

Mindreading, Privileged Access and Understanding Narratives

Szabolcs Kiss¹ and Zoltán Jakab²

1 Institute of Psychology, University of Pécs
7624 Pécs, Ifjúság útja 6, Hungary
kiss.szabolcs@t-online.hu

2 Institute of psychology for Special needs, Loránd Eötvös University
1071 Budapest, Damjanich u. 41-43, Hungary
zoltan.jakab@barczi.elte.hu

Abstract

In this paper we first offer a task analysis of the false belief test including the bidirectional relationship between mindreading and language. Following this we present our theory concerning Quinian bootstrapping of the meaning of mental state terms and relate it to the task-analytic framework. Finally we present an experiment on ascribing privileged access through minimal narratives which is intended to serve as a test of our theory.

1998 ACM Subject Classification J.4 Social and Behavioral Sciences

Keywords and phrases mindreading, false belief attribution, privileged access, minimal narratives

Digital Object Identifier 10.4230/OASICS.CMN.2014.88

1 Introduction

In this paper we present a theory concerning the acquisition of mental state terms and their usage in the explanation and prediction of behaviour. In order to see the broader context of our theory first we discuss the task analysis of the famous false belief test. This is important because one can interpret the passing of this verbal test as a manifestation of the possession of the mental term of belief. As we are interested in the acquisition of mental terms, in our task analysis we will focus on the bidirectional relationship between mindreading and language. Following this we present our theory of Quinian bootstrapping of the semantics of mental state words. We will see that the third stage of this bootstrapping process is the formation of a so-called folk functionalist theory of mental state terms. We describe an experiment that we did in order to examine the formation this folk functionalist theory. This experiment addressed the understanding of the notion of privileged access in children. We were curious whether children can explain and predict the behaviour of protagonists in minimal narratives; to this end we examined the usage of mental state terms in our subjects' interpretation of a character's behaviour in those narratives.

2 The task analysis of the false belief test

During the last thirty five years there has been an explosion of research into the naïve theory of mind (ToM) of primates, children and adults. At present, research in this area is conducted under various different labels within the cognitive sciences such as naïve psychology,



© Szabolcs Kiss and Zoltán Jakab;

licensed under Creative Commons License CC-BY

5th Workshop on Computational Models of Narrative (CMN'14).

Editors: Mark A. Finlayson, Jan Christoph Meister, and Emile G. Bruneau; pp. 88–105

OpenAccess Series in Informatics



OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

intuitive psychology, everyday psychology, common sense psychology, folk psychology, belief-desire psychology, natural psychology, interpretation, mentalization, metarepresentation and mindreading. In the present paper, we shall use these terms as roughly synonyms. By theory of mind we mean the basic ability to attribute or impute mental states (e.g., beliefs, desires, intentions, thoughts, emotions and so on) to ourselves and others in order to explain, predict, interpret and influence the behaviour.

In everyday life there are a lot of occasions where we read the minds of other people around us. One such situation is the classic false belief scenario. The famous false belief task was suggested by philosophers (e.g., [12]) in response to [17] question concerning the possibility of the chimpanzee's theory of mind. At that time, philosophers thought that the real criterion of theory of mind is the case when the organism can ascribe a mental state to the other which is different from its own. This is the situation in the classic change of location false belief task [24]. In this task, children watch the following scenario: Maxi puts his chocolate in location A in the kitchen and then leaves the scene. In his absence, his mother removes the chocolate from location A and puts it in location B. Then Maxi returns and the child is asked where he will search for his chocolate. The basic finding is that three-year-olds say that Maxi will look for his chocolate in the new B place while only four-year-olds can correctly indicate location A in their verbal responses.

Why do three-year-olds fail on this task and what are the cognitive requirements of passing this test at four years of age? In other words, what kind of task analysis can we provide for the success on this false belief test? At present, there are nearly twenty different cognitive explanations for these questions and the most important ones are listed below.

1. According to Leslie (e.g., [39]), humans have an innate theory of mind module which manifests itself in pretend play between 18-and 24-month-of-age. In Leslie's view, this innate mindreading module is not sufficient to pass the famous false belief test because the latter also requires the so-called *selection processor* which is responsible for inhibiting the reality-based response (i.e., that the chocolate is in location B). So three-year-olds have an intact theory of mind module but their selection processor does not yet work appropriately. The opposite is true in the case of children with autism: these children do not have a mindreading module but they possess a selection processor which is at work when these children pass the false photograph test. So children with autism can represent public, external representation such as an outdated photograph and their deficit is specifically with representing mental states. In other words, they have a domain-specific deficit.
2. According to Perner [29], passing the false belief test requires metarepresentational ability on the child's part. Perner explicitly commits himself to Fodor's Representational Theory of Mind [3] which simply holds that mental states are representational states. So when the child forms a representation about a representational state such as a belief, she constructs a metarepresentation. This is not enough, however, for managing on false belief tasks. Following Frege, Perner argues that in order to pass the false belief task the child needs to understand the distinction between the reference and sense of a representation. In his task analysis the four-year-old child understands that the reference of Maxi's representation is that the chocolate is in location B while the sense of his mental representation is that the chocolate is in location A. And the child believes that Maxi will act in accordance with his sense of representation.¹ According to Leslie (e.g., [39]), humans have an innate theory of mind module which manifests itself in pretend play between 18-and 24-month-of-age. In Leslie's view, this innate mindreading module is not sufficient to pass the famous false belief test because the latter also requires the so-called selection processor which is

responsible for inhibiting the reality-based response (i.e., that the chocolate is in location B). So three-year-olds have an intact theory of mind module but their selection processor does not yet work appropriately. The opposite is true in the case of children with autism: these children do not have a mindreading module but they possess a selection processor which is at work when these children pass the false photograph test. So children with autism can represent public, external representation such as an outdated photograph and their deficit is specifically with representing mental states. In other words, they have a *domain-specific* deficit.

