Logic in psychology: With applications to false-belief tests[1][2]

Torben Braüner
Roskilde University, Denmark

March 1, 2019

[1] The work on second-order false belief tests is joint work with Patrick Blackburn and Irina Polyanskaya.

[2] Thanks to the VELUX FOUNDATION for project funding (VELUX 33305)

Plan of talk

I **The first-order and second-order Sally-Anne tests**

——————— Case 1 ———————

II **Natural deduction for hybrid modal logic**

III **The first-order Sally-Anne test, formalized**

IV **The second-order Sally-Anne test, formalized**

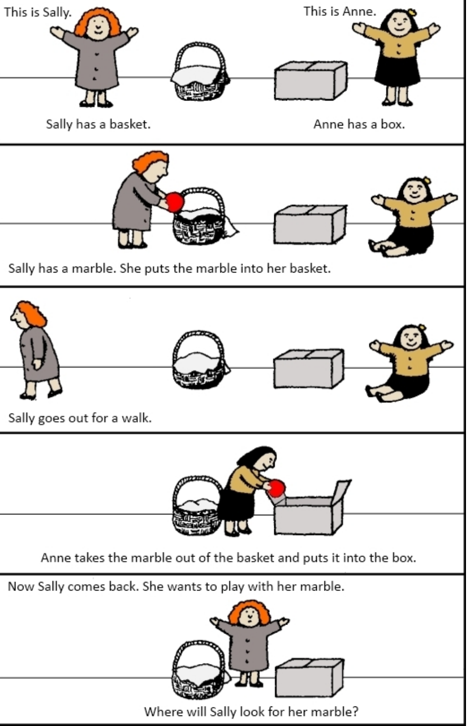
——————— Case 2 ———————

V **Comparing Sally-Anne to three other second-order false-belief tests**

VI **The four false belief tests: Empirical results**

Part I

The first-order and second-order Sally-Anne tests



The (first-order) Sally-Anne test measures a child's capacity to ascribe false beliefs to others

Goes back to Wimmer and Perner (1983)

Most children above the age of four give the correct answer

Baron-Cohen, Leslie, and Frith (1985) showed that autistic children have a delayed ability to answer correctly

Note: Autism Spectrum Disorder (ASD) is a psychiatric disorder with the following diagnostic criteria.

A. Persistent deficits in social communication and social interaction.

B. Restricted, repetitive patterns of behavior, interests, or activities.

Cf. *Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-V)*, published by the American Psychiatric Association.

(One in 59 U.S. children has ASD)

First-order versus second-order false-belief tests

First-order (age 4): The experimental subject has to realize that someone can hold a false belief about the world
"*Where does Sally believe the marble is?*"

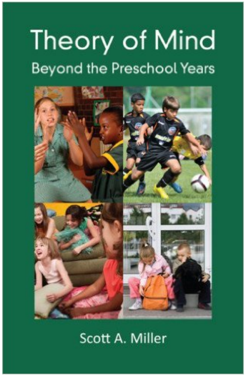
Second-order (age 5-7): The subject has to realize that someone can hold a false belief about someone's belief about the world

Second-order version of Sally-Anne test:
Sally looks through the keyhole when she is out

Second-order test question:
"*Where does Anne believe that Sally believe the marble is?*"

We have entered the world of recursion!

Second-order Theory of Mind



- Underlies much complex social behavior such as peer coordination and understanding non-literal language like idioms and irony
- But there are far fewer second-order false belief tests and they are less varied in design than their first-order cousins
- Much less is known and many conclusions are tentative, see Miller (2012)

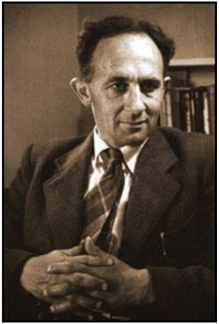
Based on hybrid logic we argue that the second-order test requires genuine modal reasoning but the first-order does not (psychological import, cf. CogSci 2016 paper)

—————————————— **Case 1** ——————————————

Part II

Natural deduction for hybrid modal logic

Hybrid logic was invented by Arthur Prior (1914-1969)



- Prior emphasized the internal perspective of modal logic
- "Perspective" is a keyword in this talk
- First key idea in hybrid logic:
  add nominals to the modal language, propositional symbols true at precisely one world/time/person/state/location:
  for example patrick and irina
- Second key idea in hybrid logic:
  build satisfaction statements,
  formulas like $@_{patrick}\,philosopher$ and $@_{irina}\,psychologist$
- Examples like this are typical of Prior's egocentric logic. They let us shift to another person's perspective

We want to formalize the reasoning in the Sally-Anne tests

Main assumption of our work:

Giving a correct answer to the Sally-Anne test involves a shift to the perspective of a different person and back

To formalize this reasoning with "local" perspectives, we use hybrid logic as follows:

- The perspectives of persons are represented by points in the Kripke model
- Nominals stand for such person perspectives
- Satisfaction operators can shift to a different perspective
- A natural deduction "perspective shifting" rule enables reasoning about what is the case from a different perspective

The perspective-shifting rule...

