Recursive belief manipulation and second-order false-beliefs

Torben Braüner (torben@ruc.dk), Patrick Blackburn (patrickb@ruc.dk), and Irina Polyanskaya (irinap@ruc.dk)
Roskilde University, Denmark

Abstract

The literature on first-order false-belief is extensive, but less is known about the second-order case. The attainment of second-order false-belief mastery seems to mark a cognitively significant stage, but what is its status? Is it an example of *complexity only* development, or does it indicate that a more fundamental *conceptual change* has taken place? In this paper we extend Braüner's hybrid-logical analysis of first-order false-belief tasks (Braüner, 2014, 2015) to the second-order case, and argue that our analysis supports a version of the conceptual change position.

Keywords: Second-order false-belief tasks; hybrid logic; natural deduction; complexity only; conceptual change; belief formation; belief manipulation; recursion

Introduction

First-order false-belief tasks are a widely studied family of reasoning tasks in cognitive and developmental psychology. A well known example is the *Sally-Anne task*:

> *A child is shown a scene with two doll protagonists, Sally and Anne, having respectively a basket and a box. Sally first places a marble into her basket. Then Sally leaves the scene, and in her absence, Anne moves the marble and puts it in her box. Then Sally returns, and the child is asked: "Where will Sally look for her marble?"*

Children above the age of four typically handle this task correctly: they say that Sally will look in the basket, which is where Sally (falsely) believes the marble to be. Younger children, on the other hand, say that Sally will look in the box: this is indeed where the marble is, but this information is not available to Sally and hence the response is incorrect. For children with *Autism Spectrum Disorder (ASD)*, the shift to correct responses usually occurs at a later age.

The attainment of first-order false-belief mastery is a milestone in the acquisition of *Theory of Mind (ToM)*, the capacity to ascribe mental states such as beliefs to oneself and others, and some researchers account for ASD using some form of a *ToM deficit hypothesis*; see (Baron-Cohen, 1995). Many first-order false-belief tasks have been devised, and over the past 35 years both correlational and training studies (involving both typically developing and children with ASD) have yielded robust results (concerning, for example, the link between ToM and linguistic abilities; see the meta-study (Milligan, Astington, & Dack, 2007)).

Second-order false-belief tasks are less well studied. Consider the following version of the *second-order Sally-Anne task* (the bold font highlights the new text added to the first-order version just given; the bracketed [Sally will] marks the shift in word order from 'will Sally'):

> *A child is shown a scene with two doll protagonists, Sally and Anne, having respectively a basket and a box. Sally first places a marble into her basket. Then Sally leaves the scene, and in her absence, Anne moves the marble and puts it in her box.* **However, although Anne does not realise this, Sally is peeking through the window and sees what Anne is doing.** *Then Sally returns, and the child is asked: "Where* **does Anne think that** *[Sally will] look for her marble?"*

Again there is a transition age. Children above the age of six typically handle this task correctly: they respond that Anne thinks that Sally will look in the basket, which is where Anne (falsely) believes that Sally believes the marble to be. Younger children respond that Anne thinks that Sally will look in the box: this is where Sally knows the marble to be, but Anne does not know that Sally knows this and hence the response is incorrect. For children with ASD, the shift tends to occur at a later age, if indeed it happens at all.

The mastery of second-order false-beliefs is another key step in the acquisition of ToM, but less is known about it and many conclusions are tentative (Miller, 2009, 2012). The direction of the causal relation between second-order false-belief task competence and linguistic competence remains unclear, and though there have been second-order correlation studies on children with ASD, we know of no linguistic training studies on this population investigating the link between language and second-order false beliefs. Finally — the issue we address — there is no consensus on the status of the shift from first-order to second-order mastery. Starting with (Sullivan, Zaitchik, & Tager-Flusberg, 1994), some researchers have viewed it as a reasonably straightforward addition to first-order mastery: the acquisition of second-order mastery occurs when the child has sufficiently strengthened his or her information processing capacities, such as working memory and sequencing; following (Miller, 2009, 2012) we call this the *complexity only* position. Other researchers, starting with (Perner & Wimmer, 1985), have argued that the transition marks a more fundamental cognitive shift; again we follow Miller and call this the *conceptual change* position.

