validiteten i testproceduren, relativt tidigare varianter på FBT.

I studie III följde vi upp experimenten från studie II med ögonrörelsemätning, med nya grupper av barn, detta hade två mål. Vi ville se 1) om barnen tittade annorlunda på filmen om de tittar tillsammans jämfört med att titta själva, och 2) om de tittar mot den plats där Maxi borde leta även om de med ord svarar den plats där flygplanet faktiskt finns. Vi fann att barn som tittar tillsammans med experimentledaren tittar mer på filmen (fler fixeringar) och proportionerligt mer på de kritiska händelser som leder till att Maxis flygplan blir flyttat. Vi fann också att trots att de flesta barnen i denna studie svarade att Maxi skulle leta där flygplanet fanns i verkligheten (alltså fel svar), så tittade de flesta barnen mot rätt plats (där Maxi tror att flygplanet är) i det ögonblick berättelsen gjorde klart att Maxi ville ha sitt flygplan. Detta är i linje med övriga resultat i litteraturen, som tyder på att barn förstår implikationerna av andras avvikande perspektiv om de får reagera spontant eller interagera med andra, men att de har svårt att verbalisera denna kunskap som svar på en direkt fråga.

Tillsammans tyder dessa studier på att på att när barn delar perspektiv med andra, så har de bättre möjlighet att förstå implikationerna av situationer när perspektiv inte är delade. De empiriska studierna (II och III) visar vidare att när barn ska resonera i samtal om konsekvenserna en tredje persons avvikande perspektiv, så underlättar det om samtalspartnern har samma perspektiv som barnet på de kritiska händelserna. Det verkar alltså som att barn under fyra års ålder framför allt har svårt att hantera när flera agenter samtidigt (experimentledaren och Maxi i vårt fall) har perspektiv som de inte delar med barnet. Utöver detta ser vi att barn är mer intresserade av en berättelse om de följer den tillsammans med en vuxen, vilket tycks leda till starkare fokus på den viktiga informationen i berättelsen.

Acknowledgements

The list of whom I'd like to thank is extensive. First and foremost, my PhD supervisors – my main supervisor **Magnus Lindgren**; for encouraging my ideas, helping me believe in them; having patience with extensive Socratic dialogue about questions perhaps not always very interesting to you, nevertheless seeing where to cut the Gordic knots in the end; showing that there are always better ways in which to fail; academic mentoring in general; and countless over-lunch discussions about cooking, literature, and Scottish politics (to name a few subjects). My co-supervisor **Ingar Brinck**; for constant request for clarification and theoretical sharpening, helping me see inconsistencies even in what I thought was the most trivial statements (which of course wasn't); great feedback at various stages; and for encouraging my work all the way since I took your pragmatics and philosophy course in 2004. My collaborator-turned-co-supervisor **Elia Psouni**; for organising data collections and making them happen (sometimes too fast!), essentially enabling the acquisition of the data included in this thesis – a great effort. Thanks to you all!

Thanks to the MSc students involved in the two empirical papers: **Leni Boström & Martin Persson, Lisa K Sidén & Maria Wallin, Janna Gunnäng & Mathilde Hyddmark**. And thanks to **Riley Sims** for sharing your knowledge on eye-tracking children. Thanks to all the **pre-schools**, **children**, and **parents** who made the empirical studies possible. Also, the people who have reviewed my thesis work at various times deserve huge thanks – **Paula Rubio-Fernández** for inspiring feedback on my final draft a few months ago; and **Fredrik Björklund** and **Annika Wallin** who provided feedback on an even earlier outline of this thesis, a few years back.

Then of course my colleauges, with whom I have shared a lot both at and outside of work – my officemates at different time periods, **Arvid Erlandsson**, **Robin Hellerstedt**, **Sabina Gušić** and **Rugile Sokolova**. My "neighbours" **David Marcusson-Clavertz**, **Per Davidson** and **Rachel Elsinga**, and the Monday Club – **Martin Hassler Hallstedt**, **Anna Lindqvist** and **Ellinor Schad**. Not to forget the members of the **CCL (Cognition, Communication and Learning) Linneaus environment**, the **CogComLab**, and the **Swedish Cognitive Science Society** (SweCog). All of these people I have had great benefit of having around during my PhD time.

Additionally, both CCL and SweCog (and thus ultimately the **Swedish Research Council**) deserve thanks for the funding of various excursions and conferences.

Then of course my **Mom** and **Dad**, always nearby despite the physical distance, have provided great support.

The people involved in my finding my way to do research in psychology deserve special thanks. Tracing back – My grandmother **Ing-Marie**, who passed away a little more than a year ago, the first to show me the thrill of pursuing knowledge, when I at the age of 9 or so found a fish in the Lule river which I could not identify (it turned out it was a lamprey). My **Dad**, being a licenced psychologist, naturally always encouraged my interests in the subject – but also always warned me, *"make sure you don't get stuck in academia!"* I haven't heeded that advice so far – perhaps it was mostly thanks to **Gunnar Windahl**, who inspired me to question fictious dichotomies such as mind/body and science/humanities, and encouraged me to go to Lund in the first place. And **Petra Björne**, then a PhD student supervising my master's thesis at LUCS, was instrumental in shaping my interest in what became my research subject.

Finally, I have been lucky or competent enough to have a life partner who already did the PhD marathon, knowing exactly what it means to sometimes be immersed in work 14 hours a day, and always available within hours even at those occasions when we spend time in different countries, thanks to the invention of instant messaging. **Johanna**, without you I might not have gotten this far.

Dalby 2016-05-01

List of papers

This dissertation is based on the following three papers, referred throughout the text by their roman numerals.

I.

Falck, A., Brinck, I., & Lindgren, M. (2014). Interest contagion in violation-of-expectation-based false-belief tasks. *Frontiers in Psychology*, *5*(JAN), 1–5. doi:10.3389/fpsyg.2014.00023

II.

Psouni, E., Falck, A., Boström, L., Persson, M., Sidén, L. & Wallin, M. Together I can! Joint attention boosts 3-4 year olds' performance in a verbal false-belief test. *Manuscript in preparation.*

III.

Falck, A., Gunnäng, J., Hyddmark, M., Psouni, E. Three-year-olds' processing of a false-belief story - evidence from a visual world paradigm. *Manuscript in preparation.*

Introduction

Human social functioning, far from being manifestations of a monolithic theory of other people's minds, is a mixed bag of cognitive and emotive abilities. Today we possess a reasonable mechanistic understanding of many basic functions such as social attention and emotional contagion. We also have a firm research tradition on how we consciously reason about how other people's actions are influenced by what they are taken to believe. What are still missing from the picture are the intermediate levels, a way to bridge basic non-cognitive social functioning with more advanced cognitive abilities such as the understanding of others' beliefs. In my PhD research presented here, I suggest and investigate, theoretically and empirically, directions for bridging this gap.

I present a theoretical account of how interest in objects and events can be contagious, in that we become interested in what interests others without necessarily reflecting on the other person as the source of our interest. This reasoning can be applied to non-verbal false-belief experiments, offering an explanation of infants' performance in social-cognitive experiments without presupposing that they possess an adult-like theory of mind (Paper I). I also present and discuss data from preschool children, suggesting that three to four year old children's performance of standard false-belief tasks is improved if a caregiver shows – and consequently shares – interest in the same task (Paper II). Then I further expand on this finding, via an eye-tracking investigation of pre-schoolers' attention when they follow a story in which the protagonist's perspective is a key plot feature, while manipulating the caregiver's attentional behaviour meanwhile (Paper III).

In the following, I will review and discuss evidence on how people come to share information with others. Then I will review and discuss the evidence on how we act in situations when other people's beliefs differ from ours, that is, when information is not shared. Then I will briefly remark on methods, before introducing the included papers. Finally I will discuss the importance and implications of the present findings.

When information is shared

Many interpersonal matters are about sharing. With sharing I refer to two or more people having something in common – in (including but not restricted to) perception, action, or thought. First, I discuss how attentional targets can be shared, that is, concurrently attending to the same physical object or event. Then I build on sharing of attentional targets to establishing that a perspective can be shared.

Social attention

Social attention is the phenomenon that we tend to attend to others, and to where they attend. Essentially, it includes two attentive behaviours: attention to people, and attention to the objects of other people's attention. We are good at finding people in our visual environment. Not only do we attend faces and eyes automatically and often without reflection, also people's actions are easily detectable. Finding people by their activities is efficient and relies on quite scarce visual information, such as the movements of key joints (Johansson, 1973). We are also prone to attend to the targets of other's object-directed actions, including but not limited to what they look at. A paradigmatic mechanism behind finding the target of other's attention is gaze following, the tendency to react in a reflexive way on shifts of other's gaze. Gaze following in humans is closely connected to the contrast differences found in the visual image of the human eye, something that is evident from several laboratory experiments (e.g. Friesen & Kingstone, 1998; for a review see Birmingham & Kingstone, 2009). The strong contrast difference is perhaps one of the reasons why infants attend others' eyes early in development – parents' eyes are presumably among the most salient sources of contrast in the infant's visual environment. While infants react to eye contact from birth (Farroni, Csibra, Simion & Johnson, 2002), they start to follow gaze after eye contact at 3 months and without preceding eye contact at 6 months (Brinck, 2008). The effects of being attended, such as when establishing eye contact, has recently gained much interest (Schilbach, Timmermans et al. 2012; Risko, Richardson & Kingstone, 2016). These situations are the first instances in which infants have the possibility to experience others' attention, by being an attentional object rather

than a subject (Reddy, 2009). These encounters pave the way for infants' following attention to other objects.

Attention following in humans does however not require visual access to conspecifics' eyes. With or without gaze, we readily parse directed actions, and the targets of others actions is often the object of our attention. For example, the literature on motor attention suggests homologues both on the behavioural and neural level (Rushworth, Johansen-Berg et al. 2003). In adults, mismatching head and gaze cues slows responses compared to congruent cues (Langton, 2000), suggesting that head direction is an attentional cue in its own right. Brooks and Meltzoff (2005) showed that at 9 months, infants follow head turns regardless of whether the turning head has closed or open eyes, while at 10 months they start to take into account whether the eyes are open, not following a head with closed eyes. This suggests that not only gaze shifts, but also other aspects of directed behaviour directs attention already at a young age. Yu and Smith (2013) showed that when playing with toys together with a parent, 12-month-olds seldom look at the parents face – instead they look at the manipulated objects, no matter whether it was currently manipulated by the child or the parent (or both). While this is perhaps not surprising at face value, Yu and Smith also made the interesting observation that gaze to the manipulated object was largely synchronised between parent and child, so that they both were looking at the same object at the same time as either of them was handling the object. Thus, it appears that attention as conveyed by physical manipulation and gaze can be redundant with each other, and either can be a basis for sharing attention, at least from the first birthday onward. Pierno, Becchio and colleagues (2006) show evidence of motor-related activity related to grasping in subjects scanned with fMRI, when they observe a person looking to a graspable object compared to the same person looking the other direction. Becchio, Bertone and Castiello (2008) review a range of evidence suggesting that looking at objects as a consequence of attentional cueing by gaze alters the perception of these objects. Consequently, sharing an attentional object does not only imply having access to the same information, the perception of the target is also dependent on pragmatic information from how the information is attended by the interacting parties (see Brinck, 2014). This is not to imply, however, that they fixate the same object all the time, only that they attend to it concurrently at some times throughout the same episode, and that the object is not hidden or removed from their visual fields.

Modes of sharing information

As triadic situations may involve both attention to the shared object and the interacting people attending to each other, I will discuss how different types of sharing plays out in such situations. In the following, I will borrow some

terminology from Brinck (2001; 2004; 2008; Zlatev, Brinck & Andrén, 2008). Brinck defines 'attention contact', as focused attention on each other's attentional state (Zlatev, et al., 2008). A paradigmatic case of this phenomenon is eye contact, in the sense of looking at each other's eyes to the extent that both parties are aware of looking at each other. Consequently, if eye-contact is followed by an attentional shift by both parties to an external object, both will be aware not only that both are looking at the object, but also of each other's attention to the object. Eye contact therefore makes the mutual attention to each other visible to the agents, and consequently, after eye contact the agents' attention to the object is not only simultaneous, but shared. However, not all episodes of sharing attention involve eye contact. Brinck (2004; 2008) distinguishes two types of sharing of attentional objects between two agents which I here call Alice and Bob;

Individual sharing, consisting of Alice reacting to Bob's attention (consciously or not), so that Alice and Bob come to have the same attentional target, without Bob attending to Alice's attention,

Mutual sharing, consisting in Alice and Bob attending to the same target and both attending to each other's attention to the target.