What is hypothetical reasoning? Reasoning when you put yourself in another person's shoes

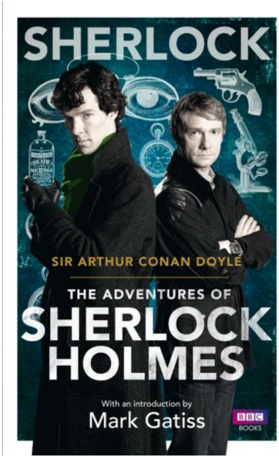
Example: Chess players visualize the board from the opponent's side, taking the opposing pieces for their own and vice versa

In other words, such a chess player

1. switches to the opponent's perspective
2. makes a decision of what to do in the opponent's situation
3. switches back again
4. predicts that the opponent will make the decision in question

Of course, the player has to make adjustments for relevant differences when taking the opponent's perspective

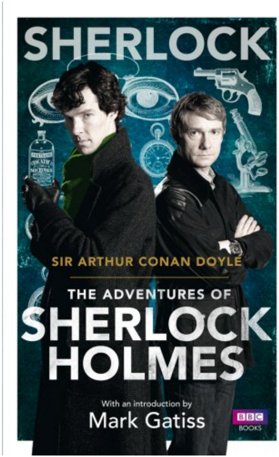
Also Sherlock Holmes does it...



You know my methods in such cases, Watson. I put myself in the man's place, and, having first gauged his intelligence, I try to imagine how I should myself have proceeded under the same circumstances.

Quotation from A.C. Doyle (1894)

Also Sherlock Holmes does it...



You know my methods in such cases, Watson. I put myself in the man's place, and, having first gauged his intelligence, I try to imagine how I should myself have proceeded under the same circumstances.
In this case the matter was simplified by Brunton's [the suspect's] intelligence being quite first-rate...

Quotation from A.C. Doyle (1894)

Natural deduction for propositional logic plus further rules

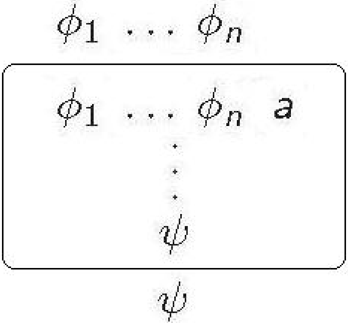
Focus on the *Term* rule: Enables hypothetical reasoning about what is the case from a particular perspective denoted by the point-of-view nominal a

$$\dfrac{\phi_1 \quad \ldots \quad \phi_n \qquad \begin{array}{c}[\phi_1] \ldots [\phi_n][a] \\ \vdots \\ \psi\end{array}}{\psi}\,(Term)^*$$

$*$ $\phi_1, \ldots, \phi_n$, 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 a.

Discharged assumptions are indicated by putting brackets $[\ldots]$ around formulas

The *Term* rule delimits a subderivation which is clear with alternative syntax like boxes in linear logic



Cf. also Jørgensen, Blackburn, Bolander, and Braüner's work on Seligman-style tableu systems (LPAR 2013, AiML 2016, Journal of Logic and Computation 2016)

Part III

The first-order Sally-Anne test, formalized

We want to formalize the reasoning in the Sally-Anne test

We use the symbolizations

B	Believes that ...
S	Sees that ...
$p(t)$	The marble is in the basket at the time t
$m(t)$	The marble is moved at the time t

and the following four "Belief formation" principles

(D)	$B\phi \to \neg B \neg \phi$	
$(P1)$	$S\phi \to B\phi$	Seeing leads to believing
$(P2)$	$Bp(t) \wedge \neg Bm(t) \to Bp(t+1)$	Inertia
$(P3)$	$\neg Sm(t) \to \neg Bm(t)$	

Loosely based on principles of Stenning and Van Lambalgen (2008) and also Arkoudas and Bringsjord (2008)