3. Apperly [10] argues that the central feature of the false belief test is perspective-taking. In order to pass this test the child needs to adopt the perspective of Maxi and predict the searching behaviour on this basis. In other words, the child must overcome her egocentric bias that the chocolate is in location B.
4. According to Gopnik and Wellman [8], the false belief test requires the concept of belief. On their view, concepts are embedded within intuitive theories. A crucial feature of cognitive development is that these naïve theories change. In fact, this whole research field is called theory of mind which shows the influence of this particular, so-called theory-theory position.
5. A further analysis of the false belief test builds on the notion of executive functions. ‘Executive function’ is an umbrella term covering, among other things, the following notions: the inhibition of pre-potent responses, planning, the temporal organisation and monitoring of actions, sequencing behaviour, sustained attention, working memory, impulse control, etc. Within the false belief test the executive component is the inhibition of the reality-based response (the chocolate is in location B). This executive element is similar to the above discussed task of the selection processor.
6. It is important to realise that the false belief test eventually is a kind of narrative. So it should come as no surprise that if we highlight the story component of the false belief scenario then even three-year-olds are able to pass the task [11].
7. According to the task analysis of Frye, Zelazzo and Palfai [13], the false belief task can be broken down to understanding of conditionals. If Maxi wants to find his chocolate then he should search at location A. Indeed, these authors report significant correlations between the understanding of conditional statements and passing the false belief test.
8. De Villiers and de Villiers (e.g., [22]) provided experimental evidence to support the thesis that the acquisition of the syntax of the so-called complement-taking predicates is a cognitive prerequisite for passing the false belief test. Deaf children whose parents are not using any sign language are delayed at the false belief test.
9. According to Riggs, Peterson, Robinson, and Mitchell [30], there is a strong correlation between passing the false belief test and understanding of counterfactual statements. Within the false belief scenario the real situation is that the chocolate is in location B and the case that Maxi should look for location A is counterfactual with respect to this actual case.
10. We have seen above that in order to pass the false belief task children must possess the concept of belief. Where does this concept come from? According to Paul Harris [35], the child needs to participate in conversations and she can construct the notion of belief from these dialogues. So the experience in participating in conversations is a necessary precondition of passing the false belief test.
11. Both Paul Harris and Alvin Goldman [27] are committed to the so-called simulation theory of mindreading. According to them, the child can pass the false belief test via a kind of mental simulation in which she uses her own mind introspectively to predict the

protagonist's behaviour. In doing so, the child can rely on her imaginary identification with the protagonist.

12. The specific role of introspection in passing the false belief test is not restricted to simulation theory. Kuhn [15] also argues that self-observation plays a decisive role in the accomplishment of this task. On this view, the child forms the concept of belief via introspection.
13. In Fonagy's (e.g., [46]) view, there is a strong correlation between the security of attachment as measured by the strange situation in infancy and the age at which children can pass the false belief test. In fact, secure children can pass this task earlier than their insecure peers.
14. Astington [59] provides a social constructivist account of mindreading which is necessary for passing the false belief task. On this view, the concept of false belief emerges first at the interpersonal level and only later does it become interiorised into the individual's mind.
15. According to Fodor [4], folk psychology is innate and the basic inborn mindreading apparatus is present in the competence of the three-year-old child but at the same time she has performance limitations such as different cognitive heuristics which mask this competence. Moreover, Fodor suggests various hypothetical experiments in order to confirm his view.
16. There exists an explanation according to which the false belief test is a kind of meta-memory task. Notice, that in order to pass this task the child must attribute a memory to Maxi (i.e., He remembers where he put his chocolate at the beginning.)
17. A further explanation of the success on this test requires the notion of intention. The argument being simply that Maxi wants to find his chocolate. So this task involves the ascription of intention to the protagonist.
18. Dan Sperber (e.g., Sperber et al. [20]) argues that the false belief test can be approached as a task of epistemic vigilance. On this view, it is an important cognitive developmental achievement when the child gives up her basic trust that played an important role in early communication and begins to take into account the possibility of misleading information such as a lie or error.
19. Finally, Helming, Strickland and Jacob [32] argue that the classic false belief task is a normative task, that is, three-years-olds interpret the test question as "Where should Maxi look for his chocolate?" And the correct answer to this question is location B (since that is the actual location, and arguably it is a norm that one should look for something where it actually is).

By this quick review of positions we intend to illustrate why Bloom and German [43] are right when they suggest to abandon the false belief task as a test of theory of mind. On the one hand, the classic false belief test contain elements that are not specific to theory of mind (e.g., executive functions), and on the other, there are important mindreading developments before passing this task (e.g., understanding visual perspective). These different explanations are summarised in Table 1.

We collected the cognitive explanations of the verbal false belief task. Since 2005 we have a growing body of experimental evidence concerning the infants' ability to attribute false beliefs to others (e.g., Kovacs et al. [1]; Onishi and Baillargeon [34]) as it is demonstrated in various non-verbal tasks. These violations-of-expectations experiments are subject to different interpretations. One of them is that infants have an implicit theory of mind. Indeed, Onishi and Baillargeon argue that 15-month-old infants have a representational understanding of

■ **Table 1** Cognitive explanations of the false belief test

Authors	Explanations	Evidence
Leslie	Innate theory of mind module	Theoretical and experimental
Perner	Metarepresentation	Conceptual and experimental
Apperly	Perspective-taking	Experimental
Gopnik and Wellman	Conceptual change	Experimental
Russell	Executive functions	Experimental
Lewis	Narratives	Experimental
Kuhn	Introspection	Theoretical
Frye, Zelazo and Palfai	Understanding of conditionals	Experimental and correlational
De Villiers	The syntax of complement-taking predicates	Experimental
Harris	Conversations	Observations
Riggs	Understanding of counterfactuals	Experimental and observational
Fonagy	Secure attachment	Experimental and observational
Goldman	Simulation and imaginary identification with the protagonist	Philosophical and experimental
Astington	Social constructivism	Experimental
Fodor	Cognitive heuristics	Hypothetical experiments
	Metamemory	Conceptual
	Attribution of intention to the protagonist	Conceptual
Sperber	Epistemic vigilance	Experimental
Jacob	The test as a normative task	Conceptual

mind. Ruffman and Perner [57] provide an alternative, strictly behavioural explanation to this rich interpretation.