The bulk of our paper is theoretical. We extend Braüner's (Braüner, 2014, 2015) hybrid-logical analysis of first-order false-belief tasks to the second-order case, and argue that our analysis lends weight to a version of the conceptual change position. But the backdrop to our theoretical work is an ongoing training study on Danish speaking children with ASD which investigates whether training in linguistic recursion can lead to improvement in second-order false-belief mastery. We briefly link the theoretical analysis with our experimental work in the concluding discussion.

Logical analyses of false-belief tasks

There have been few previous applications of logical methods to false-belief tasks. The pioneering work is due to

Stenning and Van Lambalgen, who analyse the Sally-Anne and other first-order false-belief tasks using non-monotonic closed world reasoning (Stenning & van Lambalgen, 2008). The first-order Sally-Anne task has also been formalized using an interactive theorem prover for a many-sorted first-order modal logic (Arkoudas & Bringsjord, 2008). Applications of logical models to second-order false-belief tasks are even rarer: the clearest example is the use of Dynamic Epistemic Logic in (Bolander, 2014), though the use of game theory in (Szymanik, Meijering, & Verbrugge, 2013) to investigate performance in higher-order social reasoning is also relevant.

Our approach differs from these papers in a key respect: it takes the notion of *perspective shift* as fundamental. The work just cited formalizes false-belief reasoning from a global perspective; the hybrid logical approach of (Braüner, 2014, 2015) formalizes the local shifts of perspective required by the experimental subject when reasoning about the agents in the scenario (that is, Sally and Anne). The intuition is this. Correctly handling the first-order Sally-Anne task seems to involve taking the perspective of another agent, namely Sally, and reasoning about what she believes. So to speak, you have to put yourself in Sally's shoes.[1] In the following two sections we explain why *natural deduction* in *hybrid modal logic* enables us to formalize the reasoning underlying the first-order Sally-Anne task perspectively.
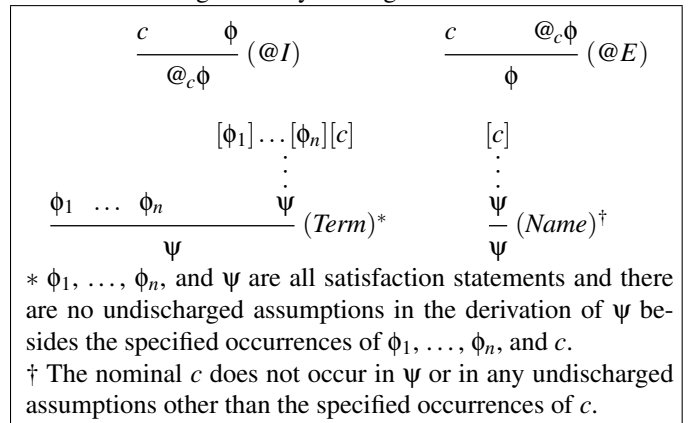
Hybrid modal logic

Modal logics are a family of logics in which the truth-value of logical formulas is evaluated relative to models consisting of propositional information distributed over a collection of abstract points; in Braüner's approach to false-belief tasks, these abstract points are taken to be people, and the information distributed over them is their beliefs, what they see, and so on. Modal formulas are evaluated locally: truth-value evaluation begins from the perspective of one particular person and then goes on to take into account other people's perspectives.

Modal logic makes use of various (application dependent) modal operators; here we will use such operators as B (to express that someone believes something) and S (to express that someone sees something). For example, $B\phi$ expresses that someone believes the information ϕ (whatever that might be). Modal operators can be applied recursively: $BB\phi$ expresses that someone believes the information that $B\phi$, that is, someone believes that someone believes ϕ.