Mutual sharing implies not only sharing an attentional object, but also the sharing of the attention to the object. This includes but is not limited to sharing as a consequence of eye contact, in which Alice seeks eye contact with Bob, then directs attention to a third target which Bob subsequently follows attention to. This latter phenomenon would correspond to what is often in the literature called joint attention (JA; Brinck, 2001).

Here it is evident also that we can be influenced by others' attention without being aware of it. Individual sharing does not require awareness of the sharing, but applies for any level of appreciation of the source of the attentional influence. That is, individual sharing can be a case of *contagion* – a social influence which is not necessarily acknowledged by the receiver. In paper I of this thesis I and colleagues define *interest contagion*[1] as follows:

> The tendency to automatically find interest in events, situations, or objects that others show interest in, as conveyed by behaviourally manifest attention, and, consequently, to converge with respect to interest. (*Paper I, p. 3*)

[1] We borrowed the structure of the definition from Hatfield, Cacioppo & Rhapson (1994), who define *emotional contagion* as "The tendency to automatically mimic and synchronize expressions, vocalizations, postures, and movements with those of another person's and, consequently, to converge emotionally". In their theory as in our, behaviour is the basis of influence and the influence consists in contagious transfer of behaviour.

Interest contagion implies (individual) sharing of an object. Note that the definition is lean on the level of attention needed *to* the other person. It suffices to attend to the person showing interest in an event, situation or object, in order to pick up the direction of their attention. Therefore it follows that individual sharing of attention can happen as a consequence of interest contagion. If Bob is attending to the target of Alice's attention, as a consequence of Alice's attentional behaviour drawing Bob's attention, then Bob is individually sharing Alice's attention to the object. Thus, we argue in paper I that instances of individual sharing of a third object can happen without explicit knowledge of the other person being the source of our object-directed interest.

Sharing perspectives

When engaging in a joint activity together with others, our overt attention to the shared space lets us influence each other without noticing, typically so that we attend to the same information. Attending to the same information in a scene renders it likely that we will perceive the scene in a more similar way compared to attending to different information, thus, it implies that our perspectives on the scene converge. If Alice and Bob attend to the cat on the mat, they will be likely to subsequently remember the cat on the mat, want to cuddle it, worry about it scratching the mat, and so on. This aligns with Wilby (2010), who argues that in order to share a belief, we need only to attend to the same information at the same time (see also Heyes, 2014a). As we have seen, and which Wilby (2010) acknowledges, we do not need to attend to each other's attention in order to share an attentional object. While we may well be aware of the fact that we are sharing information, this is not a prerequisite to share perspectives. Rather, a prerequisite is that our attentional processes are linked to each other – it suffices that my attention affects yours (and the other way around) in order to share attention to an object, for this I do not have to process your attention as an object of cognition (Wilby, 2011; see also Michael, 2011). Whether individual or mutual, sharing of an attentional object therefore implies causal links between attentional processes in two or more subjects. Awareness of these causal links is possible but not necessary for qualifying the sharing of attention.

It explains why doing things together is typically not harder compared to doing things in isolation, even though there are other people involved which may require coordination (see also Butterfill, 2013). In many everyday triadic interactions peers attend to each other more or less, while typically focusing on the shared object. This suggests a structured alternation between individual and mutual sharing during many everyday triadic situations. The time course of mutual sharing depends on the participants' span of awareness and their ability to reciprocate. For example, if we go to a movie theatre together, we would have

quite few occasions where we would reciprocate or comment the joint view, directing each other's attention. But we would still see it as an episode of mutual sharing, and for most people, watching together will nevertheless enhance the experience, possibly leading to an increased arousal and attentional vigilance. As long as a mutually shared episode is being experienced in a way distinct from other forms of sharing, there is an experiential basis for being aware of the episode as mutually shared. Consequently, it does not rely on a constant awareness of the other person's attention. The account of sharing information presented here is cognitively lean on the level of attention needed to the people with which we share. Whereas we do attend to our conspecifics in a mutually shared situation, we do not necessarily cognize on them or become aware of attending to them, neither do we cognize on the fact that we are influenced by their attentional behaviour. In other words, the mechanisms discussed here does not rely on us having attributed attentional capacities to other people, as attribution is a cognitive operation taking another person and a capacity as arguments.

Whereas it has been argued that attributing attention to others is a cognitive precursor to attributing beliefs to them (e.g. Tomasello, 1995), it is not needed for interest contagion to happen, or for mutual sharing of the target of attention. Thus, throughout an episode of sharing (individually or mutually) attention to an event or an object, we are also sharing a perspective, at least in a basic sense. This does not imply, however, that every aspect of the perspective is shared. Alice might like cats whereas Bob might be allergic. Whereas attentional targets can afford different actions depending on individual preferences or differences in learning history, the spatial organisation of things will generally be agreed upon within a shared space, leaving Alice and Bob with a shared perspective on the spatial layout (e.g. a shared belief that the cat is on the mat). Next we turn to situations in which shared perspectives do not obtain.

When information is not shared

So far I have discussed situations in which people share attentional targets, and thus have the same *perspective on* the shared scene (in a basic sense). An important question in (developmental and) social psychology is how we handle situations in which perspectives are not shared, that is, acting socially in situations where our perceptual history differs from those of our peers.

Reasoning about others' beliefs

In a seminal experiment, Wimmer and Perner (1983) investigated how children aged 3-9 years reason about what other people know. They showed children a picture-based story, in which Maxi[2], 5 years old, helped his mother put away a piece of chocolate in a blue cupboard. However, when Maxi was outside playing, his mother moved the chocolate to the green cupboard. Then, the story goes, Maxi comes back, remembers the chocolate, and wants to eat a piece of it. Where will he go and look for it? Children were asked this question, as to see whether they expected Maxi to look for his chocolate where he left it. None of the children below 4 years of age answered this alternative, instead they claimed that he was going to look in the green cupboard, where the chocolate actually was. This was the first structured investigation of children's understanding of how people can be mistaken about reality, acting instead on their *false beliefs*. Subsequent research showed a stable pattern – children below 4 years of age generally answered that Maxi will look where the chocolate is, while older children typically answered that he would look where he left it. Notably, children's answers are typically given with confidence – they tend to answer either the old location or the new one, and persist in their choices both before and after the 4-year-old limit.

Baron-Cohen, Leslie and Frith (1985) made a puppet adaptation of Maxi, the so called Sally-Anne task, replicating Wimmer's and Perner's findings. They also tested the same experiment with children who were diagnosed with autism, finding that these children typically did not succeed on this task until at earliest 5 years of

[2] I.e. "Little Max" in German.

age (or later, or never, depending of severity of autistic symptoms[3]). Subsequent research showed that the ability to answer "correctly" on FBT questions is related to a range of cognitive factors, including language ability (Astington, 2000; for a meta-analysis see Milligan, Astington & Dack, 2007) and general intelligence (Bradmetz, 1998). Moreover, false-belief tasks are passed later by deaf children not exposed to sign language compared to those who are (Peterson & Siegal, 1995), suggesting a causal role of language development. Executive functioning is a collective term for abilities of attentional control and self-regulation. These abilities have also shown stable links to performance on classical FBTs (Carlson & Moses, 2001; Frye, Zelazo & Palfai, 1995), though as we will see, this relation is confined to a specific type of FBTs.

In a meta-analysis, Wellman, Cross and Watson (2001) included a large amount of task manipulations, examining changes to the original procedure which could help children to answer correctly to the FB question. They found in general little effect of the task manipulations which had been tried to that date. These manipulations were generally focused on changing the wording of the FB question, making the protagonist's false belief more salient by making their false belief arise due to apparent deception, and increasing the level of interaction with the child. Letting the child partake in the story and hide the critical object from the protagonist led to improved performance, suggesting that the degree to which the child is involved in the story plays a role. Making the deception obvious by other means also helps, as in the "Naughty Snakey" paradigm (Bowler, Briskman & Grice 1999). In this study the object was moved not by the experimenter but by an antagonist agent named "Naughty Snakey", which was presented as intending to "spoil our game". In their study children performed better when "Naughty Snakey" moved the object compared to the experimenter who asked the question. This suggests that task changes that help the child identify with the protagonist, the experimenter, or both can be beneficial to their performance on the false-belief question test. Or alternatively, exaggerating the deceptive component of the task may make clearer that the protagonist misses information. As we shall see, making any larger improvement to the children's performance involves more radical changes to the paradigm than had been tried in 2001.

[3] In the latest clinical consensus definition of autism as of DSM-5, also what was formerly called Asperger's syndrome is included in autism. Baron-Cohen's et al. studies used an earlier definition as an inclusion criterion, which does not include Asperger's syndrome.

Observing information-constrained actions

All the above variants of the FBT relied on verbal reports. Famously, Onishi and Baillargeon (2005) showed that 15 month old infants looked longer at an adult actor reaching for an object where it actually was, compared to when reaching where the actor last had seen it. This clever experiment employed what was already an established technique in infant research – infants look longer at situations they are either not familiar with or are not expecting. In a range of experiments, this has been employed together with simple illusionist-style tricks in order to pinpoint ages where infants start assuming that objects persist while being temporarily hidden, that they re-appear behind a barrier according to a continuous trajectory, and similar physical phenomena (for a review see Baillargeon, 2008). Consequently, an actor could put a toy in one of two boxes, and did so repeatedly, until the infant got used to seeing the actor reach into a box containing a toy (thus losing interest and looking away after a short time). Then, the actor hid behind a curtain and the toy was moved by a magnet, suggesting that it moved by itself. Eight groups of children then saw different completions of the sequence – four in which the actor saw the toy move to one of the boxes, and four in which the toy moved without the actor seeing were it ended up, thus giving the actor a false belief about its location. Half of the children then saw the actor reach for the box in which the toy actually was, and half saw the actor reach into the one where it wasn't. For a full sketch of the design, see Figure 1. The infants who saw the actor reach into the box in which she last saw the toy looked for a shorter time than those seeing her reach into the other box, regardless of the toy's real location. This suggests that children expect people to look for objects where they last saw or interacted with them, essentially suggesting some form of sensitivity that people can have false beliefs. Later studies built on this paradigm replicated the finding in various versions, using both live and puppet-style agents (Scott & Baillargeon, 2009; Song et al. 2008; Surian et al. 2007).

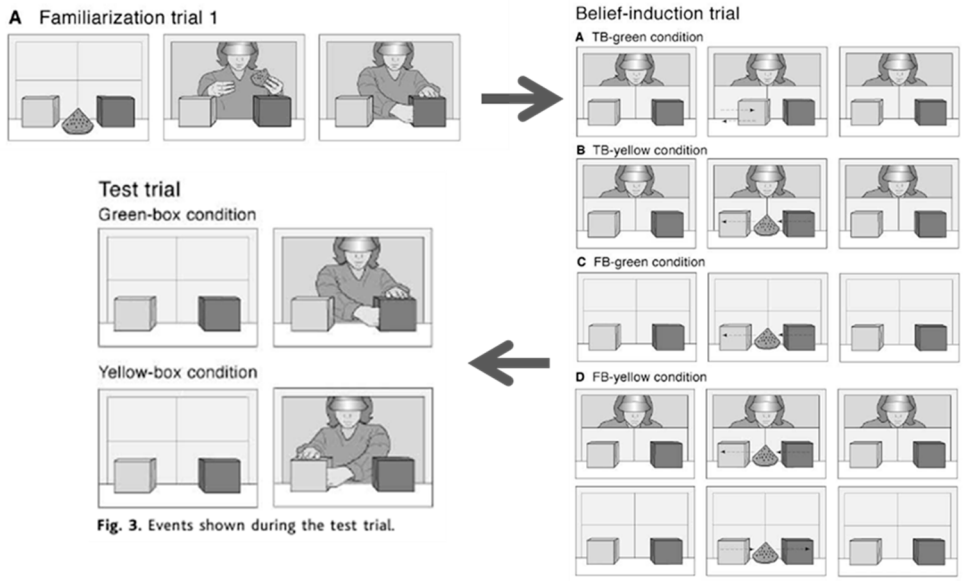

Figure 1. Structure of Onishi & Baillargeon's (2005) violation-of-expectation study on false belief understanding in infants. Reprinted with permission from AAAS.

In another study using a different experimental paradigm, eye-tracking was used to monitor 25-month old children's looking in advance to an actor's expected search for an object (Southgate, Senju & Csibra, 2007). The scene was similar to that of Onishi and Baillargeon (2005), though the measure was different, and the object was removed (not moved) by a puppet. At the agents return, toddlers did typically look in advance towards the place where the agent saw the object last. In contrast to the Onishi and Baillargeon study, the authors did not include a true-belief condition (since the toy always disappeared from the scene), but they had (false-belief) conditions in which the ball ended up in either of the locations before disappearing from the scene. The authors argue that they did not want to induce a bias by having a true location, and that because of this the two FB conditions would suffice.