At present there are various attempts to explain the existing gap between early mindreading and later success on the verbal false belief task. For instance, one such attempt is Alan Leslie's theory of an innate mindreading module and his selection processor discussed above.

3 The relationship between social cognition and language

What is the connection between mindreading and language? The relationship between social cognition and language is in the focus of several researchers (see e.g., Astington and Baird, [60]). One view is that certain mindreading abilities are the cognitive prerequisites for acquiring language. Below, we offer a rough summary of these different theory of mind preconditions and the various aspects of language being explained by them.

1. According to Bloom [42], understanding of the speakers' referential intention is necessary for word learning. To put this into a broader perspective we can say that the acquisition of the mental lexicon is dependent upon one's theory of mind.
2. Baldwin (Baldwin and Moses, [2]) demonstrated experimentally that older infants can take into account the other's direction of gaze when they learn a new word. This means

that in a situation where there are more new objects present infants will attach the heard new label to that particular object to which the speaker paid attention rather than that the one they themselves looked at. So again learning a new word requires theory of mind ability; in this case the understanding of attention.

3. Tomasello [19] argues that the so-called nine-month-olds' revolution in understanding intentions is a cognitive prerequisite for language acquisition. At nine-month-of-age infants begin to demonstrate volitional behaviour and this is the base on which they can ascribe intentions to other people around them. Theoretically speaking, this is a kind of mental simulation by which the infants understand others.
4. Beckwith [56] claims that the emergence of the concept of other minds is a precondition for the formation of the so-called Experiencer thematic role. Thematic roles (such as Agent, Theme, Experiencer, Goal, etc.) are the semantic labels for various arguments in the predicate-arguments linguistic theories. In particular, the so-called psychological verbs (such as fear, love, hate) take the Experiencer role as one of their arguments.
5. Hamvas [21] presented experimental evidence showing significant correlation between passing the false belief test and detecting the violation of Gricean conversational maxims. So we can argue that the concept of false belief is required for certain pragmatic abilities.
6. Similarly, Happé [26] showed that the notion of false belief is necessary for understanding metaphors. In fact, she experimentally tested the basic tenets of Relevance theory (Sperber and Wilson, [18]) arguing for the mindreading basis of conversational pragmatics (see also Kiss, [53]).
7. Györi [38] also presented experimental evidence for his view that mindreading is a prerequisite for understanding irony.

These different positions are summarised in Table 2. But the above list is only one side of the coin. The other side is that many linguistic factors play a role in the emergence of theory of mind. Below, we summarize these linguistic prerequisites and the elements of mindreading that are explained by them. So we can conclude that the connection between mindreading as a micro-level phenomenon and language (as a macro-level phenomenon) is a bidirectional one.

1. Astington and Baird [60] collect various papers which all claim that language matters for theory of mind. Their book came out in the same year when Onishi and Baillargeon [34] reported their experimental findings concerning the infants' implicit theory of mind discussed above. So, of course this book cannot address the issue of the existence of a preverbal mindreading competence.
2. Within the task analysis of the well-known false belief test we saw the strong and provocative proposal by de Villiers and de Villiers [22] according to which the syntax of complementation is necessary for passing this test. So they view syntax as a linguistic precondition of mature theory of mind.
3. We have also discussed Harris's [35] theory that claims that conversations are inevitable for the mental construction of belief. Here again we can see that some kind of linguistic practice is responsible for certain social cognitive achievements.
4. O'Neill [33] argues that the linguistic given/new distinction is mandatory for the development of certain pragmatic skills. But it should be noted that since the pioneering work of Paul Grice pragmatics and mindreading are so closely connected that one can hardly disentangle these two basic competences.
5. In a relatively early work, Astington [58] dealt with the connection between narratives and the child's developing theory of mind. Of course, it is needless to say that the ascription of mental states to the protagonist is a central feature of all stories.

■ **Table 2** Social cognitive prerequisites for language

Authors	Special mindreading ability	What aspect of language is explained
Bloom	Understanding referential intention	Word learning
Baldwin	Gaze-following	Acquisition of the lexicon
Tomasello	Nine-month-olds' revolution in understanding intentions	Word learning
Beckwith	The emergence of the concept of other minds	Formation of the Experiencer thematic role
Kiss Sz.	Passing the false belief test	Relevance theory
Hamvas E.	Passing the false belief test	Grice: Conversational maxims
Happe	Passing the false belief test	Understanding metaphors
Györi M.	General theory of mind	Understanding irony

■ **Table 3** Linguistic prerequisites for social cognition.

Author	Linguistic prerequisites	What aspect of mindreading is explained
Astington and Baird	Language in general	Theory of mind
DeVilliers	Syntax	Passing the verbal false belief test
Harris	Conversations	The emergence of the concept of belief
O'Neill	The given/new distinction	Pragmatics
Astington	Narratives	Mindreading stories
Kiss	Lexical semantics	The meaning variance of mental terms

6. Kiss [54] describes the acquisition of the meaning of mental terms within the so-called theory of mind research field. (See also the next section of the present paper for more details on this.)

These views are presented in Table 3. In this section we looked at answers to the following general question: What is the exact role of language in the formation of theory of mind? As our primary interest in this paper is the acquisition of mental state terms, we would like to offer another five possible theoretical roles for language in ToM.

1. First, one could suggest that language in general, and the mental terms in particular, are merely labels for our independently existing concepts (Fodor, [5], [6]).
2. Second, we may assume that language is an “invitation to form categories”, thereby playing a facilitatory role
3. Third, R. Mitchell ([48], p. 41) says in the spirit of Wittgenstein ([36], par. 384): “In a sense, language creates mental states: “You learned the concept “‘pain’ when you learned language” Note however, that in this quotation Wittgenstein refers to concepts, but Mitchell is concerned with mental states in general.
4. Fourth, language may be taken as a replacement of expressive behaviour.
5. Fifth, language can be understood as a medium for representing mental states.