Hybrid modal logics are extended modal logics in which it is possible to refer to individual points. Hybrid logic is crucial to this paper as it gives us the tools needed to refer to Sally and Anne, and to model their reasoning. Technically, hybrid logic is built by extending ordinary modal logic with *nominals*, which are special propositional symbols interpreted to

---

Figure 1: Hybrid-logical rules

$$\frac{c \qquad \phi}{@_c\phi}\ (@I) \qquad\qquad \frac{c \qquad @_c\phi}{\phi}\ (@E)$$

$$\frac{\phi_1 \ \cdots \ \phi_n \qquad\quad \overset{[\phi_1]\ldots[\phi_n][c]}{\vdots}\ \psi}{\psi}\ (Term)^* \qquad\qquad \frac{\overset{[c]}{\vdots}\ \psi}{\psi}\ (Name)^\dagger$$

* $\phi_1, \ldots, \phi_n$, and ψ are all satisfaction statements and there are no undischarged assumptions in the derivation of ψ besides the specified occurrences of $\phi_1, \ldots, \phi_n$, and c.

† The nominal c does not occur in ψ or in any undischarged assumptions other than the specified occurrences of c.

---

be true at exactly one point (so here, a person). Nominals should be thought of as naming the unique person they are true at. For example, we shall use s as a nominal true at Sally; in effect it is a 'name' or 'constant' that picks out Sally.

One other piece of hybrid logic machinery will be important: *satisfaction operators*. If ϕ is an arbitrary formula and s is the nominal that names Sally, then a new formula $@_s\phi$ can be built. The $@_s$ prefix is called a satisfaction operator, and the formula $@_s\phi$ is called a *satisfaction statement*. The satisfaction statement $@_s\phi$ says that the formula ϕ is true at a particular person, namely Sally, as she is the person s refers to. We call the nominal (here s) used in satisfaction statements the *point of view* nominal.

Let's make this just a little more concrete. Prefixing the belief formula $B\phi$ with the satisfaction operator $@_s$ yields the formula $@_sB\phi$ which says that Sally believes the information ϕ. Satisfaction statements of this form will play an important role in our analysis.

These informal examples should help guide the reader in what follows. For more on hybrid logic see (Blackburn, 2000) and (Braüner, 2011).

Seligman's natural deduction system

Logicians have developed many (very different) methods of formal proof; in this paper we make us of the one called *natural deduction*. Natural deduction was originally developed to model the structure of mathematical argumentation, but there is now some experimental backing for the claim that natural deduction is a mechanism underlying human deductive reasoning more generally; see (Rips, 2008). Be that as it may, we shall argue that the change in proof structure required as we move from modelling the first-order Sally-Anne task to modelling the second-order version supports the view that a conceptual change takes place in the developmental shift from first- to second-order false-belief mastery.[2]

---

[1]This might suggest that we are adopting the simulation-theory view of ToM, cf. (Gordon, 2009). But we are agnostic here: 'perspective taking' may seem reminiscent of 'simulation', but their relationship (if any) is unclear. Our talk of putting ourselves in someone else's shoes should be read in a pre-theoretic sense: it expresses an intuition that we wish to formally model in hybrid logic.

[2]So, broadly speaking, our work falls into the "mental logic" approach in the psychology of reasoning. The major alternative is the "mental models" approach, which views the construction of models as the mechanism underlying human reasoning. We earlier

Two main ideas drive natural deduction. The first is that there are two different kinds of rule for each logical connective: one to introduce it, the other to eliminate it. The second, crucial to this paper, is that *hypothetical reasoning* (or *conditional reasoning*) is hardwired into the very core of natural deduction: we can make an assumption, work out its consequences, and then discharge it (get rid of it).[3] We use the natural deduction system for hybrid logic obtained by extending the standard natural deduction system for classical propositional logic with the rules in Figure 1; the symbol c in these rules is an arbitrary nominal (that is, the name of a person).[4]

The rules $@I$ and $@E$ in Figure 1 are the introduction and elimination rules for satisfaction operators. The $@I$ rule says that if we have the information c (so we are reasoning about the person called c) and we also have the information ϕ, then we can introduce the satisfaction operator $@_c$ and conclude $@_c\phi$, which says that ϕ holds from c's perspective. The $@E$ rule says: suppose that when reasoning about the person named c, we also have the information that $@_c\phi$. Then we can eliminate the satisfaction operator $@_c$ and conclude ϕ.