This study design might require a stronger sense of understanding beliefs, since the anticipation of an agent's action was tested, and not only the reaction to that action. Another important feature in this study was that the agent did not disappear from the scene when missing the information - instead she simply redirected her attention to an event behind her back (a phone ringing), turning her head away from the moving of the object. This suggests that it is the attention of the agent which guides the toddlers' expectations, rather than the agent's presence or absence.

The so called *anticipatory looking* measure used by Southgate and colleagues was not new in 2007. Clements and Perner (1994) used a similar measure to assess 29-month old (and older) children's anticipatory looking to an agent acting on their false belief. Their story was a more traditional FB story with many verbal narrative elements, and they also followed up the anticipatory looking measure with a verbal question. Crucially, the anticipatory looking was not elicited by the agent's returning per se as in the Southgate and colleauges (2007) study, but by a verbal prompt "*I wonder where he [the protagonist, a puppet mouse] will look?*". Interestingly, they found looking patterns indicative of an understanding of the agent's false belief in children from 35 month onwards, but in the age group 29-34 months they did not – only three of the 11 children in this group looked where the agent last saw the object. Southgate et al. discusses this discrepancy and suggests it could be the "I wonder…" prompt which led the children to think of the true location of the object, perhaps by misinterpreting it as a question directed to the child (Southgate et al. 2007). However, this has not been tested experimentally.

Helping someone trying to act on erroneous information

The above reviewed studies all dealt with stimuli which were simply observed by the infant, only addressing how infants react to watching others act. However, social encounters typically imply interaction (Carpendale & Lewis, 2006; Gallagher, 2001, 2004; Reddy, 2009; Reddy & Morris, 2004). For this reason the question has been raised whether infants react properly to others' false beliefs also in interaction. Buttelmann, Carpenter and Tomasello (2009) put infants and children in a situation where they could help an adult experimenter (called E2 in their study) to open a box. In one condition ("true belief"), E2 left a toy in a box, and then watched it being moved to another by a second experimenter (E1 in the Buttelmann et al. study). The child was then invited to lock both boxes with a key, which the child could keep. E2 then tried to open the empty box, and the child was invited to help by unlocking the box. Most infants and pre-schoolers were helpful in opening the empty box. In another condition, "false belief", E2 left the room for a short while. This time E1 again moved the toy, but without E2 seeing this, in an overtly sneaky way to involve the child in playing a trick. The boxes were locked again by the child. Then, E2 came back and tried to open the empty box, which was also the box in which E2 had left the toy. Now, however, most 18 month old (but not 16 month old) infants did not open the box which E2 tried to open, but instead opened the other box (containing the toy). This suggested that the infant understood that E2 wanted his toy and did not know that it had been moved.

However, this interpretation was recently challenged by Allen (2015), who noted that E2's false belief was confounded with the trick-playing inherent in the false-belief condition. He used Buttelmann and colleauges' false belief condition together with two other conditions, both in which E2 had missed seeing the moving of the toy, with pre-schoolers at ages 3-5:

In the *clairvoyance* condition, E2 tried to open the box which actually contained the toy (which he would have thought was empty), instead of the empty box (which he would have thought contained the toy). This was the same as the false-belief condition in every other respect.

In the *hands full* condition, the scenario was changed so that E2 would want an empty box. Apparently a toy-thief named Oscar-the-Grouch was in the vicinity, stealing all toys, so E1, E2 and the child had to hide all their toys. Now, they filled the first box with toys so that it was just possible to close. Then E2 left the room in order to fetch some more toys to hide, during this pause E1 suggested that they move the already hidden toys to the other box, which they did. E2, coming back with hands full of toys, now tries to open the box which was previously empty (though now being full of toys).

Allen found that children in the "clairvoyance" condition tended to open the box with the toy, even though the actions of E2 suggested that he was really after the empty box (assuming that he acted on his false belief). This suggests that children's actions were not driven by an interpretation of E2's false belief per se, but rather by a tendency to want to help finding the salient object (see also Helming, Strickland & Jacob, 2014). So what about the *hands full* condition, in which E2's intention to actually open the empty box was given a rationale? In this condition results were mixed, with a spread over age. Most 3-year-olds opened the box with the toys (suggesting that the typical action of 'retrieving toys' were still more salient), while 5-year-olds tended to open the empty box (consistent with understanding the thief-scenario and E2's false belief simultaneously).

Allen suggests, that the helping paradigm of Buttelmann et al. (2009) really taps into a sensitivity to the typical desire in a given scenario rather than an appreciation of false beliefs per se. Interestingly, in the Buttelmann et al. (2009) control condition 18-month-olds and 2.5-year olds had no problem opening the empty box, without an overall story suggesting that this would be a relevant target. Nevertheless, when such a story clearly existed in Allen's 'hands full' scenario, three-year-olds apparently had problems. It seems that the social context explanation provided by Allen would not suffice to explain this discrepancy. We will get back to this question in the discussion section of this thesis.

Experimenter interaction and tracking of perspectives

Furthermore, it has been shown that the interaction with the experimenter can affect children's performance in different ways, suggesting that some of the differences in performance may be reflecting consequences of the experimenter's actions on the child's attention. In Rubio-Fernández & Geurts' (2013) Duplo task, children took part in a verbally administered false-belief story in which the experimenter interacted with the children, testing the children by letting them enact the story by themselves. In this task 3-year olds passed the task in great numbers, however, there were also children not passing – raising the question of how this paradigm relate to the infant studies showing expectations of belief-guided actions.

The Duplo task did not, as it seems, improve the performances of children around 3 years of age unconditionally, instead producing results more in line with Clements and Perner's study. This suggests that even though the lack of experimenter interaction simplified the test phase, children still had a comparably harder time with successfully predicting the protagonist's action. Thus, the difference between the Duplo task and the infant studies becomes interesting in itself. While the three-year-olds' performance in the Duplo task broke down when the Duplo figure disappeared from the scene, rather than was overtly looking away, the infant studies have been employing a large number of ways in which the agent's belief becomes invalidated. Rubio-Fernández and Geurts (2013) discussed this result in terms of tracking of perspectives – suggesting that both loss of visual access to the protagonist and experimenter's actions can disrupt children's contact with the protagonist's attention (perspective), and thus explaining their results parsimoniously.

Theoretical intermission – what does it mean to understand other perspectives?

Before going on to the studies I will present some brief theoretical remarks.

Characterisations of different types of tasks

Several definitions have been suggested in order to differentiate FBT measures, in ways which corresponds to how they can be mechanistically explained. Clements and Perner (1994) discussed anticipatory looking as an implicit measure, as opposed to the standard FBT requiring the child to convey their expectation to an experimenter by pointing or telling. Similarly[4], Onishi and Baillargeon (2005) introduced their VoE task as a non-verbal alternative to the standard verbal task, discussing the differences in terms of task demands of the verbalisation of the response (see also Leslie, 2005). There is a problem with dividing the tasks as verbal and non-verbal, as the notion leaves unclear whether *verbal* refers to the way the information is conveyed to the child, or to the way the child responds. As there exists tasks leading to successful performance in infants in which some information is conveyed verbally (Buttelmann et al., 2009; Song & Baillargeon, 2008), as well as non-verbal response tasks which seems hard enough for older children (Call & Tomasello, 1999), the verbal component cannot be the only relevant difference.

Baillargeon et al. (2010) focused on how the measure is obtained from the child, pointing out that it is asked for in the standard task, whereas it is read from the children's behaviour in the newer tasks used with infants. They refer to the standard tasks as "elicited response tasks", and the newer non-verbal tasks as "spontaneous-response tasks". The study by Buttelmann et al. (2009) does not fit

[4] Whereas the description of the task differences used by Clements and Perner (1994) and Onishi and Baillargeon (2005) can be construed as similar, their respective theoretical interpretations of the consequences of the difference cannot. This is outside the scope of this point, however.

this pattern, as the child is responding to the experimenter trying to fetch their object, so the response can be argued to be elicited. Consequently, they refer to it as an indirect-elicited response task, while arguing that similar mechanistic demands are put on the child as in the spontaneous-response tasks. Also the Duplo task can be viewed as an indirect-elicited-response task, as the experimenter invited the child to play with the Duplo figure by the means of verbal interaction (Rubio-Fernández & Geurts, 2013). Importantly, the spontaneous/elicited response framework denotes the tasks by the differences in how the child responds. However, it also depends on how the response is obtained from the child. If the child is asked to direct the answer to an experimenter, young children (~ 3 years) fail, whereas if the child is asked to help the protagonist – by engaging in a joint activity with the protagonist (Buttelmann et al., 2009), or by identifying with the protagonist (Rubio-Fernández & Geurts, 2013), they succeed. This suggests that the important distinction may not be so much how the child expresses the expectation, but whom the child is acting towards when responding, and which relation they have.

Critique of the false-belief task's scope

One problem with the FBT, and its role in understanding beliefs, is that it only concerns beliefs which are false. The reason for this is methodological, as true beliefs are confounded with reality. The problem is the following: if we ask how children respond to what others believe, and the information about what they believe can be obtained some other way, then we do not know that we are looking at how they understand or react to beliefs. Therefore, to see whether Alice expects Bob to act according to his experience but not according to reality we must construct a situation where the only reasonable action by Bob depends on what Bob has experienced, and where that experience was of a state which is no longer the case. The FBT is thus an instrument to get at beliefs, which uses falseness as a way of making sure that it is a belief we are looking at. Moreover, the canonical concept of a belief implies that it is Bob's belief, not just any possible state of the world remembered or imagined by Alice.

However, there is no principled reason that our everyday concept of a belief needs to be the best starting point for explaining socio-cognitive abilities. Fabricius and Kahlil (2003; see also Hedger & Fabricius, 2010) tested children in a standard FBT with the addition of a true-belief condition, constructed by a small adjustment to what happened with the object: when the agent was away, the object was moved but then moved back to its old location, so that the agent had an outdated belief which was still true. They found that 4-year olds expected the agent to be mistaken

in their search, as if they expected outdated information to lead to error unconditionally. These results suffer from a lack of replications, however, Fiebich (2013) discusses this with some directions regarding why. She argues based on these tasks that others' beliefs are understood by different mechanisms depending on whether they are true or false. But in FBTs such as those reviewed here, true beliefs consistently originate from a context shared between the story protagonist and the observer, whereas false beliefs always refer to a context which was not shared. This suggests that the basis of such mechanisms may have something to do with whether information is shared.

The meta-representation debate

The most persistent debate regarding infants' performance on the various non-verbal FBTs regard whether they meta-represent other's beliefs – that is, entertain parallel representations of other people's view of the world. For example, Baillargeon et al. (2010) hold that infants have an innate modular capacity for representing other's beliefs, which is active when observing others act (explaining what happens in the non-verbal FBTs). Modular in this context means informationally encapsulated (Fodor, 1983), that is, being impenetrable by cognitive process. A cognitive process can thus only work on belief representations as wholes, including meta-representations (i.e. representations of what others' represent). Consequently, meta-representing another person's perspective involves keeping in mind how the current situation looks to them.

Under this assumption, three-year-olds' failures on the verbal FBT is typically explained as a limitation of their executive functioning when verbalising the response (See also Leslie, 2005; Onishi & Baillargeon, 2005). This view has some support from the fact that verbal FBT performance correlates with executive ability (Carlson & Moses, 2001; Frye, Zelazo & Palfai, 1995). However, several alternative accounts exists – e.g. that older children meta-represent but not infants (e.g. Perner & Roessler, 2012), that infants and adults form behaviour-based representations which only approximate beliefs (Apperly & Butterfill, 2009; Butterfill & Apperly, 2013), or that representations are not needed at all (Gallagher, 2001; 2004; Gallagher & Hutto, 2008). I will however not go into this debate in detail, rather, I intend to focus on a specific prediction which the meta-representational account seems to entail:

If we automatically and effortlessly form representations of others beliefs when interacting with them, and entertain them in parallel (at least as long as the overall context is the same), we should function similarly in situations where others' beliefs are shared as when they are not shared.

This follows from the defining aspect of modularity that modular representations are monolithic, i.e. they can only be processed as wholes by a cognitive system. If we already have a mechanism to represent the world from another person's point of view, and this is modular in the canonical sense of the word, there is no reason why other's actions should be processed differently when they have the same information as I have, compared to when they lack access to information which I have. Consequently, the conjecture that beliefs are modular seems more or less taken for granted in many accounts of perspective-taking. Moreover, the consequences are similar of some non-metarepresentational accounts, for example Butterfill and Apperly's (2013) minimal theory-of-mind approach. They argue that whereas infants do not represent what others believe, they represent situations in which an agent has previously seen an object. However, such a representation still describes a situation which looks in a certain way to another person.