4 Quinian bootstrapping of mental terms

Now we would like to explore the role of Quinian bootstrapping in the acquisition of the meanings of mental terms such as “happy”, “think”, “believe” etc. Let us briefly introduce Quinian bootstrapping as developed by Carey ([50], [51]). From Carey’s analysis we pick out the importance of a placeholder structure. The placeholder structure consists of symbols whose meanings are initially determined in relation to each other. In this structure, many symbols are connected to each other and these connections are represented in long-term memory. At the beginning of the bootstrapping process the symbols have no meaning but later the so-called modelling processes give meaning to these symbols. According to Carey, these modelling processes can be analogy, inductive inference, thought experiment, abduction. Carey introduced Quinian bootstrapping in her explanation of the acquisition of numeral list representation and rational number as well as certain aspects of intuitive physics. Quinian bootstrapping makes possible the creation of new representational systems with novel concepts that were not available in the earlier conceptual machinery. In a word, Quinian bootstrapping underlies radical conceptual change both in the history of science and individual development.

How can we characterise the meaning of mental state terms? Folk functionalism is an intuitive theory in which the meanings of mental terms are organised. According to this common sense functionalist theory, the meaning of a mental term consists of the input and output conditions of the given psychological state as well as the mental state’s connections to other mental states. For instance, the meaning of the term pain consists of the cause of the pain (e.g. touching a hot stove), the pain’s relationship to other mental states (the desire to get rid of the pain) and the pain’s connections to behaviour (pain elicits wincing).

How does the child learn the meaning of mental terms? On the first stage of the ontogenetic acquisition of the semantics of mental terms the child already uses mental terms, but she is not fully aware of their meanings yet. At this stage the child uses mental terms referring to the behavioural components of the mental state only. For instance, the term happiness refers only to behavioural manifestations such as a smile. This phenomenon is called semi-successful reference by Beckwith [56]. This is consistent with Wittgenstein’s view according to which the attribution of mental states is always based on behavioural criteria. The phenomenon of semi-successful reference of mental terms is also in line with Wittgenstein’s well-known remark that we use words whose meanings become clear only later.

Clearly, in this case we can see the learning process of Quinian bootstrapping at work. One of the central components of this bootstrapping process is the existence of a placeholder structure (see above). According to Carey, the meaning of a placeholder structure is provided by relations among external, explicit symbols. In our case, these external, explicit symbols are mental words and expressions represented in the child’s long-term memory. So, the child represents many mental words and lexical items whose full and complete meanings become available only at later stages of this bootstrapping process.

Carey [51] asks one of us what kind of evidence do we have in order to support this first stage in the acquisition of mental terms? Here and now we can provide an anecdotal evidence for this stage. Once one of us observed a two-year-old girl who was clearly playing with her mother’s scarf and during this play she put it around her neck and said: “I have a sore throat.” This observation is consistent with the view that at the beginning the child uses external, behavioural criteria for attributing a mental state to herself. Furthermore, P. Mitchell et al. ([47], p. 329.) writes:” Indeed, children’s earliest use of mentalistic terms usually links with observable behaviour”.

At the second stage of the change of meaning of mental terms the child discovers the inner subjective component (feeling or qualia) of mental lexical items and she realises that the reference of mental terms includes this component. In other words, the child recognises the phenomenological or experiential qualities of mental terms. So within this recognition introspection serves as a kind of modelling process using Carey's terminology of Quinian bootstrapping. In this stage, mental terms have gone through *meaning variance* in relation to the first stage, but the child does not yet possess the full representation of the meaning of mental terms found in the folk functionalist theory (see [52] for more details).

The third stage is the acquisition of this common sense functionalist theory. It is the result of a long learning process during which the child comes to understand the relationship between mental states and their eliciting conditions and the interconnections of mental states to each other and to their behavioural consequences. This is the acquisition of a coherent theory by which the child understands specific causal processes such as the fact that perception leads to the fixation of beliefs, or that beliefs can bring about other beliefs by means of inference and that beliefs and desires cause actions together.

Later in this paper we present an experiment in which we tested the working of this folk functionalist theory in children from 5- to 7 years of age. In that experiment we were interested in how children can explain a protagonist's behaviour by citing mental states. In other words, we asked how children use mental terms in order to explain or predict behaviour.

In sum, mental terms go through changes of meaning during semantic development. The successive naïve psychological theories of children determine the meanings of mental terms. This meaning variance of mental terms is similar to the meaning variance of scientific terms discussed by philosophers of science (e.g., [45]). (In fact, we have borrowed the expression of meaning variance from this philosophical tradition.) As identical terms gain different meanings in different theories, the changes of meaning lead to the problem of incommensurability between various theories. In this way, we can extend the notion of incommensurability to the child's developing theories of mind as well (see [54]).

We have presented the theory of the meaning variance of mental terms in cognitive development. We briefly touched upon the role of introspection in the acquisition of the meaning of mental terms and at the same time we committed ourselves to the so-called theory-theory of mindreading. So the question remains what is the proper role of introspection in the development of theory of mind? As we said above, a crucial aspect of the Quinian bootstrapping of learning the semantics of mental terms is the so-called modelling process by which the mental terms represented in long-term memory get their meanings. And in our view, this modelling process is a kind of introspection. On the one hand, toddlers or young children have mental concepts (e.g., such as a belief); on the other, they have many mental terms in their memory. The cognitive task for the developing child is to find a mapping between the concepts and the mental terms. This mapping is achieved via self-observation. In a word, we do not hold the provocative view of Gopnik [7] according to which first-person self-knowledge is illusory or the less radical approach of Carruthers [44] which claims that the mind is not transparent to itself. As we said before introspection has an important role to play in the acquisition of the meaning of various mental terms.