But it is the *Term* rule that drives our analysis. This rule lets us switch to another person's perspective using hypothetical reasoning: the bracketed expressions $[\phi_1]\ldots[\phi_n][c]$ in the statement of the rule are (discharged) assumptions. The key assumption is c, which can be glossed as: let's switch perspective and temporarily adopt c's point of view. Incidentally, when using the *Term* rule we must make at least one assumption c, but we can make several, and this is often necessary to drive the proof through. The remaining (discharged) assumptions $[\phi_1]\ldots[\phi_n]$ in the rule's statement are additional assumptions we may wish to make about the information available from c's perspective.[5]

The rule works as follows. Suppose that on the basis of assumptions $\phi_1\ldots\phi_n, c$ we deduce ψ from c's perspective. Then the *Term* rule tells us that if $\phi_1\ldots\phi_n$ are available in the original perspective,[6] then we can discharge the assumption (which we do by bracketing them, thus obtaining the $[\phi_1]\ldots[\phi_n][c]$ displayed in the statement of the rule) and con-

clude ψ unconditionally in the original perspective.

The *Term* rule is subtle and powerful.[7] Indeed, as we shall now see, the hybrid logical analysis of the first-order Sally-Anne task boils down to a single application of *Term*, namely one modelling the shift from the experimental subject's perspective to Sally's.

Formalizing the first-order Sally-Anne task

In this section we describe how (Braüner, 2014) formalizes correct reasoning in the Sally-Anne task. Let us call the child performing the task (the experimental subject) Peter. Let t_0, t_1 and t_2 be three successive times where t_0 is the time at which Sally leaves the scene, t_1 is the time at which the marble is moved to the box, and t_2 is the time after Sally has returned. Here's the intuition we wish to model. To answer the question, Peter views matters from Sally's perspective and reasons as follows. At the time t_0, when Sally leaves, she believes that the marble is in the basket since she sees that it is. As she sees no action to move it, when she is away at t_1 she also believes the marble is in the basket. At t_2, after she has returned, she still believes that the marble is in the basket since she was out of the room when Anne moved it at the time t_1. So Peter concludes that Sally believes that the marble is in the basket.

In our formalization we make use of the predicates $l(i,t)$ and $m(t)$ and the modal operators B and S mentioned above. The argument i in the predicate $l(i,t)$ denotes a location, and the argument t in $l(i,t)$ and $m(t)$ denotes a timepoint. We take time to be discrete, and use $t+1$ to denote the successor of t.

$l(i,t)$	The marble is at location i at time t
$m(t)$	The marble is moved at time t
S	Sees that ...
B	Believes that ...
s	Nominal naming Sally

We also make use of four *belief formation* principles:[8]

(D)	$B\neg\phi \rightarrow \neg B\phi$
(P1)	$S\phi \rightarrow B\phi$
(P2)	$Bl(i,t) \wedge \neg Bm(t) \rightarrow Bl(i,t+1)$
(P3)	$\neg Sm(t) \rightarrow \neg Bm(t)$

Principle (D) is a common modal principle which says if we believe something to be false, then we don't believe it.

Principle (P1) states that beliefs may be formed as a result of seeing: that is, seeing leads to knowing. This is principle (9.2) in (Stenning & van Lambalgen, 2008), page 251.

Principle (P2) is reminiscent of both principle (9.11) in (Stenning & van Lambalgen, 2008), page 253, and axiom $[A_5]$ in (Arkoudas & Bringsjord, 2008), page 20. Principle (P2) formalizes a "principle of inertia" saying that a belief that the predicate l is true is preserved over time, unless it is believed that an action has taken place causing the predicate

---

mentioned (Bolander, 2014) which analyses first- and second-order false-belief tasks in Dynamic Epistemic Logic. As this is a formalism designed for a (sophisticated form of) model manipulation, this work is probably best classified as a mental models approach. There are also more practically-motivated intermediate approaches, for example the POLYSCHEME agent architecture (Kurup, Bignoli, Scally, & Cassimatis, 2011), which integrates multiple representations and has both rule-based and model-based features.

[3]For a full formulation of the discharge mechanism, we refer the reader to (Braüner, 2011).