The consequences of the conjecture that beliefs are modular are perhaps best demonstrated with the help of an example. As we shall look deeper into in paper I, an experiment by Kovács, Téglas and Endress (2010) purported to show evidence of a modular meta-representational capacity in 7-month-olds. They presented infants with scenes which either showed a ball disappearing from a scene when an animated agent (a blue Smurf) was watching, or a very similar scene without the agent. Infants were less surprised (as measured by looking time) when the ball was shown (unexpectedly as it had disappeared) in the conditions where the agent had seen it disappear compared to the conditions in which the agent did not see it disappear. The authors suggested that this showed that the infant was representing the agent's belief about the ball in parallel with their own, so that the agent's belief influenced the infant's own belief.

However, the infants did not react differently when the ball was revealed while the agent was watching compared to while he was not watching, suggesting that the infants never attributed the belief to the agent (experiment 2), or at least did not form an expectation about how the agent would act (see Kampis, Somogyi et al. 2013 for data on this phenomenon in 10-month-olds). Kovács (2015) argues that the content of another's belief can be separated from the owner of the belief, so that one can represent that *someone* believes that the cat is on the mat, or that Bob believes that *X*, which would solve the problem that the 'beliefs' in their study was never sensitive to *who* was doing the believing. By this modification, the content of the belief is still regarded as a representation of the world being in a specific state, still maintained in parallel with our own view of the world. The question is whether this is the most parsimonious account of the evidence, or whether the need for dissecting beliefs this way in order to explain the youngest infants' performance compromises modularity.

True belief bias, sharing bias, or experimenter distraction?

For a long time, the assumption was that children who fail verbal FBTs do so because they cannot inhibit their true knowledge of the toy's location (e.g. Carlson & Moses, 2001). This, however, has recently been questioned on multiple grounds. Rubio-Fernández (2015) showed that answering negation questions ("where isn't Sally's doll?") lead adult participants to respond slower to subsequent references to the same location. A false belief question "Where will Sally look for her Doll", however, did not show lingering inhibition in the same way, but rather facilitation of the same location. Instead, Rubio-Fernández (2015; Rubio-Fernández & Geurts, 2013) suggests that the main reason 3 year old children fail verbal FBTs is that they fail to keep track of what the protagonist has seen or not, as soon as the object is mentioned by the experimenter. This would potentially explain the discrepancy between infants' performance on non-verbal FBTs and older children's performance on verbal FBTs. The results from Rubio-Fernández & Geurts (2013; 2015) highlights another aspect, that the experimenter's actions are powerful cues to information in the FBT, also when these actions lead the children to irrelevant or misleading information.

As attention is involved in determining one's current perspective, and the experimenter's attention influence that of the child, it is a fair question to which extent children must reflect on the experimenter's perspective. This view has gained considerable momentum in recent years (Carruthers, 2013; Helming, Strickland & Jacob, 2014; Gallagher 2015). Thus, the response which must be successfully inhibited is rather one created by the experimenter's actions, which entails an idiosyncratic situation which does not resemble any relevant real-world tasks. This would still explain the relation between executive functioning and verbal FBT performance. This question, however, transcends the representation debate. While Helming et al. (2014) suggest that the tendency to be misled by the experimenter is due to the child automatically and compulsively representing the experimenter's perspective (which in turn represents the true location of the object), Gallagher (2015) argues that the child's response is due to sharing a perspective on the true location with the experimenter, in virtue of the child's and adults' connected attentional processes.

In the meanwhile, I suspected that the very fact that the experimenter knows the answer to the FBT question suggests that the question itself is perceived as weird by the child. Why ask about something you obviously know about? This is implied in the standard FBT not only because the experimenter is attending to the same information as the child, but also because the experimenter narrates the story. Consequently, sharing a perspective on a false-belief story does not only imply

that the child's attention to information is connected to that of the adults', it also regulates when the child needs to reflect on the experimenter's perspective or not. This model, however, stands in contrast to the meta-representational account of young children's performances. If we assume that the meta-representation account of perspectives holds, we would expect the perspectives of others to be equally accessible to us whether they are shared or not shared.

The aim of this thesis

The general lesson learned from experimental and theoretical work on sharing attention and interest on one hand, and different versions of false-belief tasks on the other, is that the former is easy while the second is comparably hard. In three studies I explored the impact of sharing interest on situations in which others' mistaken beliefs matters. The main argument is that if we consider the implications of sharing information by other means than attributing mental states to others, we can explain how children behave in a range of situations in which they apparently understand the implications of others' having false beliefs.

Methodological remarks

Here follows some discussion and elaboration on the methods used in the different works, with emphasis on what is not already described in the included articles.

A note on parsimony: two views

The first paper (I) is a theoretical piece in which the argument is built on parsimony. That is, trying to find the explanation with the fewest unfounded assumptions, still explaining all the current data. Citing the conflicting accounts of infants' FBT performance given by Onishi and Baillargeon (2005) on one hand, and Perner and Ruffman (2005) on the other, Apperly and Butterfill (2009, p. 954) note that claims of parsimony is often associated with different conflicting accounts. Moreover, we shall see that we in paper I handle parsimony somewhat differently from those mentioned above. This suggests that parsimony isn't theoretically neutral in relation to questions of the nature of infants' FBT performances. Specifically, the plausibility of the hypothesis that infants meta-represent other beliefs depends on how parsimony is understood. I suggest two readings of the principle, which I take as representing the different takes on parsimony compatible with the different positions in the representation debate.

Rule parsimony. The number of functional connections or rules needed in order to explain a specific behaviour, such as *"people look for things where they think they will find them"* or *"people look for things where they last saw them"*.

Process parsimony. The number of processes or mechanisms needed in order to explain a specific behaviour, e.g. attention, imagination, recognition, recall.

In the *rule parsimony* interpretation, the functional information processing structure is what parsimony operates upon. The goal then is to simplify the logical structure of the cognitive machinery. This way a belief processing module (cf. Leslie, 1987) is a parsimonious assumption compared to the notion that we need a rule for each context. Put simply, it takes more rules to link stimulus and behaviour without the assumption that people have minds than with it. The rule parsimony interpretation allows only redundancy on the functional level. This was the main line of Chomsky's (1957) argument against behaviourism – he suggested that for a mind to work through stimulus-response associations only, more rules

are needed compared to a cognitive system which could represent. This is how Onishi and Baillargeon (2005) argue their interpretation of their results to be parsimonious: *"[t]o explain [...] the present results, it is more parsimonious to assume that infants attribute to others beliefs that can be shaped and updated by multiple sources of information than to assume that infants form an extensive series of superficial expectations linking different perceptions to different actions."* (ibid, p. 257). On this account, the belief representation becomes a third variable which alleviates the need for a larger number of functional rules. Importantly, what mechanisms are in place or not is of less importance in this reading, as cognition is modelled on its informational constraints.

Process parsimony, on the other hand, concerns the cognitive process rather than the functional links upon which it operates. In this reading, assuming a cognitive module for representing the beliefs of others is unparsimonious if the same function can be realised only with reference to processes which already have an empirical foundation. This is the reading of parsimony used in Paper I of this thesis. Reducing outcomes of implicit false-belief tasks to sharing attention together with contextual reinstating of agent-object relations thus enables the belief-processing module to be omitted. Instead, the model refers only to processes having separate empirical support in the proper developmental age range.

Thus, the difference between these readings of parsimony corresponds to a difference in how cognition is conceived. If cognition, including social cognition, is essentially a set of rules connecting stimuli to responses, then the rule parsimony interpretation would apply, whereas the process parsimony view would be of limited use. This is because rules are domain-constrained – a rule about how a person would act depending on a specific belief, or a rule on how a person would act depending on a specific history of encounters (allegedly more complex than the former), would be hard to recruit to other domains of knowledge than person-knowledge. On the other hand, if we allow cognition to emerge from non-cognitive processes (Barsalou, Brezeal & Smith, 2007; Mahner & Bunge, 1999; Kosslyn & Koening, 1993), the existence of these processes is what needs explanation. Thus, it suffices to explain which processes are needed for an ability to emerge and how they can co-exist, something which can be grounded in existing evidence from other types of tasks.

The Maxi false-belief film

For use in studies II and III a film-based false-belief story was developed[5]. The story was made in a style similar to the popular Swedish children's show Alfons Åberg[6], with still pictures rather than animation, and a narrator describing the events. See Paper II for the full story with pictures and text. The story introduced Maxi, who has recently got an airplane which is now his favourite toy. Maxi joins his friend to play outside, and puts the airplane away in a red coffin to the left in his room. While Maxi is out, his dad enters the room, cleaning and vacuuming. At a certain point, Maxi's dad is picking up the airplane from the coffin, moving it to the dresser on the other side of the room! This event is exaggerated by the narrator (panel 9) in an excited tone. Then Maxi comes back, depicted standing in the middle of the room (panel 4), and it is announced that he wants to play with his airplane again. The narrator then announces that Maxi is going to do something, by saying that "Maxi takes[7]…" [here the film stops]. So, apparently the internet connection was lost and the film froze, preventing us from knowing what happens next! Here, the experimenter steps in and asks what the child thinks that Maxi is going to do: *"Where do you think he will look for his airplane?"*. During the entire part when Maxi is outside, he is shown in a distance outside the window (panels 7-9). This was in order to keep him visible to the child during the entire hiding event, showing that he does not know what Dad is doing, in line with the observation by Rubio-Fernández & Geurts (2013) that seeing the protagonist during the critical hiding event improves children's responses. We reasoned that this feature would be especially important, since several of our hypotheses pertained to the ability to follow Maxi's perspective.

Memory for Maxi

The Memory for Maxi scale was developed[8] based on Nepsy-II "Memory for story" (Korkman, Kirk & Kemp, 2007). It consisted of 13 items addressing the different elements of the Maxi story film, about which the child was interviewed

[5] By Elia Psouni, Leni Boström, Martin Persson and myself, based on an idea by myself and EP. MP created all the pictures. Furthermore, the film was piloted in LB's and MP's MSc. thesis.

[6] English: Alfie Atkins, see http://www.alfons.se/english/about-alfie/ (as of 2016-02-07)

[7] "Maxi tar…" in Swedish, where "tar" could in a very specific (and unusual) sense mean "takes [his airplane from the coffin]", but it is more often used in a broader meaning of the type "takes [any type of action]". So in this particular context it signals that Maxi is going to act upon his desire to play with the airplane, but not specifically that he is going to fetch something in a specific place.

[8] By Elia Psouni, Lisa K. Sidén and Maria Wallin. The Memory for Maxi scale was piloted as part of LKS's and MW's Lic. Psych. thesis.

after the FBT question. It included items both relating to the critical features of the story, such that "Maxi went out to play" (after putting away his airplane), and "Dad moved Maxi's airplane", but also non-critical features such as "Maxi had an airplane". Each item was assessed on three inclusive levels: first children were asked to tell what happened in the story. For each item they did not mention spontaneously, they received a direct question ("what was the boy's name"). If they failed to answer this too, they got a two-choice question: "*Was his name Maxi or Taxi?*". The number of free recall, supported recall and forced choice correct responses were coded separately, allowing us to analyse memory on three different levels. One containing only the number of items freely recalled, one containing the number of items recalled either freely or with help, and finally one score based on the number of items remembered by any means (freely, with support, and correct answer to a two-choice question). Counting only free recall measures children's spontaneous retelling of the story, adding supported recall tests their ability to recall items, and finally adding the forced choice alternative poses a way to see whether they are able to recognise features which were not recalled.

The eye tracking measures

In study III we used the eye tracking technique in order to determine where children looked, and when, during the showing of the Maxi film. By this technique, infra-red light is emitted by the eye tracker, which reflects to the subject's cornea and is recorded with a camera specific to this type of light. We used the SensoMotoric Instruments (SMI) RED 250 eye tracker, which samples eye movements at a rate of 250Hz. It belongs to the remote class of eye trackers, e.g. it is not needed to attach anything to the subject. It is thus non-invasive, and the infra-red light emitted is weak compared to normal light, rendering it a very safe technique. It needs calibration, however, which is done by asking the subject to look at dots showing up on the screen in specific positions known to the eye-tracking software. This is followed by a validation procedure, during which new dots are presented, and the deviation from the expected eye position can be assessed. As this requires that the subject look at the dots both during the calibration and the validation event, we acknowledged that this might pose a problem if the children are not fully attentive during the (to the film-watching) comparably boring calibration and validation procedures. For this reason, we had a second validation prepared. During the initial scene of the film, when Maxi was presented against a white background, we verified qualitatively whether the children were looking at him in a reasonable way, e.g. looked consistently at his figure and also scanned his face as would be expected. As Maxi's face corresponded to about 2.5*2.5 degrees of visual angle, we could therefore verify the calibration with an accuracy of 2.5 degrees.