In our view, the best approach to the development of mindreading is a complex of introspection-based simulation and theory-theory. Our commitment to theory-theory is a kind of semantic determination because we emphasize that both mental concepts and words are embedded in coherent intuitive theories. In addition, the role of introspection as discussed above leads us to accept simulation theory (e.g., [27]).

5 The cognitive developmental investigation of the attribution of privileged access to mental states

Experimental studies of privileged access originate from work by Jürgen Habermas. Within the framework of universal pragmatics developed by Habermas the notion of the ideal speech situation plays a crucial role. In our earlier work [53], we discussed this concept from the point of view of the empirical research on theory of mind in children. Here we would like to address one important aspect of the ideal speech situation, namely that we attribute privileged access to the other person concerning her intentional states. Furthermore, we assume that the other person is able to verbally report these mental states. In other words, we tend to assume that in this ideal situation the other subject has infallible inner eye for her psychological states. This in fact means that we attribute an ability of introspection, therefore first-person authority, to other people. The objects of introspection are intentional states, mental states, certain propositional attitudes (like belief, desire, etc.), psychological states, reasons, intentions, motives, thoughts, emotions, etc.

In this idealisation we do not assume inaccessible, unconscious mental states; instead we assume that the mind of the other is transparent to itself. It is important to see that this is a kind of accountability and eventually responsibility. However, this idealisation exists only in folk psychology because many theorists within psychology deny first-person authority. (See e.g. the whole school of psychoanalysis, [44, 7, 49, 16, 25], and the famous work of Wittgenstein).

What is the relationship between ToM and the attribution of privileged access? From the point of view of empirical research we would like to mention two points. First, within the research on theory of mind development different experiments focus on the questions of how and when children become able to recognise mental states as the reasons or causes of behaviour. These experiments presuppose that once the child comes to recognise that intentional states underlie actions she immediately imputes privileged access to those mental states.

Second, these experiments also presuppose that when the child interprets the other person's behaviour in terms of beliefs and desires, she also presumes that the verbal report of the other concerning the intentional states behind the behaviour would correspond to these mental states; that is to say the child does not take self-deception or confabulation into consideration at the beginning.

Gergely (personal communication) raises the possibility that intentional causation and the above-discussed idealisation are conceptually inseparable. First let us see the option when we take this to be the case. We can only say that this is probably true from a developmental point of view. One of the most important milestones within the research on social representation is the empirical work by Moscovici [55] cited by László [28] that reports the finding according to which Freudian concepts appear in the folk psychology of adults. From this research we know that the notion of the unconscious is present within the everyday reasoning about human action in adults. This implies that adults do not always assume accountability and the ideal speech situation. On the contrary, they tacitly assume that there may exist reasons or other intentional causes which, although play an important role in causing action, are not accessed in a privileged fashion (and hence are not reportable).

From this point of view it is a very important question how the child gives up the ideal speech situation and when she comes to suppose unconscious intentional reasons behind action. We do not know of any empirical data relevant to this issue. Flavell et al. [23] examined the understanding of the unconscious by children. According to their results, this

understanding develops in the early elementary school years but this investigation did not address the understanding of the unconscious causes of action in children. It seems that the abandonment of the ideal speech situation is the result of cultural learning.

Of course, according to the second theoretical option intentional causal understanding and idealisation are conceptually separate even in ontogenesis. Let us look at one aspect of this idealisation according to which the child presumes accountability of intentional actions. It is possible that children first understand that mental states cause behaviour and only later do they assume that others can verbally report on intentional states. Maybe the ideal speech situation in this sense requires the active participation in conversations. It is worth noting again that Paul Harris [35] argues that the understanding of belief in three-year-olds is due to the fact that these children already have enough active experiences of conversations. According to him, the two-year-old child forms the mental concept of desire due to her own agency. From the vantage point of the ideal speech situation two-year-old children could understand desire as a mental state but they do not assume that the other person is able to verbally report on it. At this age the child is not able to anticipate imaginary conversations and she could expect the ideal speech situation later only on the basis of internalisation of dialogues and the development of imagination.

In our earlier work [61] we studied the relationship between the attribution of privileged access and perspective taking. In the experiment reported below we used a different method to investigate the attribution of privileged access and by implication the ontogenetic emergence of the ideal speech situation. We place the exploration of the attribution of privileged access into the framework of experimental philosophy. It is worth mentioning that in the novel collection of traditional and experimental philosophical papers [9] naïve theory of mind plays an important role.

Our research questions are the following: When and how do children impute privileged access to others concerning their mental states? What is the relationship between ToM and the attribution of privileged access? In other words, is mindreading a necessary condition of the attribution of privileged access? Or to put it slightly differently, is theory of mind a prerequisite for the ascription of privileged access?

It is worth mentioning that P. Mitchell et al. [47] studied the emergence of privileged access in children and found that even five-year-olds assigned more self-knowledge to themselves than to an adult. This means that 5-year-olds have an understanding of first-person authority. This study, however, did not investigate the attribution of privileged access to others, only to oneself. We explored the attribution of privileged access through minimal narratives. In this experiment, we studied the relationship between mindreading, privileged access and understanding non-literary narratives. Today we witness many attempts to study the different aspects of this relationship. One particular and well worked-out example is the work of Hutto [14]. He argues that the most important part of common-sense psychology is its narrative nature. At the same time, the connection between theory of mind and literary narratives is also an essential research topic within contemporary cognitive science (see e.g. [37]). This relationship is a very important issue in contemporary cognitive narratology. The attribution of mental states (i.e. ToM.) and privileged access to the protagonist is a central feature of all stories. According to Bruner [41], there are two psychological landscapes in every story. The landscape of action consists of the arguments of an action, i.e. its actor, intention, goal, situation, etc. whereas the landscape of consciousness comprises the mental states (e.g. knowledge, belief, feeling, thinking) of the participants. In the experiments reported below, both landscapes played an important role. So far we have seen that understanding narratives has relevance for passing the famous false belief test and the connection between language and mindreading.