[4]This is a modified version of Seligman's original natural deduction system for hybrid logic (Seligman, 1997) and is taken from Chapter 4 of (Braüner, 2011). We omit the rules for propositional connectives such as $\wedge$, $\rightarrow$ and $\neg$ as (a) they are standard and (b) we prefer the more perspicuous proof tree obtained by 'hiding' the simple propositional reasoning involved in the Sally-Anne task.

[5]The *Name* rule tells us that if we can prove the information ϕ by adopting some *arbitrary* perspective c, then ϕ also holds from the original perspective. As we won't use this rule in our analysis, we refer once more to (Braüner, 2011) for further discssion.

[6]Indicated by the premisses $\phi_1\ldots\phi_n$ listed just above the horizontal line in the statement of *Term* given in Figure 1.

[7]A subtlety worth emphasizing is that (as is stated in Figure 1) the assumptions $[\phi_1]\ldots[\phi_n]$ must all be satisfaction statements, otherwise the rule is not sound. We refer the reader to (Seligman, 1997) and Chapter 4 of (Braüner, 2011) for further discussion.

[8]This terminology is from (Stenning & van Lambalgen, 2008). The distinction they draw between belief formation and belief manipulation is central to our paper, and we shall return to it shortly.

Figure 2: Formalization of the child's correct response in the first-order Sally-Anne task

$$
\cfrac{
  \cfrac{
    \cfrac{[s]\ [@_sSl(basket,t_0)]}{Sl(basket,t_0)}\ (@E)
  }{Bl(basket,t_0)}\ (P1)
  \qquad
  \cfrac{
    \cfrac{
      \cfrac{
        \cfrac{[s]\ [@_sS\neg m(t_0)]}{S\neg m(t_0)}\ (@E)
      }{B\neg m(t_0)}\ (P1)
    }{\neg Bm(t_0)}\ (D)
  }{}\ (P2)
}{}
$$

$$
\cfrac{
  Bl(basket,t_1)
  \qquad
  \cfrac{
    \cfrac{
      \cfrac{[s]\ [@_s\neg Sm(t_1)]}{\neg Sm(t_1)}\ (@E)
    }{\neg Bm(t_1)}\ (P3)
  }{}\ (P2)
}{
  \cfrac{
    \cfrac{[s]\quad Bl(basket,t_2)}{@_sBl(basket,t_2)}\ (@I)
  }{}
}
$$

$$
\cfrac{@_sSl(basket,t_0)\quad @_sS\neg m(t_0)\quad @_s\neg Sm(t_1)\qquad\qquad @_sBl(basket,t_2)}{@_sBl(basket,t_2)}\ (Term)
$$

to be false.[9]

Principle (P3) encodes the constraint that if someone didn't see the marble being moved, then they can't have come to believe that it moved. Obviously this is not generally true, but the point of the formalization is to capture how Peter reasons about the Sally-Anne scenario (*Well, she can't have seen it, so I guess she won't believe it*).

The perspectival reasoning involved in the Sally-Anne task can be formalized as the derivation in Figure 2. We have already given Peter's informal perspectival reasoning; the formal proof mirrors it in full detail using a single application of *Term* in which the assumptions about s model the shift to Sally's perspective. The first two premises $@_sSl(basket,t_0)$ and $@_sS\neg m(t_0)$ taken together say that Sally at the earlier time t_0 saw that the marble was in the basket and that no action was taken to move it. The third premise, $@_s\neg Sm(t_1)$, says that Sally did not see the marble being moved at the time t_1 (since she was absent). Note that when applying the belief formation principles, we simply use them as rules. For example, when applying $P1$ (that is $S\phi \to B\phi$, or seeing leads to knowing) we simply use it to license the move from $S\phi$ on one line to $B\phi$ on the next.[10]

From belief formation to belief manipulation

The proof tree just discussed may seem complex, but it has a simple structure. The bulk of the reasoning on the (rather messy) right-hand-side of the proof tree consists of correctly sequencing applications of belief formation principles until the crucial formula $@_sBl(basket,t_2)$ — *Sally believes the ball is in the basket* — is deduced. What turns this into a formalisation of correct reasoning in the Sally-Anne task is the way the sequencing of belief formation principles is perspectivized. The right-hand-side sequencing occurs between the

initial assumptions of s (which perspectives it as *Sally's* reasoning) and the final application of *Term* which lets us conclude that the crucial formula is also true from *Peter's* point of view. In short, the analysis consists of *Belief Formation + Perspectival Reasoning* correctly combined.