As the eye-tracker produces a stereo stream of X, Y coordinates of the positions of each eye's focus, post-processing is needed in order to detect and differentiate oculomotor events. This was done by the SMI BeGaze software, using the standard settings. We were interested in the fixations, which correspond to segments in time during which the eye is relatively still and thus looking at a specific feature, which is also the time during which visual processing takes place (Holmqvist, Nyström et al. 2011).

Summary of the papers

The papers consist of two branches, differing in how they relate the social properties of interest, as displayed by a subject's manifest visual attention, to infants and children's understanding of others' differing perspectives.

The first paper (I) is concerned with spontaneous reactions to social stimuli, outlining how interest contagion can create memory effects in non-verbal FBTs, leading to what seems to be socially sensitive performance.

The other two papers (II & III) concern implications of showing interest towards a third entity in a social context, exploring what happens with children's social behaviour when interest has been shared – Paper II focuses on their interaction with the experimenter when discussing a verbally narrated false-belief story, and Paper III deals with their processing of the same verbal false-belief story as a function of experimenter's interest in the story.

Paper I (Falck, Brinck, Lindgren)

This paper concerns how interest contagion, as defined in the introduction of this thesis, leads to effects on memory in certain situations in a way explaining some evidence of infants' understanding of others' beliefs.

The background of this paper comes from Kovács et al. (2010) experiment mentioned above. It takes its starting point in an interesting property of Kovács et al. experiment – that it captures the sensitivity to agents' attentional focus without requiring prediction of an action by the agent, or even an assessment of an action. Instead, this experiment used simple reality judgements as test cases (reaction times for adults and looking times for infants), measuring only the reaction to a world state – a ball either shown or not shown at a specific location. The interpretation given by Kovács et al. (2010 implies a kind of meta-representation: "[F]rom 7 months on […] humans automatically *compute other's beliefs* and seem to hold them in mind *as alternative representations of the environment*" (my emphasis). On their account, the agent's belief is represented in parallel with our own and is thus influencing our behaviour (though it is not necessarily being attached to the agent, see Kovács, 2015).

In contrast, I and colleagues argue that in order to explain infants (and adults) behaviour on this task in the most parsimonious way, we need to assume two processes. First, we need 1) a mechanism for recognition memory that allows differential sensitivity or readiness to accept a specific world state, and we need 2) a tendency to direct our attention to targets of agents' attention. Additionally, both mechanisms must exist in 7-month olds. (1) is evident in children from 4 months onward (Baillargeon, 2008), (2) exists from 6-7 months (see introduction). We thus argue that the explanation of infants' behaviour as an effect of recognition memory and social attention is parsimonious because it relies on mechanisms known to be present in infants (e.g. process parsimony).

We also suggest that interest contagion functions as scaffolding for more elaborate abilities, in particular the ability to use an agent's overt interest as a cue for memory retrieval. This would imply a bridge between this basic ability to share information, and a later developing ability to understand that information is not always shared. Crucially, in the Kovács et al. (2010) experiments which we try to explain the outcome of, the sensitivity to what the agent had seen (and showed interest in) was not modulated by whether the agent was around when the infants looking times and adults' reaction times were measured (Kovács et al., 2010, experiment 2). This suggests that the reactions was more driven by the agent's effect on the observer's (experiment participant's) object processing, rather than of the observer's understanding of the agent's role in the situation.

This can be contrasted to experiments showing that infants' and adults' reactions are sensitive not only to the past history of agent-object interactions, but also to the presence of the agent. This can be seen as a step closer to a rational understanding of the scene, as if the agent was expected to act in a certain way. The Onishi & Baillargeon (2005) study included a condition (FB-Yellow, see fig 1), which could hardly be explained as a function of interest contagion only. We gesture at the end of the paper at a solution in which the agent serves as a cue for remembering the previous situation in which the agents' belief was true, when they were last associated with the object in its previous location. While such explanations has been proposed before (e.g. Perner & Ruffman, 2005), what might be a problem for such an account is the broad range of other studies which has shown remarkable contextual variety in where and when infant show abilities suggesting that they (on some level) understand others' false beliefs (Baillargeon et al., 2010), especially the eye-tracking experiment showing that toddlers look in advance to anticipate a mistaken agent's action (Southgate et al. 2007). We do not solve this in full in the paper, but we note that the agent's interest is likely to be instrumental for its role as a contextual cue.

Paper II (Psouni, Falck, Boström, Persson, Sidén, Wallin)

In this paper an implication of interest transfer is investigated, namely, how sharing of attentional targets can influence children's reasoning about people's actions when these are constrained by the information they have access to (their beliefs). Specifically, in this paper and the next, I and colleagues explore the effects of sharing an attentional target (and thus a perspective) with the experimenter in verbal FBTs.

We employed the Maxi film, described above, in order to be able to manipulate the experimenter's interest in the film, and thus knowledge of it, while still being able to present the story as a verbal narrative. Crucially, the experimenter was released from their typical role as a narrator, which is a persistent feature of many standard verbal FBTs. This solves a concern with the standard FBT, namely, that who is narrating the story also knows the end of the story, and thus the answer to the FB question. Thus we could get a clean test of the role of the experimenter's attention to the story. In two experiments, we tested children's reasoning about Maxi's expected action, while manipulating to which extent the experimenter showed interest in the story together with the child.

Experiment 1

The first experiment was designed to manipulate the experimenters' and child's sharing of the FBT story, in a way which could disentangle the experimenter's attention to (and hence perspective on) the FBT story from its mere presence. We tested children in three conditions[9].

Watching together, in which the experimenter announced "Let's watch a film", sat beside the child, keeping attention to the film the entire time.

(Test leader) Inattentive, in which the test leader sat beside the children, but did not watch the film, instead they were "working with their papers", which they also explained to the child in the beginning.

(Test leader) Absent, in which the test leader left the room, excusing themselves with that they had to work with their papers.

The reasoning behind these conditions was to parametrise the level of presence of the experimenter, so that we could isolate both a condition in which the experimenter was clearly not seeing or knowing the events, and a condition in which the experimenter was overtly uninterested (as opposed to absent). Better

[9] I use different naming here compared to the articles, as to be able to use the same names overall. (1) is named "test-leader jointly attending" in paper 2, and (2) is named "watching alone" in paper 3.

performance in the *watching together* condition would indicate that the children is following the story better in this condition, that they find the FBT question less strange because the experimenter has better reason to wonder about a story which they attended, or because the experimenter's perspective in general is more accessible when the story is shared. If we instead would see worst performance in this condition, it would indicate that the mere fact that the experimenter knows the current location of the object makes the children think of that location when answering the question (see Helming, Strickland & Jacob, 2014, for possible mechanisms).

Findings

We found that children performed better in the *watching together* condition, in contrast to both other conditions. This suggests that children's reasoning about Maxi's perspective is improved when the entire story has been shared with an experimenter compared to not shared, moreover, the experimenter's overt interest is instrumental rather than their mere presence. What we do not know is whether the beneficial effect of watching together pertains to better attention all or parts of to the story, or to different pragmatic dynamics of the test situation in which the FB question is asked. Therefore we followed up with a study to test the children's memory of the FBT story, keeping two of the three experimental conditions from experiment 1.

Experiment 2

This experiment addressed the remaining questions from experiment 1. Were children performing better because their attention to the film was better in general compared to the conditions in which the experimenter was inattentive, or were their performance gain specific to the false-belief reasoning which follows the FBT question? In order to answer this, we repeated the two 'present' (*Watching together* and *TL inattentive*) conditions, while also adding a memory measure, in order to tap in to their non-mentalistic understanding of the events in the story. The Memory for Maxi scale (see methods) was used to assess children's memory and story comprehension. We reasoned that if children are in general less focused to the story when the experimenter is looking away, we would see better memory overall for the *Watching together* condition compared to *inattentive*, but not if the effect on FBT performance had to do with the interaction with the experimenter when the FB question was asked. Finally, we allowed the possibility that specific aspects of the story would be remembered better, as the Memory for Maxi scale contained items pertaining to different aspects of the story.

Findings

We replicated the central finding from experiment 1; the children in the *watching together* group were more likely to answer the FBT question correctly. The effects on memory of watching together were confined to free recall of specific items, however. Thus, we could not support the idea that an increase in attention or arousal led to a better comprehension of the film. However, we noted that the children who watched together with the experimenter were more likely to freely recall both that Maxi went out to play and that the airplane had been moved. This could reflect a stronger memory of this fact, but it could also reflect a larger willingness to discuss this critical fact with the experimenter. These free recall items did not mediate the relation between experimental condition and FBT performance, however, as adding these dampened the contribution from experimental condition only marginally (see Paper II tables 3 and 4). They did dampen the contribution from WPPSI information sub score, however. This suggests that the improvement in FBT performance is either due to differences in how the child responds to the experimenter's question, or due to some processing variable which we have not been able to capture.

Paper III (Falck, Gunnäng, Hyddmark, Psouni)

Paper III aimed to further investigate the potential of the Maxi film, by adding an eye tracking measure, while keeping the experimental manipulation from the previous study (2 conditions). This work had roughly three goals:

To create and validate a way of studying the process of watching a false-belief story.

To see whether children's looking pattern at the moment Maxi is announced to be searching for his airplane reflects an understanding which does not show up in their verbal answers, e.g., a non-verbal spontaneous FBT.

To investigate children's watching of the film in the two conditions, watching alone or watching together. Could the difference in verbal FBT performance seen in paper II be explained by their scanning of the scene being different?

Findings

We found that children as a group looked significantly more to the empty location, which corresponds to them 'passing' the non-verbal and spontaneous FBT. In contrast, fewer children than in the previous experiments passed the verbal task, reflecting a floor effect. This rendered mediation analyses of experimental condition, children's visual processing, and verbal FBT performance impractical. Anyhow, we found that children watching together had a larger number of

fixations, but no longer fixation time in total, suggesting that they were more active in integrating the different elements of the story. Finally we also saw some correlation between looking behaviour during the event when Maxi's search action is prompted (the non-verbal FBT) and looking behaviour when Maxi's airplane was moved – children's looking to Maxi outside the window during the moving of the toy were positively correlated with their looking towards the empty location during the search event.

The floor effect in verbal FBT performance was surprising, as we had a similar setup here as in the studies reported in paper II. The children were younger than in the paper II studies, but not much younger – In experiment 1, if we choose the two comparable conditions (*watching together* and *TL inattentive*) we had 50% pass rate at 43.5 months; in experiment 2 (both conditions) 46% pass rate at 44.7 months, whereas in this experiment we saw 24.1% pass rate at 41.0 months. This discrepancy cannot be readily explained by the younger age of the present group of subjects, as selecting the 9 oldest children in this data set (so that we arrive at a mean age of 44.8 months), still increases the pass rate to only 33% (3 children passing). We speculate in the paper that it might have had to do with the calibration procedure, which took several attempts for most children. Children may have found the situation strange because of this. It could also be the RED 250 eye tracker setup, even though it is a standard computer monitor with the eye-tracker placed underneath, it is somewhat large and bulky. In experiments 1 and 2 the film was presented with a laptop, which is perhaps more similar to the situations in which 3-year-olds normally watch films.

General Discussion

All the included studies have been concerned with enhancements of perception and/or cognition by the means of sharing interest. Either individually, as in the case with 7-month old infants and adults watching an agent acting upon an object, or mutually, as in 3-4-year olds watching a film together with an experimenter.

First I will elaborate on the account given in paper I, drawing some consequences of the reasoning presented in the paper and connecting to other problems and findings in the literature. Second, I will discuss the implications of the results from papers II and III, especially the question of what is the primary mechanism behind the children's performances on the verbal FBT. Then I will point to some notable findings outside the scope of the main objectives of the thesis, before concluding.

Consequences of interest contagion for how to understand others' belief-guided actions

Paper I argues that interest in objects is contagious, and that this is neglected in theories of perspective taking, i.e. ToM cognition and its 'implicit' counterparts. Interest in objects as a consequence of contagion may provide a means of arriving at shared information, which in turn solves part of the problem of understanding 'beliefs'. This happens non-cognitively, and is not reliant on attributing abilities to the agent. It suffices to be sensitive to expressions of interest (see also Fenici, 2015).

Consider the following properties of the influence on interest contagion on recognition memory: Interest contagion is possible in situations in which an agent's interest is bodily manifest. This might lead the observer to attend to the target of the agent's interest, which affects recognition memory in the observer – interest in an object leads to increased processing of the same object, making it more recognisable in all contexts. This explains Kovács et al.'s (2010) result. Interest contagion also has another effect, namely, that the observer will individually share information about the target of the agent's interest. Thus, the agent and the observer will arrive at similar beliefs (see also Heyes, 2014a), which suggest that a basic building block of belief understanding is already in place –

they converge on a belief, though they do not (by this mechanism alone, at least) become aware of each other's belief. I claim two novel traits in this approach:

1) Sharing the agent's belief is not due to perspective **taking** because there is no perspective being 'taken' in any cognitive way, rather, agents object-directed actions causes copying of behaviours in the same direction. Neither is the other's perspective adopted in any implicit way (which would be the consequence of implicit perspective-taking/Theory-of-Mind/Metalizing), since only behaviours are copied. There is no need for an observer to 'think about' the agent (which does not exclude that they can, of course, if they have the cognitive capacity and the situation affords it).