6 Experiment

6.1 Methods

Subjects. Seventy five children participated, in the experiment: 25 5-year-olds (mean age: 5:4, range 5:0 to 5:9; 14 boys and 11 girls), 25 6-year-olds (mean age: 6:6, range 6:0 to 6:11; 10 boys and 15 girls), and 25 7-year-olds (mean age: 7:3, range 6:0 to 9:0; 12 boys and 13 girls). The 5-year-olds and the 6-year-olds attended preschools in Pécs, Hungary and the 7-year-olds were first graders also in Pécs in Hungary.

Materials and procedure. We used two stories from the study of Bartsch and Wellman [31] and we added a privileged access question to each of them. The first story is the following: “Here is Jane. Jane is looking for her kitten. The kitten is hiding under the chair. But Jane is looking under the piano. Why do you think Jane is doing that?” This is a backward reasoning task because children have to infer backwards from the protagonist’s action to her underlying mental states. (E.g. Jane thinks that the kitten is under the piano.)

Following this, the corresponding privileged access question was: “What do you think, if we asked Jane why she is looking for her kitten under the piano what would she say?” This question relates to the possible verbal report concerning the mental states underlying Jane’s action.

The second story is the following: “Sam wants to find his puppy. His puppy might be hiding in the garage or under the porch. But Sam thinks his puppy is under the porch. Where will Sam look for his puppy, in the garage or under the porch?” This is a forward reasoning task because children have to infer forwards from the mental states of the protagonist to his action. (E.g. Sam will look for his puppy under the porch.) So the Jane’s story and the Sam’s story are different in terms of the direction of reasoning that is required from the child’s part.

The privileged access pair of the above question looked like this: “What do you think, if we asked Sam why he is looking for his puppy under the porch/garage what would he say?” Each story was accompanied by a colourful drawing in order to help children to understand the minimal narratives. The order of the two stories was counterbalanced among participants.

Coding. Originally we developed a coding system which consisted of eight categories. The different answer categories were the following: (1) Does not know; (2) Mental states (e.g., “She is looking for her kitten in the kitchen because she *thinks* it is there); (3) Further reasons for action (e.g., “... because she wants *to play with it*”); (4) The physical environment of the drawings (e.g., “... because the garage is bigger”); (5) Goal (e.g., “... because she wants *to find her kitten*”); (6) Reality answer (e.g., ... because her kitten *is under the piano*); (7) Perception (e.g., “... because the last time she saw her kitten it was there”); (8) Other.

Later we simplified our category system. As our primary interest is in mindreading and understanding the attribution of privileged access we divided children’s behavioural explanation responses into two categories. One of the categories relates to the explanation in terms of mental states (e.g., belief, thoughts etc.) and all the other accounts were grouped into the second category.

■ **Table 4** Percentages of the different answers to the first question in Jane’s story.

Answer categories	5-year-olds	6-year-olds	7-year-olds
Does not know	4	20	4
Mental states	52	16	44
Further reasons for action	0	0	4
Physical environment in the drawing	16	20	8
11 Goal	0	4	4
Mentioning location	20	16	12
Perception	0	16	12
Other	8	4	12

6.2 Results

The descriptive statistics of children’s answer in terms of the eight original categories is shown in Tables 4-7. As the data in this tables did not lead to significant results we reduced the categories into two as mentioned above.

Regarding our first story (the one with Jane as the main protagonist) 13 5-year-olds (52 percent) mentioned mental states in their behavioural explanation in response to the first question. Interestingly, only four of the 6-year-olds (16 percent) did the same in response to this question. Of the 7-year-olds eleven (44 percent) answered in terms of mental states to this question. In all age groups the remaining children cited other reasons in their response to this question. This result is significant (Chi-Square test: $p < .05$ ($p = 0.02$)) presumably because 6-year-olds mentioned mental states in their response less often.

In response to the privileged access question in the first story nine 5-year-olds (36 percent), two six-year-olds (8 percent), and nine seven-year-olds (36 percent) cited mental states. The remaining children did not make reference to psychological states in their answer to the privileged access question. This finding is again significant (Chi-Square test: $p < .05$) for the same reason as mentioned above.

Concerning our second story (the one with Sam as the main character) we made an analysis of the responses to the second privileged access question as a function of the correct/incorrect answer to the first question. Notice that in the story with Sam there exists a correct answer to the first question i.e. Sam will search for his puppy where he thinks it is namely under the porch. The data are presented in Table 8. This table shows that children who answered the first question incorrectly gave much less mental-state-based explanations to the privileged access question than those who answered it correctly (Chi-Square test: $p < .01$ ($p = 0.01$)).

6.3 Discussion

In response to the first question in the Jane story 6-year-olds mentioned mental states less often than did 5- and 7-year-olds. This is a surprising result which may be due to the fact that 5-year-olds mentioned the “She thinks ... or she believes ...” phrases very often. This is consistent with the general view of Wellman [40] according to whom a kind of explicit belief-desire psychology emerges at the age of 4 and 5. Maybe the emergence of this belief-desire psychology at that age means that this intuitive psychology becomes the dominant explanatory framework and it is only a later developmental achievement that the child considers other kinds of possible accounts as well. But this is only a partial explanation and further research is needed in order to explain this decrease of mental state explanations

■ **Table 5** Percentages of the different answers to the second question in Jane's story.

Answer categories	5-year-olds	6-year-olds	7-year-olds
Does not know	24	32	12
Mental states	36	8	36
Further reasons for action	12	0	0
Physical environment in the drawing	0	20	8
Goal	0	0	0
Mentioning location	24	16	20
Perception	0	20	8
Other	4	4	16

■ **Table 6** Percentages of the correct/incorrect answers to the first question in Sam's story.

	5-year-olds	6-year-olds	7-year-olds
Sam's first question answered correctly	72	68	92
Sam's first question answered incorrectly	28	32	8

in the 6-year-olds. In fact, first we need to repeat this result on another sample to firmly exclude the possibility of a sampling error.