Analagous remarks are made by Stenning and Van Lambalgen about their own analysis of first-order false-beliefs tasks; see (Stenning & van Lambalgen, 2008), page 257. They note that the bulk of the reasoning involves belief formation principles and their analysis succeeds because it is carried out using closed world reasoning; we might summarise their approach as *Belief formation + Closed World Reasoning* correctly combined. However they then go on to remark that what they call *Belief Manipulation* rules (which codify how to reason from one belief state to another) are unnecessary.

As far as first-order false-belief reasoning is concerned, we agree completely. Indeed, until now we have provided no proof rules for manipulating the belief operator B beyond the belief formation principles. And that is because, to model the first-order Sally-Anne task, we had no need of anything else. But belief manipulation rules will be needed if we are to extend our perspectival analysis to the second-order Sally-Anne task.[11] We turn to this task now.

Formalizing the second-order Sally-Anne task

First an observation. Peter (the experimental subject) has the same beliefs about Sally in the *first-order* Sally-Anne task as Anne has about Sally in the *second-order* task. For example, in the first-order task Peter believes that Sally does not know the ball has moved (and he's right) whereas in the second-order task Anne analogously believes that Sally does not know the ball has moved. Anne is wrong about this (Sally peeked through the window) but the fact remains that her beliefs about Sally match those of Peter in the simpler task.

This suggests that we should extend the first-order analysis by (a) using our first-order analysis as an analysis of *Anne's reasoning about Sally* in the second-order task and (b) embedding this pre-existing proof into a larger proof which captures

---

[9]For a discussion of how the inertia principle is applied — and just as importantly, how it is *not* applied — in the second-order Sally Anne task, see (Braüner, Blackburn, & Polyanskaya, 2016).

[10]As we remarked earlier, we do this to 'hide' the simple propositional reasoning involved. Strictly speaking, deducing $B\phi$ from $S\phi$ requires us to apply the propositional rule of *modus ponens* to $S\phi \to B\phi$. Using the belief formation principles as additional natural deduction rules enables us to omit such steps and reduce the size of the proof tree.

[11]Stenning and Van Lambalgen do not analyse second-order false-belief tasks.

Figure 3: Belief manipulation rule for the *B* operator

$$
\cfrac{B\phi_1 \quad \ldots \quad B\phi_n \qquad \begin{array}{c} [\phi_1] \ldots [\phi_n] \\ \vdots \\ \psi \end{array}}{B\psi} \ (\text{BM})^*
$$

∗ There are no undischarged assumptions in the derivation of ψ besides the specified occurrences of $\phi_1, \ldots, \phi_n$.

Peter's reasoning about Anne in the second-order task. That is, we should add another level of nesting to the perspectival analysis. Making this work requires us to introduce a recursive belief manipulation rule for *B*.

We have chosen the rule given in Figure 3. We call it BM. It is a version of a rule from (Fitting, 2007) that fits naturally with our tree-style natural deduction proofs.[12] We will also make use of the following additional notation:

D Deduces that ...

a Nominal naming Anne

and of a natural deduction formulation of the belief formation principle

($P0$) $D\phi \to B\phi$

which says that if we can deduce the information ϕ then we believe ϕ. This is principle (9.4) in (Stenning & van Lambalgen, 2008), page 251.

With this machinery in place, the reasoning in the second-order Sally-Anne task can be formalized by the proof tree in Figure 4 (where to save space, we have omitted names of the introduction and elimination rules for the @ operator). Note how this derivation is built around our analysis of the first-order case: the dots in the upper-right corner of Figure 4 mark where this earlier proof slots in.

This is indeed a formalization of the second-order Sally-Anne task. The conclusion, $@_aB@_sBl(basket, t_2)$, says that Anne believes that Sally believes that the marble is in the basket at the time t_2, and this indeed the correct response the the second-order task.