2) Sharing the agent's belief is neither about **perspectives** in the monolithic, modular sense outlined above, as our theory makes no reference to the other person's view of the world (cf. Kovács et al., 2010). Of course, the other's interest partly determines their perspective, as well as my convergence of interest with them may lead to me adopting a similar perspective. But the mechanism makes no reference to a perspective in the form of a construct attached to an individual agent. Rather, what is shared is the information which leads to the specific belief in the observer, and it is shared through the agent's focus on the same information.

Consequently, our account provides a generic, non-cognitive way of sharing information in a specific context, without awareness of sharing and without attributing psychological capacities to agents.

Interested agent as a cue for recall?

Extrapolating from the argument in paper I, I suggest that some false-belief understanding can be realised by an observer reinstating a past event cued by an agent associated with that event appearing. I suggest that the agent's manifesting interest similarly to in previous situations may activate memory of the previous situation (in which the false belief was true). This would correspond to cued recall, where the agent's interest mediates the agent's role as a contextual cue for the observer, so that the interested agent is what is activating the memory of the previous situation. This corresponds to recall in terms of memory systems, hence implies context sensitivity (Diana, Yonelinas & Ranganath, 2007). The resulting model of non-verbal FBT performance, applicable to the situation instantiated by the Onishi & Baillargeon (2005) experiment, thus differs in an important way both from the mentalistic explanations suggested by Onishi & Baillargeon (2005) and the minimalistic explanation suggested by Apperly & Butterfill (2009). On my account, the belief of the agent is not represented as a possible worldview that is attributed to an agent, not even in a simplified form of a registration (cf. Apperly

& Butterfill, 2009; Butterfill & Apperly 2013). Rather, infants' looking behaviour emerges from an interaction of memory processes in the observer.

The perceptual history of the agent is also a history of modulating our own perception when interacting with the agent. The ability to contextualise own past perceptual states with the agent's as a cue is achieved without attributing a specific ability or capacity to the agent, and it does not require explicit thinking. This explains how adaptive social behaviour can arise, without apparent cognitive effort, in situations where agents' act in accordance with their perceptual history. Moreover, it makes plausible that the difference between reactive, spontaneous handling of false beliefs as described here and explicit (e.g. instructed) reasoning about beliefs involves similar process differences as between episodic intrusions and episodic imagery. A suggestion similar to the one made here was made by Perner & Ruffman (2005) in response to the seminal infant FBT study by Onishi & Baillargeon (2005). They suggested that the infants associate the experimenter with the object's last location. This, however, has problems with explaining the results from anticipatory looking in infants around two years of age (Southgate et al., 2007). This is the reason why I suspect that at least some sensitivity to the agent's interest is required, not any appearance of the same agent will do. This does not imply that infants need to attribute interest to the agent – as shown in the introduction of this thesis; interest is a property of behaviour as much as a mental state.

Criticism and ways forward

The argument put forth in paper I can be criticised on grounds that it is not an experimentally tested account – I am rather trying to find the most plausible explanation to existing data. This is a just criticism, and the reason why I am defending it in terms of parsimony. I am suspecting that this theory cannot suggest a definitive test which can refute a mentalistic alternative – I see no way around the problem discussed by Povinelli and Vonk (2003) and Reddy and Morris (2004), that mentalistic explanations are impossible to test directly. Rather, I believe we must search for 'signature limits' (Apperly & Butterfill, 2009) - limitations on when and where the phenomenon appears. The tendency to assume that information is shared based on other cues than agents' seeing (Dumontheil et al. 2010; Moll et al., 2011) could be such a limit, suggesting that object-directedness other than seeing can give rise to similar effects as correctly understanding another's belief.

A related account of infants' performance

Heyes (2014b) criticises the mentalistic interpretations in the bulk of the infant false-belief literature, arguing that infants' behaviours can be explained by simpler mechanisms. She gives an account on Kovács et al.'s experiment which resembles ours, based on modulation of recognition memory. She argues that the phenomenon of retroactive interference, the tendency to forget an event due to the appearance of another event close in time (Pearce, 2008), arises as a consequence of most FBT designs, thus disrupting infants' memory in a way which can explain the observed behaviour.

I agree with this reading, but I would add two brief points. First, interest contagion concerns the processing of information within events, whereas retroactive interference concerns shifts of events, though their consequences for the memory of events are similar. Second, the boundaries between events which retroactive interference rely on are not arbitrary. I argue in an unpublished manuscript (Falck & Lindgren, unpublished) that event segmentation (Kurby & Zacks, 2008) when observing an agent's actions is partly constituted by the agent's shifts of interest and thus an essential part of social perception. This fact about how we perceive events has the important consequence that Heyes' (2014b) retroactive interference account, while not assuming mentalistic mechanisms, is still reliant on social information, something that she also acknowledges (Heyes, 2014a). Thus, whereas retroactive interference might well be at work in many of the infant FBT studies, this phenomenon itself is embedded in an event structure which is to a large part constituted by changes in agents' activities.

Therefore, the mechanisms proposed by me and Heyes (2014a; 2014b) are compatible but involve slightly different social information, and may thus work in concert in both FBTs and in everyday social situations. An important consequence is that these two low-level accounts of infants' social abilities, though relying on in principle domain-general mechanisms, are nevertheless profoundly social, as they make use of information intrinsic to how people regularly behave. The point is not to explain away, but to explain.

Further consequences of attentional shifts

Agents' shifts of interest can also explain the puzzling discrepancy between Buttelmann et al.'s (2009) true-belief condition and Allen's (2015) hands-full condition. Recall that in the Buttelmann et al. study there was no reason for the child to infer that the experimenter (E2) would want to open the empty box, save a mentalistic interpretation. In contrast, the Allen (2015) study gave such a reason, but the three-year-olds still failed to open the empty box. The argument played out

in Paper I of this thesis suggests a solution to this problem. The answer can be extracted from E2's overt interest during the original true-belief condition from Buttelmann et al. (2009). Since E2 made very clear that he saw the toy being moved, marking this with an "ah" when the moving was finished, he overtly signalled that he was 'done' with the toy. E2's shift of interest would then render the toy less salient for the child, making opening the empty box a more relevant response. In contrast, in the hiding (hands-full) scenario, the toys were highly salient – it was obviously critically important to hide them from Oscar-the-Grouch, the toy-stealing adversary. If three-year-olds attention is highly susceptible to the saliency given to certain objects by the social situation, while older children's attention is more flexible (and more controllable), these results would be expected, and well in line with the findings from Rubio-Fernández & Geurts (2013) reviewed above – since the toys are important now, the three-year-old will open the box with the toys, even though it is maladaptive given the current situation in which toys have to stay hidden. They simply cannot ignore what is the most important *object*, rather than *objective*, within the shared context.

Consequences of sharing the Maxi story for children's understanding of Maxi's perspective

Turning to the other two papers, I will discuss them together as all three experiments share the same experimental design structure. In the following, I will refer to the experiments as 1, 2 and 3, where the two first are the ones presented in paper II and '3' denotes the eye-tracking experiment (Paper III). Their aim was to find out what the experimenter's overt interest in, and hence apparent knowledge of, the adventures of the protagonist in a verbal FBT story means for children's performance. Children's performance was assessed with an elicited-response task in all three experiments, and additionally a spontaneous-response task in the last experiment.

How our experimental manipulation implements sharing

Importantly, the means by which the experimenter's interest is transferred to the child in these experiments differs in an important aspect from interest contagion: the shared interest in the "experimenter watching" condition is established within an explicitly triadic context, that is, the experimenter makes clear that "we" are watching the film. Moreover, it differs in the sense that the child is observing the experimenter sporadically, in contrast to how infants observe the agent in Kovács and colleagues' (2010) experiments. Thus it implies a case of *mutual sharing* in

Brinck's (2008) terminology. Recall from the section on social attention that mutual sharing still allows focus on the shared object most of the time.

The experiment session starts with the experimenter addressing the child, saying *"now we get to watch a movie"* (or: *"now you get to watch a movie, I will {sit here and work with my papers | go outside and work with my papers}"*). Our film-watching starts as a joint attentional act, as the experimenter is declaring "now we get to watch a film", whereupon the film is started and the experimenter turns their attention towards it (in the *watching together* condition). The first episode of attention contact (eye contact) occurs while preparing to watch the film together and is initiated by the experimenter. Following it, attempts at attention contact from the experimenter are not repeated. Any attempts by the child at making contact, verbally or by looking, are acknowledged by the experimenter with a humming sound or a brief look, without making lasting eye contact.

Thus we had no reason to expect the children to look at the experimenter in a way making them miss important information. We can safely assume that as a rule, children will look to the experimenter only when the child finds no particular interest in the scene shown, or when something interesting happens and the child has already grasped the critical information, just wanting to check that the adult is looking. This ensued, as confirmed by watching the video uptakes, that the children focused their attention to the film most of the time. Regardless of which form of reciprocity each child experiences with the experimenter in the film-watching, it seems intuitively correct to describe the experimenter and child as watching the film together. The fact that the story progresses regardless of the experimenter's actions releases the experimenter from the role as a narrator.

Consequently, it did also differ from the standard FBT in terms of reciprocity between the child and the experimenter. In the standard FBT in which the experimenter enacts the entire story, it is done in interaction with the child. In the original Maxi task, every scene is shown to the child in an ostensive manner. Here is Maxi, Maxi puts away his chocolate, Maxi wants to play outside… in this event, as the experimenter is narrating to the child, the experimenter is also attending to the child and making sure that the child attends the information. This qualifies the entire story as a joint attentional episode, in which joint attention is established repeatedly to different features of the story. Attention contact takes place a number of times, and whenever the experimenter has made sure that the child is attending to the information, s/he continues. This is not as much a design choice as a prerequisite for the situation to work at all – as the child and experimenter is playing in a shared space, it would be weird not to establish rapport and make sure that the child follows each step of the story. Joint attention is re-established again for each scene in the standard FBT, making each scene a joint attention episode. The experimenter, in directing the child's attention, becomes reinforced in their

role as a narrator. This implies an authority which puts the experimenter and child in a certain un-equal relation, reinforced by the fact that the experimenter provides the information constituting the story. Our procedure removed this feature, while still allowing the experimenter and child to share the elements of the story in a social context, without implying the additional authority of the experimenter being a narrator.

Results summary

We found that performance was better when the child and the experimenter watched the film together, before the experimenter asked the FB question.

Experiment 1 additionally showed that it did not matter whether the experimenter was away, thus clearly missing out on information, or merely inattentive, which would have rendered the experimenter's information access less obvious. Compared to the *watching together* condition, the FBT performance in the *inattentive* condition led to the numerically worst performance, though not significantly (p = .30 for the *inattentive-absent* contrast). A cautious interpretation of this difference would suggest that an uninterested experimenter is more detrimental to performance compared to an absent one. This is compatible with a parametric scaling of the interest in the story, suggesting that an interested experimenter (*watching together*) is facilitatory and an uninterested experimenter is costly to children's performance, compared to a baseline of watching the film alone. The evidence from experiment 1 suggests that it is indeed the experimenter's interest in the story which matters, together with its implications for sharing a perspective.

Experiments 2 and 3 (the eye-tracking experiment) were designed to tap not only into verbal FBT responses, but also into children's processing of the story. The Memory for Maxi scale in experiments 2 and 3 found no evidence for a general impact of sharing the story on memory, if we added up all forms of recall (free + supported) for all items. We did see increased likelihood of free recall of the fact that Dad moved the airplane in the watching together condition. The eye-tracking experiment (3) further showed that children in the watching together condition looked more to Dad when he moved the airplane.

The eye-tracking experiment, importantly, added a spontaneous FB measure in terms of children's predictive looking when Maxi was about to search for his airplane. This showed that more children looked towards where he had left it, than towards where it really was, whereas their verbal responses overwhelmingly pointed to where it really was (22 out of 29 children). This is in line with previous findings from three-year olds when a spontaneous measure is employed (Clements & Perner, 1994; Garnham & Ruffman, 2001; Rubio-Fernández & Geurts, 2013).

Moreover, we saw that children who looked more at Maxi outside the window when the airplane was moved were more inclined to look at the empty (correct) location in the search event. This is in line with Rubio-Fernández & Geurts (2015) finding that three-year-olds need visual access to the protagonist during the movement event in order to succeed at a spontaneous-response FBT.