A similar finding emerged in response to the privileged access question in the first Jane story. In response to this question we expected children to verbally report on Jane's mental state but we got relatively few mental state accounts and the general trend was from the children's part to rationalise Jane's behaviour. The decrease of mental state responses in the 6-year-olds can be explained as above.

Regarding our second story with Sam we can say that those children who could predict Sam's behaviour in terms of his mental state correctly gave significantly more mentalistic explanations to the second privileged access question. This finding seems to suggest that there is indeed a close conceptual connection between intuitive theory of mind (as assessed by our first question) and the ascription of privileged access. This is the theoretical option that we discussed in the introduction to our experiment, and further research is needed to separate these two components (i.e. theory of mind and the attribution of privileged access). That is, we have to examine younger children to assess this theoretical option.

In our experiment we tested the working of the folk functionalist theory by which children interpret, explain and predict the behaviour of others using mental state terms. As we saw above, 36 percent of the 5-year-old subjects cited mental states in their responses to the privileged access questions in both stories. We have also seen that according to Mitchell et al. [47] 5-year-olds have first-person authority. This raises the theoretical possibility that children at first have privileged access concerning their own mental states and they attribute privileged access to others only later. But this can mean that as soon as the child discovers that she has privileged access she becomes able generalize this principle to others. This was our original motivation to involve 5-to-7-year old subjects in our study: on the one hand it seems fairly clear that they can attribute privileged access to themselves, and we wanted to see how easily they proceed to understanding others' privileged access. Our results suggest that there may be some delay in this generalization.

■ **Table 7** Percentages of the different answers to the second question in Sam's story.

Answer categories	5-year-olds	6-year-olds	7-year-olds
Does not know	24	32	8
Mental states	36	24	36
Further reasons for action	4	12	12
Physical environment in the drawing	4	4	4
Goal	4	4	0
Mentioning location	16	20	28
Perception	0	4	4
Other	12	0	8

■ **Table 8** Responses to the privileged access question in the Sam story as a function of the answers to the first question: absolute number (per cent)

	Mental-state-based explanations	All other responses
First question correct	23 (31)	35 (47)
First question incorrect	1 (1.33)	16 (21)

Acknowledgements. We wish to thank Zsófia Bacskai, Dominika Czérna, Brigitta Hány, Adrienn Jörg, Tímea Kórizs, Petra Kovács and Anna Soós for their help in data collection, and three anonymous reviewers for their helpful comments on an earlier version of this paper.

References

- 1 Kovács Á., Téglás E., and Endress A. The social sense: Susceptibility to others' beliefs in human infants and adults. *Science*, 330(6012):1830–1834, 2010.
- 2 Baldwin D. A. and Moses L. J. Early understanding of referential intent and attentional focus: Evidence from language and emotion. In Ch. Lewis and P. Mitchell, editors, *Early understanding of mind*. Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1994.
- 3 Fodor J. A. Propositional attitudes. *The Monist*, 61(4):501–523, 1978.
- 4 Fodor J. A. A theory of the child's theory of mind. *Cognition*, 44(3):283–296, 1992.
- 5 Fodor J. A. *Concepts: Where cognitive science went wrong*. Oxford University Press, New York NY, 1998.
- 6 Fodor J. A. *LOT2: The language of thought revisited*. Oxford University Press, New York NY, 2008.
- 7 Gopnik A. How we know our minds: The illusion of first-person knowledge of intentionality. *Behavioural and Brain Sciences*, 16(1):1–14, 1995.
- 8 Gopnik A. and Wellman H. M. The theory theory. In L. A. Hirschfeld and S. A. Gelman, editors, *Mapping the mind. Domain specificity in cognition and culture*. Cambridge University Press, Cambridge, 1994.
- 9 F. Allhoff, R. Mallon, and S. Nichols, editors. *Philosophy. Traditional and experimental readings*. Oxford University Press, Oxford, 2013.
- 10 I. Apperly. *Mindreaders. The cognitive basis of "theory of mind"*. Psychology Press, Hove and New York, 2011.
- 11 Lewis C. Episodes, events, and narratives in the child's understanding of mind. In C. Lewis and P. Mitchell, editors, *Children's early understanding of mind: Origins and development*. Erlbaum, Hillsdale NJ, 1994.

- 12 Dennett D. Beliefs about beliefs. *Behavioural and Brain Sciences*, 1(4):568–570, 1978.
- 13 Frye D., Zelazo P. D., and Palfai T. Theory of mind and rule-based reasoning. *Cognitive Development*, 10(4):483–527, 1995.
- 14 Hutto D. D. *Folk psychological narratives. The sociocultural basis of understanding reasons*. MIT Press, Cambridge MA, 2008.
- 15 Kuhn D. Theory of mind, metacognition, and reasoning: A life-span perspective. In P. Mitchell and K. J. Riggs, editors, *Children's reasoning and the mind*. Psychology Press, Taylor and Francis Group, 2000.
- 16 Montgomery D. Wittgenstein's private language argument and children's understanding of the mind. *Developmental Review*, 17(3):291–320, 1997.
- 17 Premack D. and Woodruff G. Does the chimpanzee have a theory of mind? *Behavioural and Brain Sciences*, 1(4):515–526, 1978.
- 18 Sperber D. *Relevance. Communication and cognition*. Blackwell, Oxford, 1986.
- 19 Sperber D. *The cultural origins of human cognition*. Harvard University Press, Cambridge MA, 1999.
- 20 Sperber D., Clement F., Heintz C., Mascaro O., Mercier H., Origg G., and Wilson D. Epistemic vigilance. *Mind and Language*, 25(4):359–393, 2010.
- 21 Hamvas E. Do children notice the violations of conversational maxims?, 2000. Manuscript, in Hungarian.
- 22 De Villiers J. G. and De Villiers P. A. Language for thought: Coming to understand false beliefs. In D. Gentner and S. Goldin-Meadow, editors, *Language in mind*. MIT Press, Cambridge MA, 2003.
- 23 Flavell J. H., Green F. L., Flavell E. R., and N. T. Lin. Development of children's knowledge about unconsciousness. *Child Development*, 70(2):396–412, 1999.
- 24 Wimmer H. and Perner J. Beliefs about beliefs: Representations and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1):103–128, 1983.
- 25 Wimmer H. and Hartl M. The cartesian view and the theory view of mind: Developmental evidence from understanding false belief in self and other. *British Journal of Developmental Psychology*, 9(1):125–128, 1991.
- 26 F. G. E. Happé. Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition*, 48(2):101–119, 1993.
- 27 Goldman A. I. *Simulating minds. The philosophy, psychology, and neuroscience of mindreading*. Oxford University Press, New York NY, 2006.
- 28 László J. *The science of stories. An introduction to narrative psychology*. Routledge, London, 2008.
- 29 Perner J. *Understanding the representational mind*. MIT Press, Cambridge MA, 1991.
- 30 Riggs K. J., Peterson D. M., Robinson E. J., and Mitchell P. Are errors in false belief tasks symptomatic of a broader difficulty with counterfactuality? *Cognitive Development*, 13(1):73–90, 1998.
- 31 Bartsch K. and Wellman H. Young children's attribution of action to beliefs and desires. *Child Development*, 60(4):946–964, 1989.
- 32 Helming K., Strickland B., and Jacob P. Making sense of early false belief understanding. *Trends in Cognitive Sciences*, 18(4):167–170, 2014.
- 33 O'Neill D. K. Talking about "new" information: The given/new distinction and children's developing theory of mind. In J. W. Astington and J. A. Baird, editors, *Why language matters for theory of mind*. Oxford University Press, New York NY, 2005.
- 34 Onishi K. and Baillargeon R. Do 15-month-old infants understand false beliefs? *Science*, 308(5719):255–258, 2005.