Moreover, Peter (who is now working away at the second-order task) can establish this. The first two premises used in the application of *Term* with which the proof concludes, $@_aS@_sSl(basket, t_0)$ and $@_aS@_sS\neg m(t_0)$, say that at time t_0, Anne saw that Sally saw that the marble was in the basket and that no action was taken to move it — this is the case since both Anne and Sally were present.

The third premise used in the concluding application of the *Term* rule, $@_aD@_s\neg Sm(t_1)$, says that Anne deduced that Sally did not see the marble being moved at the time t_1 — this is the case since Anne was present but Sally was absent at that time (and Anne did not see sneaky Sally peeking).

Finally, note the crucial role the belief manipulation rule BM plays in glueing the two levels of perspectival reasoning together. The embedded proof (which reasons from Sally's

perspective) yields the conclusion $@_sBl(basket, t_2)$, the correct response to the *first-order* task. But this is not enough: Peter must be able to deduce that holds from Anne's perspective. The application of BM prefixes the belief operator to form $B@_sBl(basket, t_2)$, and the very next step of the proof shows that this belief holds from Anne's point of view too.

Concluding discussion

We do not deny that second-order reasoning is more complex than first-order — the previous section with its embedded proof and use of the BM rule showed this clearly. Nonetheless, our analysis also suggests that the transition to second-order competence marks a more significant development than is suggested by the complexity only position: the full *reification* of beliefs. Attainment of first-order false-belief competence marks the stage at which the child becomes aware of the fact that beliefs held by other agents can be false; second-order competence marks the stage where beliefs become objects in their own right that can be recursively manipulated. This shift is mirrored in our analysis: we jumped from a logic of *Belief formation + Perspectival Reasoning* to one which allows unrestricted *Belief Manipulation* as well.

This is a significant advance. Beliefs are special objects. The child must learn that they can be embedded one inside another, and acquire the competence to carry out novel logical manipulations — and something like the BM rule, essentially a tool for handling beliefs recursively, seems to be required for this. It is tempting to speculate that at this stage of development some sort of "recursion module" is adapted to handle these strange new objects — but be that as it may, in typically developing children the reasoning architecture seems to be enriched at around the age of six in ways that suggest that a genuine conceptual change has taken place.

Recursively stacked beliefs lie at the heart of this transition, and this brings us back to our empirical work (Polyanskaya, Braüner, & Blackburn, 2016). Our logical investigations were carried out as part of an ongoing training study involving Danish speaking children with ASD. Our study is driven by the hypothesis that the delay (or even failure) experienced by children with ASD to attain second-order false-belief competence is linked to difficulties in belief manipulation. We are investigating whether children with ASD use linguistic recursion as a "scaffolding" to develop belief manipulation; this might explain why some children with ASD can pass second-order false-belief tasks, an explanation which was advanced in the first-order case by (Hale & Tager-Flusberg, 2003).

At the time of writing, our study was still work in progress. Nonetheless, we hope that these remarks show that there is a link between the abstract work of this paper and the concrete reality of psychological experimentation.

Acknowledgements

---

[12]In essence we are adding natural deduction machinery for the modal logic called K, the weakest normal modal logic.

Figure 4: Formalization of the child's correct response in the second-order Sally-Anne task

$$\frac{[a]\,[@_aS@_sSl(basket,t_0)]}{\dfrac{S@_sSl(basket,t_0)}{B@_sSl(basket,t_0)}}\,(P1) \quad \frac{[a]\,[@_aS@_sS\neg m(t_0)]}{\dfrac{S@_sS\neg m(t_0)}{B@_sS\neg m(t_0)}}\,(P1) \quad \frac{[a]\,[@_aD@_s\neg Sm(t_1)]}{\dfrac{D@_s\neg Sm(t_1)}{B@_s\neg Sm(t_1)}}\,(P0) \quad \frac{[@_sSl(basket,t_0)]\,[@_sS\neg m(t_0)]\,[@_s\neg Sm(t_1)]}{\vdots \quad @_sBl(basket,t_2)}$$

$$\frac{@_aS@_sSl(basket,t_0) \quad @_aS@_sS\neg m(t_0) \quad @_aD@_s\neg Sm(t_1) \qquad\qquad \dfrac{B@_sBl(basket,t_2)}{@_aB@_sBl(basket,t_2)}\,(BM)\;[a]}{@_aB@_sBl(basket,t_2)}\,(Term)$$

The vertical dots in the upper-right corner represent the derivation in Figure 2. So this proof contains two applications of *Term*: the concluding application, which is shown, and the one inside the earlier proof, which is not.