The question which remains, however, is why do we get better performance on the verbal FBT when children watch together with the experimenter? We may discern at least four possible mechanisms behind these findings. Let us consider them in turn.

Attention in the encoding phase

At least two mechanisms pertaining to the encoding phase can be discerned.

(1) Attention to the story in general

Perhaps children's general attention to the story is improved, as a consequence of an increase of arousal or general attention because of sharing attention with the experimenter.

(2) Attention to specific items

Perhaps sharing attention with the experimenter renders the children more attentive to specific details which are important for the gist of the story, such as the fact that Maxi's airplane is moved in his absence.

Evidence for encoding-phase mechanisms

In the eye tracking study, children who watched together fixated the story more often overall, which may suggest a general increase in attention. We did not see a general memory improvement in any of the two studies employing the memory for Maxi scale, experiment 2 and 3, however this measure might be much less sensitive. More scanning is not necessarily connected to better memory, as many processes contribute to visual scanning (Holmqvist et al., 2011).

Children who watched together were more likely to freely recall that Dad moved Maxi's airplane in experiment 2 (though not in the eye tracking experiment, though the power there is substantially lower due to much lower N). Moreover, in the eye-tracking experiment, children who watch together look more towards Dad when he moves the airplane[10]. If we allow us to interpret these findings across experiments, assuming that they reflect the same process, it would suggest that the

[10] Significantly more when he picks up the airplane from the chest, and not far from significantly more when he puts it down in the dresser.

children who watch together with the experimenter are more interested in this event already before they understand that Maxi will come back for his airplane. If children are more interested in this event, that would explain why they talk more about it afterwards – either because of improved memory for the event, or because they are more eager to tell about what they found interesting (free recall). Generally, the evidence seems to favour (2) over (1). However, as the contribution of experimental condition was not substantially reduced when adding the Memory for Maxi items into the regression (Paper 2 table 4), the free recall of these items cannot explain the relation between watching together and passing the verbal FBT. It is possible that the differences in the scanning of the scene can explain variance in verbal FBT performance, but we need more data to answer that question, as we presently found a floor effect in FBT performance. Next, we turn to the test phase.

Pragmatics and shared perspectives in the test phase

Alternatively, the mechanism behind the improvement in FBT performance has to do with the test situation – either through the children finding the FB question more pragmatically plausible in the watching together condition, or through the answer being less demanding to formulate given that it refers to information which was shared.

(3) FB question pragmatics[11]

Releasing the experimenter from the role of narrator had two goals. First, we wanted to make the experimenter able to come and leave at will. Second, we wanted to make the asking of the FB question more believable, by making the experimenter seem ignorant about the end of the story.

If the experimenter is narrating the story, they obviously know how the story ends. Asking about something you already know indicates that something is fishy with the question, perhaps that it is designed to test their knowledge, which children from the age of two can sense (Grosse & Tomasello, 2012). This, we reasoned, could confuse the children. Either they may just think that the answer is non-trivial, so that they should choose the other location, or perhaps they start to reason about the experimenter's perspective in order to find out what they are after. However, even with this improvement of the procedure, our different conditions may still mean different pragmatic contexts for the interpretation of the FB question. Let us call this the pragmatics account. Its rationale is as follows.

[11] When developing the arguments in this section I had great help from discussions with Dr. Paula Rubio-Fernández.

When we have watched a story together, we have already established a shared interest in the story. Moreover, the Maxi story is of the kind that would be expected to end with a twist – clearly, Maxi is being deprived of critical information, and this is emphasised by the narrator. In a sense, it is not so different from many narratives which children encounters, narratives that indeed has been hypothesised as being crucial to children's learning to reason about situations in which people have or lack information (Hutto, 2008). Thus, the question of what Maxi will do is reasonably plausible, provided that the experimenter is also an audience of the story. The story has exactly the dramaturgy of stories for young children. Then, having watched together, and the film stops, a reasonable thing to ask is what happens next, more so if one has shared interest in the story.

In contrast, if interest is not shared, either because I am absent or looking away, asking the question implies at least two things. First, I am suddenly showing interest in the details of a story which I first did not care about. Second, though the experimenter in all our conditions had heard about Maxi and his airplane from start, and also heard that he was going to look for it in the end, it is not clear that the experimenter would care to ask so specifically about the airplane. Rather, perhaps a more open question would be expected, such as "what will happen now?", or even "what happened"[12]. It is thus possible, that this strangeness of asking interestedly about something one did not show interest in before, creates a weird situation which then triggers a complex reasoning process, as the child tries to figure out what the experimenter means (see van Overwalle & Baetens, 2009). Consequently, it is possible that our non-shared conditions impose a less pragmatically believable situation, confusing the children when they are to answer the FB question.

I will make a final note before moving on. When watching together, and the film stops, both the experimenter and the child knows the final location of the airplane. In contrast, in the conditions where the experimenter has not watched the film, the experimenter would not have a stance on the final location of the airplane. In the *referential bias* account of verbal FBTs suggested by Helming et al. (2014, p. 169), an important feature is what the experimenter knows of the object's final location. They argue that the child may in fact be answering from the experimenters' perspective on the object's location, predicting that an experimenter sharing the protagonist's perspective should improve children's answers. We did not test that possibility here, but we see that the experimenter knowing the object's final location does not incur a cost compared to not knowing anything about it.

[12] Recall that by design, we wanted the FB question to be as similar to the previous literature as possible, as we wanted to test whether it really misled the children, even if the situation was less pragmatically complex.

(4) Talking about the shared context

The other test-phase mechanism we suggested was that it might be easier to answer a question about a situation which is shared, compared to a situation which is not shared. This would imply that the main difference pertains to the test phase in which the FB question is asked. Following the conjecture that sharing information does not require perspective-taking, there is no reason why the experimenter's perspective on the toy should interfere with children's answers in the standard FBT. This is in line with what we see in these experiments. When the experimenter asks the question after sharing the entire film, everything in the story belongs to the shared context – Maxi, the airplane, Dad, the movement event.

In contrast, when the experimenter has just re-established contact with the child in the uninterested conditions, then the shared context involves only the airplane at its present location. But if the experimenter has been uninterested and therefore not shared the story, the child would have to consider not only what they think Maxi will do, but also which information they need to tell the experimenter in order to adequately answer the question about Maxi and the airplane. In this case, since the child has to reason about Maxi in the context which was shared with the experimenter, but also needs information from a context which was not shared, answering the question requires a switch of contexts akin to perspective taking (figure 2), which in turn disrupts their perspective-taking with Maxi. As only the final location of the airplane is associated with the shared context in which the question is asked, and none of the events needed in order to understand the story is, perhaps it is too cognitively demanding for the children to actualise the relevant information in order to answer the question.

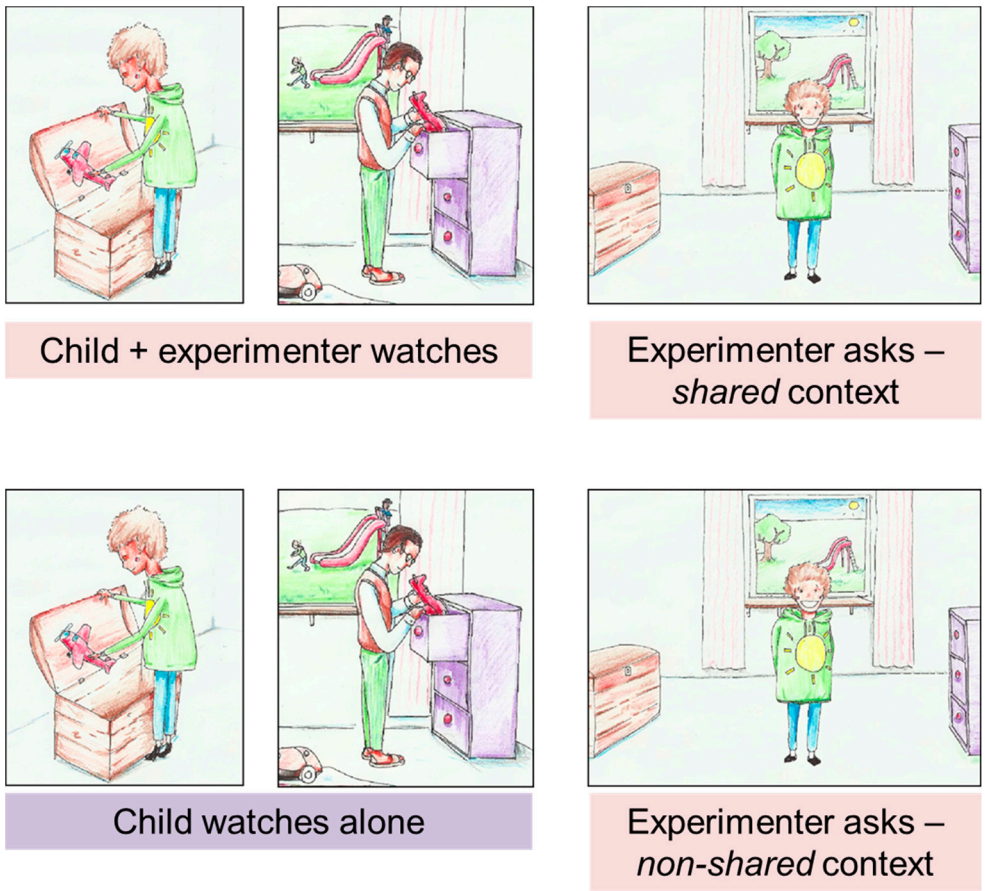

Figure 2. In the non-shared context, the child has to take the experimenter's knowledge into account, before being able to answer the experimenters question about Maxi. When the critical events in the story belong to the non-shared context, perhaps the current location of the airplane is the most accessible to the child. Pictures by Martin Persson.

Evidence for test-phase mechanisms

The evidence for these accounts is mostly negative, based on effects predicted by the encoding-phase accounts which we did not find. The strongest evidence for the test-phase account is perhaps the results from the spontaneous FBT (predictive looking), as this measure did not show any correlation with experimental condition (watching together vs alone). If the experimental condition affected the processing of the story so that more critical information for understanding Maxi's belief was attended during the story, we would expect this to surface also in the spontaneous FBT. Instead, we found such critical information in the number of looks to Maxi outside, though it did not correlate with the experimental manipulation. Moreover, our results tally with findings in the literature that spontaneous FBT responses

does typically not relate to executive and verbal ability, whereas elicited FBT responses does.

The shared-context account can explain the higher likelihood that children in the watching together condition mention that Maxi's airplane has been moved. As the difference between conditions was observed for free recall only, it may reflect a stronger tendency for the children to want to talk about this critical fact, which could be related to the pragmatics of the test situation (and mechanistically unrelated to the increased number of fixations to the moving event). Having understood Maxi's mistaken perspective (as evident from the correct FBT response), the moving event and the related information that Maxi went out to play becomes a relevant plot point to mention. Similarly, it is also possible that having shared the story with the experimenter encourages talking about the most important facts, as both have watched the film and have reason to discuss the story. Then, the answer to the question of what happened in the story would include the most important fact first, which is the fact that Maxi's airplane was moved, as this can be argued to be the most memorable event in the story. One may argue that children would instead be more eager to tell about the story to an experimenter who did not share it, predicting the opposite result. However, the Memory for Maxi procedure makes evident that the experimenter knows the answers as soon as they start asking leading questions, which renders the experimenters ignorance and hence curiosity less believable.

Ways forward

Taken together, the relative contribution of mechanisms in the test phase compared to the encoding phase is still unclear. An experiment to test the test-phase account would be possible, by having two experimenters of which one watches together with the child, and one comes in just before the film ends. Then, depending on experimental condition, either the experimenter who watched with the child (the *Co-watcher*) or the experimenter who just came in (the *Latecomer*) asks the critical FB question. This experiment would allow us to disentangle the encoding-phase account from the test-phase account (Table 1).

Table 1. An experiment to distinguish encoding-phase from test-phase mechanisms.

	Expected FBT performance	
Exp. condition	*Encoding phase mechanism*	*Test phase mechanism*
Co-watcher asks	High	High
Latecomer asks	High	Low

If the attention in the encoding phase is the main mechanism behind our results, it would suffice that one person watches the film together with the child, regardless of who is asking the question, for performance to improve above the in-attentive conditions in the present studies. In contrast, if it is mostly a question of pragmatics in the test phase, we would expect to see better performance when the person who was sharing the story with the child asks the FB question.

This experiment could also have the experimenter who asks the question lead the Memory for Maxi interview, as it would be interesting to see whether children are more informative about the events in the film when talking to the experimenter with which they shared the story. This is under the assumption that the experimenter, by asking the questions, does not reveal more knowledge of the story than they have from the shared experience. The opposite pattern of results would also be possible, that they want to tell the experimenter who had missed the story, at least provided that the children are allowed to finish their retelling before the supported questions are asked. In a follow-up study, it would make sense to include the Memory for Maxi scale with the small modification that we wait for all children to finish the free retelling before starting to fill in. This way we increase the chances that the Memory for Maxi interview is not regarded as a test by the child, but rather as a conversation about the recent experience.