- 35 Harris P. L. Desires, beliefs and language. In P. Carruthers and P.K. Smith, editors, *Theories of theories of mind*. Cambridge University Press, Cambridge, 1996.
- 36 Wittgenstein L. *Philosophical investigations*. The MacMillan Co., New York NY, 1965.
- 37 Zunshine L. *Why we read fiction. Theory of mind and the novel*. The Ohio State University Press, Columbus OH, 2006.
- 38 Györi M. *Autism and cognitive architecture. Domain specificity and psychological theorising on autism*. Akadémiai, Budapest, 2006.
- 39 Leslie A. M. The theory of mind impairment in autism: Evidence for a modular theory of development? In A. Whiten, editor, *Natural theories of mind: Evolution, development and simulation of everyday mindreading*. Basil Blackwell, Oxford, 1991.
- 40 Wellman H. M. *The child's theory of mind*. MIT Press, Cambridge MA, 1990.
- 41 Bloom P. *Actual Minds, Possible Worlds*. Harvard University Press, Cambridge MA, 1986.
- 42 Bloom P. *How children learn the meanings of words*. MIT Press, Cambridge MA, 2000.
- 43 Bloom P. and German T. P. Two reasons to abandon the false belief task as a test of theory of mind. *Cognition*, 77(1):825–831, 2000.
- 44 Carruthers P. *The opacity of mind. An integrative theory of self-knowledge*. Oxford University Press, Oxford, 2011.
- 45 Feyerabend P. Explanation, reduction and empiricism. In H. Feigl and G. Maxwell, editors, *Scientific explanation, space and time*. University of Minnesota Press, Minneapolis MN, 1962.
- 46 Fonagy P. and Target M. Attachment and reflective function. their role in self-organisation. *Development and Psychopathology*, 9(0):679–700, 1997.
- 47 Mitchell P., Teucher U., Bennett M., Ziegler F., and Wyton R. Do children start out thinking they don't know their own minds? *Mind and Language*, 24(3):328–346, 2009.
- 48 Mitchell R. A proposal for the development of a mental vocabulary, with special reference to pretense and false belief. In Mitchell P. and Riggs K. J., editors, *Children's reasoning and the mind*. Psychology Press, Sussex UK, 2000.
- 49 Nisbett R. and Wilson T. Telling more than we can know; verbal reports on mental processes. *Psychological Review*, 84(3):231–295, 1977.
- 50 Carey S. *The Origin of Concepts*. Oxford University Press, New York NY, 2009.
- 51 Carey S. Précis of the origin of concepts. *Behavioural and Brain Sciences*, 34(3):113–167, 2011.
- 52 Kiss S. The problem of meaning variance of mental terms in cognitive development. In L. I. Komlósi, P. Houtlosser, and M. Leezenberg, editors, *Communication and culture. Argumentative, cognitive and linguistic perspectives*. Sic Sat, Amsterdam, 2003.
- 53 Kiss S. *Mindreading*. New Mandate, Budapest, 2005.
- 54 Kiss S. The notion of incommensurability can be extended to the child's developing theories of mind as well. *Behavioural and Brain Sciences*, 34(3):134–135, 2011.
- 55 Moscovici S. *La psychoanalyse, son image et son public*. Presses Universitaires de France, Paris, 1961.
- 56 Beckwith R. T. The language of emotion, the emotions and nominalist bootstrapping. In D. Frye and Ch. Moore, editors, *Children's theories of mind*. Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1991.
- 57 Ruffinan T. and Perner J. Do infants really understand false belief? *Trends in Cognitive Sciences*, 9(10):462–463, 2005.
- 58 Astington J. W. Narrative and the child's theory of mind. In B. K. Britton and A.D. Pellegrini, editors, *Narrative thought and narrative language*. Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1990.

- 59 Astington J. W. What is theoretical about the child's theory of mind?: a vygotskian view of its development. In P. Carruthers and P.K. Smith, editors, *Theories of theories of mind*. Cambridge University Press, Cambridge, 1996.
- 60 Astington J. W. and Baird J. A., editors. *Why language matters for theory of mind*. Oxford University Press, Oxford, 2005.
- 61 Jakab Z. and Kiss S. Privileged access and perspective taking, 2014. Manuscript, in Hungarian.