References

Arkoudas, K., & Bringsjord, S. (2008). Toward formalizing common-sense psychology: An analysis of the false-belief task. In *PRICAI 2008: Trends in artificial intelligence* (Vol. 5351, pp. 17–29). Springer-Verlag.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. MIT Press.

Blackburn, P. (2000). Representation, reasoning, and relational structures: a hybrid logic manifesto. *Logic Journal of the IGPL*, *8*, 339–365.

Bolander, T. (2014). Seeing is believing: Formalising false-belief tasks in dynamic epistemic logic. In A. Herzig & E. Lorini (Eds.), *Proceedings of the European conference on social intelligence (ECSI-2014)* (pp. 87–107). IRIT-CNRS, Toulouse University, France.

Braüner, T. (2011). *Hybrid logic and its proof-theory* (Vol. 37). Springer.

Braüner, T. (2014). Hybrid-logical reasoning in the Smarties and Sally-Anne tasks. *Journal of Logic, Language and Information*, *23*, 415–439.

Braüner, T. (2015). Hybrid-logical reasoning in the Smarties and Sally-Anne tasks: What goes wrong when incorrect responses are given? In *Proceedings of the 37th annual meeting of the cognitive science society* (pp. 273–278). Pasadena, California: Cognitive Science Society.

Braüner, T., Blackburn, P., & Polyanskaya, I. (2016). Second-order false-belief tasks: Analysis and formalization. In *Proceedings of Workshop on Logic, Language, Information and Computation (WoLLIC 2016)*. Springer-Verlag.

Fitting, M. (2007). Modal proof theory. In P. Blackburn, J. van Benthem, & F. Wolter (Eds.), *Handbook of modal logic* (pp. 85–138). Elsevier.

Gordon, R. (2009). Folk psychology as mental simulation. In E. Zalta (Ed.), *The stanford encyclopedia of philosophy*. Stanford University. (On-line encyclopedia article available at http://plato.stanford.edu/entries/folkpsych-simulation.)

Hale, C., & Tager-Flusberg, H. (2003). The influence of language on theory of mind: a training study. *Developmental Science*, *6*, 346–359.

Kurup, U., Bignoli, P., Scally, J., & Cassimatis, N. (2011). An architectural framework for complex cognition. *Cognitive Systems Research*, *12*, 281–292.

Miller, S. (2009). Children's understanding of second-order mental states. *Psychological Bulletin*, *135*, 749–773.

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

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*, 622–646.

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that..." attribution of second-order beliefs by 5-to 10-year-old children. *Journal of Experimental Child Psychology*, *39*, 437–471.

Polyanskaya, I., Braüner, T., & Blackburn, P. (2016). *Linguistic recursion and Autism Spectrum Disorder*. Manuscript.

Rips, L. (2008). Logical approaches to human deductive reasoning. In J. Adler & L. Rips (Eds.), *Reasoning: Studies of human inference and its foundations* (pp. 187–205). Cambridge University Press.

Seligman, J. (1997). The logic of correct description. In M. de Rijke (Ed.), *Advances in intensional logic* (Vol. 7, pp. 107 – 135). Kluwer.

Stenning, K., & van Lambalgen, M. (2008). *Human reasoning and cognitive science*. MIT Press.

Sullivan, K., Zaitchik, D., & Tager-Flusberg, H. (1994). Preschoolers can attribute second-order beliefs. *Developmental Psychology*, *30*, 395–402.

Szymanik, J., Meijering, B., & Verbrugge, R. (2013). Using intrinsic complexity of turn-taking games to predict participants' reaction times. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th annual conference of the cognitive science society* (pp. 1426–1432). Cognitive Science Society.