Concluding remarks

The present thesis has considered different implications of the sharing of interest for infants' and children's understanding of other's perspectives. We have seen that due to the contagious aspect of interest, we have a non-cognitive basis of sharing information which does not rely on attribution of mentalistic abilities to others. Infants from 6-7 months onward are able to adopt others beliefs without effort and without needing reflection, which makes interaction fluent. Evidence from three-year-olds suggests that sharing information this way does not use up resources for reasoning about others' belief-guided actions. Instead, as long as the context provides sufficient social scaffolding, and the experimenter is not put in a position to confuse the children in the test phase because of having narrated the story, children between three and four perform quite well on reasoning verbally about a third agent's false belief.

I suggested above that infants first begin to understand others' diverging perspectives by recalling information previously shared with them, cued by the

appearance of the other person. The test-phase account of experiments 1-3 implies that discussing another agent's perspective with an experimenter is easier for children if the child and experimenter have shared the information which the reasoning process is based on. Together, these conjectures may form a basis for a general account of how we relate to others when their perspectives differ from ours, suitable for explaining the abilities of infants from at least 15 months of age to adults. If understanding an agent's false belief is essentially a matter of activating memory of information which was previously shared with the agent, the apparent puzzle about spontaneous-response and elicited-response FBTs can be solved. A spontaneous-response FBT provides direct access to the information shared with the protagonist via cued recall, whereas an elicited-response FBT provides at best only such direct access to the information which was shared with the experimenter. Consequently, during conversation with the experimenter, a controlled imaginary process is needed to activate the protagonist's perspective, which becomes even harder if the information feeding this process was not shared.

References

Allen, J. W. P. (2015). How to help: Can more active behavior measures help transcend the infant false-belief debate? *New Ideas in Psychology*, *39*, 63–72.

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.

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 (2006)*, *63*(6), 1201–17.

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

Baillargeon, R. (2008). For a Principle of Persistence in Infants' Physical Reasoning. *Perspectives on Psychological Science*, *3*(1), 2–13.

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

Barsalou, L. W., Breazeal, C., & Smith, L. B. (2007). Cognition as coordinated non-cognition. *Cognitive Processing*, *8*(2), 79–91.

Becchio, C., Bertone, C., & Castiello, U. (2008). How the gaze of others influences object processing. *Trends in Cognitive Sciences*, *12*(7), 254–8.

Bowler, D. M., Briskman, J. A., & Grice, S. (1999). Experimenter effects on children's understanding of false drawings and false beliefs. *The Journal of Genetic Psychology: Research and Theory on Human Development*, *160*(4), 443-460.

Bradmetz, J. (1998). Is the acquisition of a theory of mind linked to a specific competence beyond three years of age? *Intelligence*, *26*(1), 1–8.

Brinck, I. 2001. Attention and the evolution of intentional communication , *Pragmatics & Cognition*, 9(2), 255-272.

Brinck, I. 2004. Joint attention, triangulation and radical interpretation: A problem and its solution. *Dialectica*, 58(2), 179-205.

Brinck, I. 2008. The role of intersubjectivity for the development of intentional communication. In *The Shared Mind: Perspectives on Intersubjectivity*. Eds. J. Zlatev, T. Racine. C. Sinha, & E. Itkonen, Amsterdam: John Benjamins Publ., pp. 115-140.

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, *8*(6), 535–543.

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.

Butterfill, S. A., & Apperly, I. A. (2013). How to Construct a Minimal Theory of Mind. *Mind & Language,* 28(5), 606–637.

Call, J., & Tomasello, M. (1999). A nonverbal false belief task: the performance of children and great apes. *Child Development*, *70*(2), 381–95.

Carlson, S. M., & Moses, L. J. (2001). Individual Differences in Inhibitory Control and Children's Theory of Mind. *Child Development*, 72(4), 1032–1053.

Carruthers, P. (2013). Mindreading in Infancy. *Mind & Language*, *28*(2), 141–172.

Chomsky, N. (1957). *Syntactic structures*. Mouton de Gruyter, Berlin; New York 2002.

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

Colle, L., Becchio, C., & Bara, B. G. (2008). The Non-Problem of the Other Minds: A Neurodevelopmental Perspective on Shared Intentionality. *Human Development*, *51*(5-6), 336–348.

Diana, R. A, Yonelinas, A. P., & Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: a three-component model. *Trends in Cognitive Sciences*, *11*(9), 379–86.

Dumontheil, I., Küster, O., Apperly, I. A, & Blakemore, S.-J. (2010). Taking perspective into account in a communicative task. *NeuroImage*, *52*(4), 1574–83.

Fabricius, W. V., & Khalil, S. L. (2003). False Beliefs or False Positives? Limits on Children's Understanding of Mental Representation. *Journal of Cognition and Development*, *4*(3), 239–262.

Falck, A. & Lindgren, M. (unpubl.). Event segmentation in Kovács et al.'s "social sense" experiment. *(unpublished manuscript)*

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

Fenici, M. (2015). A simple explanation of apparent early mindreading. Infants' sensitivity to goals and gaze directions. *Phenomenology and the Cognitive Sciences*, 14(3), 497–515.

Fiebich, A. (2013). Mindreading with ease? Fluency and belief reasoning in 4- to 5-year-olds. *Synthese*, 1–16.

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development*, *10*(4), 483–527.

Garnham, W. A., & Perner, J. (2001). Actions really do speak louder than words-but only implicitly: Young children's understanding of false belief in action. *British Journal of Developmental Psychology*, *19*(3).

Garnham, W. A., & Ruffman, T. (2001). Doesn't see, doesn't know: is anticipatory looking really related to understanding or belief? *Developmental Science*, *4*(1), 94–100.

Gallagher, S. (2001). The practice of mind: Theory, simulation, or interaction? *Journal of*

Consciousness Studies, 8 (5-7): 83-107.

Gallagher, S. (2004). Understanding interpersonal problems in autism: Interaction theory as an alternative to theory of mind. *Philosophy, Psychiatry, and Psychology*. 11 (3): 199-217.

Gallagher, S. (2015). The Problem with 3-Year-Olds. *Journal of Consciousness Studies*, 22(1-2), 160–182.

Gallagher, S., & Hutto, D. (2008). Understanding others through primary interaction and narrative practice. In J. Zlatev, T. P. Racine, C. Sinha, E. Itkonen, J. Zlatev, T. P. Racine, ... E. Itkonen (Eds.), *The shared mind: Perspectives on intersubjectivity* (pp. 17-38). Amsterdam, Netherlands: John Benjamins Publishing.

Grosse, G., Tomasello, M. (2012) Two-year-old children differentiate test questions from genuine questions**.** *Journal of Child Language*, 39, pp. 192-204,

Hedger, J. A., & Fabricius, W. V. (2011). True Belief Belies False Belief: Recent Findings of Competence in Infants and Limitations in 5-Year-Olds, and Implications for Theory of Mind Development. *Review of Philosophy and Psychology*, *2*(3), 429–447.

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

Heyes, C. (2010). Where do mirror neurons come from? *Neuroscience and Biobehavioral Reviews*, *34*(4), 575–83.

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

Heyes, C. (2014b). False belief in infancy: a fresh look. *Developmental Science*, 1–13.

Holmqvist, K., Nyström, N., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (Eds.) (2011). *Eye tracking: a comprehensive guide to methods and measures*, Oxford, UK: Oxford University Press.

Hutto, D. D. (2004). The limits of spectatorial folk psychology. *Mind & Language*, *19*(5), 548–573.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, *14*(2), 201–211.

Kampis, D., Somogyi, E., Itakura, S., & Király, I. (2013). Do infants bind mental states to agents? *Cognition*, *129*(2), 232–240.

Karmiloff-Smith, A. (1992). *Beyond Modularity: A Developmental Perspective on Cognitive Science (Learning, development, and conceptual change).* MIT Press.

Korkman, M., Kirk, U., & Kemp, S. (2007). NEPSY-II (2nd ed.): *Clinical and interpretive manual*. PsychCorp: San Antonio.

Kosslyn, S. M. Koenig, O. (1992). *Wet Mind: The New Cognitive Neuroscience.* The Free Press, New York.

Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: susceptibility to others' beliefs in human infants and adults. *Science (New York, N.Y.)*, *330*(6012), 1830–4.

Kovács, Á. M. (2015). Belief Files in Theory of Mind Reasoning. *Review of Philosophy and Psychology*.

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, *12*(2), 72–79.

Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind." *Psychological Review*, *94*(4), 412–426.

Leslie, A. M. (2005). Developmental parallels in understanding minds and bodies. *Trends in Cognitive Sciences*, *9*(10), 459–462.

Mahner, M. & Bunge, M. (1997). *Foundations of Biophilosophy.* Berlin: Springer-Verlag.

Michael, J. (2011). Shared Emotions and Joint Action. *Review of Philosophy and Psychology*, *2*(2), 355–373.

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

Moll, H., & Tomasello, M. (2007). How 14- and 18-month-olds know what others have experienced. *Developmental Psychology*, *43*(2), 309–317.

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

Moll, H., Carpenter, M., & Tomasello, M. (2011). Social Engagement Leads 2-Year-Olds to Overestimate Others' Knowledge. *Infancy*, *16*(3), 248–265.

Overwalle, F. Van, & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems : A meta-analysis. *NeuroImage*, *48*(3), 564–584.

Pearce, J. M. (2008). *Animal learning and cognition: An introduction*. Florence, KY: Psychology Press.

Perner, J., & Roessler, J. (2012). From infants' to children's appreciation of belief. *Trends in Cognitive Sciences*, *16*(10), 519–25.

Perner, J., & Ruffman, T. (2005). Infants' Insight the Mind: How Deep? *Science*, *308*(5719), 214–216.

Povinelli, D.J. & Vonk, J. (2003) Chimpanzee minds: Suspiciously human?, *Trends in Cognitive Sciences*, 7 (4), pp. 157–160.

Pierno, A. C., Becchio, C., Wall, M. B., Smith, A. T., Turella, L., & Castiello, U. (2006). When gaze turns into grasp. *Journal of Cognitive Neuroscience*, *18*(12), 2130–7.

Poulin-Dubois, D., Sodian, B., Metz, U., Tilden, J., & Schoeppner, B. (2007). Out of Sight Is Not Out of Mind: Developmental Changes in Infants' Understanding of Visual Perception During the Second Year. *Journal of Cognition and Development*, *8*(4), 401–425.

Reddy, V. (2009). *How infants know minds.* Delhi: Prentice Hall India.

Reddy, V., & Morris, P. (2004). Participants Don't Need Theories: Knowing Minds in Engagement. *Theory & Psychology*, *14*(5), 647-665. doi:10.1177/0959354304046177

Rubio-Fernández, P. (2013). Perspective tracking in progress: Do not disturb. *Cognition*, 129(2), 264–272.

Rubio-Fernández, P., & Geurts, B. (2013). How to Pass the False-Belief Task Before Your Fourth Birthday. *Psychological Science*, 24, 27–33.

Rubio-Fernández, P., & Geurts, B. (2015). Don't Mention the Marble! The Role of Attentional Processes in False-Belief Tasks. *Review of Philosophy and Psychology*, (1).

Rushworth, M. F. S., Johansen-Berg, H., Göbel, S. M., & Devlin, J. T. (2003). The left parietal and premotor cortices: Motor attention and selection. *NeuroImage*, *20*(SUPPL. 1), 89–100.

Sodian, B., & Thoermer, C. (2008). Precursors to a theory of mind in infancy: perspectives for research on autism. *Quarterly Journal of Experimental Psychology (2006)*, *61*(1), 27–39.

Southgate, V., Senju, A, & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, *18*(7), 587–92.

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13- month-old infants. *Psychological Science*, *18*(7), 580–586.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development*, *72*(3), 655–84.

Wilby, M. (2010). The simplicity of mutual knowledge. *Philosophical Explorations*, *13*(2), 83–100.

Wilby, M. (2011). Embodying the False-Belief Tasks. *Phenomenology and the Cognitive Sciences*, 11(4), 519–540.

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.

Yu, C., & Smith, L. B. (2013). Joint attention without gaze following: Human infants and their parents coordinate visual attention to objects through eye-hand coordination. *PLoS ONE*, *8*(11).

Zlatev, J., Brinck, I., & Andrén., M. 2008. Stages in the development of perceptual intersubjectivity, In *Enacting Intersubjectivity: A Cognitive and Social Perspective to the Study of Interaction*. Eds. F. Morganti, A. Carassa, & G. Riva. Amsterdam: ISO Press, pp. 117-132.