Let s be the nominal for Sally. Then the correct answer can be formalized as:

$$\dfrac{@_s Sb(t_0) \quad @_s S\neg m(t_0) \quad @_s \neg Sm(t_1) \qquad \dfrac{[s] \qquad \dfrac{\dfrac{\dfrac{\dfrac{[s]\ [@_s Sb(t_0)]}{Sb(t_0)}}{Bb(t_0)}\,(P1) \qquad \dfrac{\dfrac{\dfrac{[s]\ [@_s S\neg m(t_0)]}{S\neg m(t_0)}}{B\neg m(t_0)}\,(P1)}{\neg Bm(t_0)}\,(\mathrm{D})}{Bb(t_1)}\,(P2) \qquad \dfrac{\dfrac{[s]\ [@_s \neg Sm(t_1)]}{\neg Sm(t_1)}}{\neg Bm(t_1)}\,(P3)}{Bb(t_2)}\,(P2)}{@_s Bb(t_2)}}{@_s Bb(t_2)}\,(Term)$$

Note how the *Term* instance, marked in red, delimits the hypothetical reasoning taking place from the perspective of Sally

What's going on?

Perspectival Reasoning + Belief Formation (modalized literals)

Let s be the nominal for Sally. Then the correct answer can be formalized as:

$$\dfrac{@_s Sb(t_0)\ @_s S\neg m(t_0)\ @_s \neg Sm(t_1) \qquad \dfrac{[s] \qquad \dfrac{\dfrac{\dfrac{\dfrac{[s]\ [@_s Sb(t_0)]}{Sb(t_0)}}{Bb(t_0)}(P1) \qquad \dfrac{\dfrac{\dfrac{[s]\ [@_s S\neg m(t_0)]}{S\neg m(t_0)}}{B\neg m(t_0)}(P1)}{\neg Bm(t_0)}(D)}{Bb(t_1)}(P2) \qquad \dfrac{\dfrac{[s]\ [@_s \neg Sm(t_1)]}{\neg Sm(t_1)}}{\neg Bm(t_1)}(P3)}{Bb(t_2)}(P2)}{@_s Bb(t_2)}}{@_s Bb(t_2)}(Term)$$

Note how the *Term* instance, marked in red, delimits the hypothetical reasoning taking place from the perspective of Sally

What's going on?

Perspectival Reasoning + Belief Formation (modalized literals)

Part IV

The second-order Sally-Anne test, formalized

Second-order formalization based on observation

In the first-order Sally-Anne task,
the subject is asked to figure out Sally's reasoning

In the second-order case, the subject is asked to figure out
what Anne reasons about Sally's reasoning

Our key observation: In the second-order Sally-Anne task, Anne has the role that the subject has in the first-order case

So we can recycle the first-order formalization..

From “belief formation” to “belief manipulation”

A new rule is needed for reasoning under the scope of a belief modality and for transferring information to and from the scope

In particular: What Anne believes about Sally’s belief

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

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

Version of a rule for the modal logic K from Fitting (2007). We call it the *Belief Manipulation* (BM) rule

Second-order formalization: The gory details

$$
\dfrac{
@_a S @_s Sl(basket, t_0) \quad @_a S @_s S\neg m(t_0) \quad @_a D @_s \neg Sm(t_1) \qquad
\dfrac{
\dfrac{
\dfrac{\dfrac{[a][@_a S @_s Sl(basket, t_0)]}{S @_s Sl(basket, t_0)}}{B @_s Sl(basket, t_0)}\,(P1)
\quad
\dfrac{\dfrac{[a][@_a S @_s S\neg m(t_0)]}{S @_s S\neg m(t_0)}}{B @_s S\neg m(t_0)}\,(P1)
\quad
\dfrac{\dfrac{[a][@_a D @_s \neg Sm(t_1)]}{D @_s \neg Sm(t_1)}}{B @_s \neg Sm(t_1)}\,(P0)
\quad
\begin{matrix}[@_s Sl(basket, t_0)][@_s S\neg m(t_0)][@_s \neg Sm(t_1)] \\ \vdots \\ @_s Bl(basket, t_2)\end{matrix}
}{B @_s Bl(basket, t_2)}\,(\mathrm{BM})
\qquad [a]
}{@_a B @_s Bl(basket, t_2)}
}{@_a B @_s Bl(basket, t_2)}\,(Term)
$$

The green proof is the earlier first-order Sally-Anne formalization

Two *Term* instances: the concluding instance, which is shown in red, and the one inside the earlier proof, which is not shown

One instance of the Belief Manipulation rule (modal logic K)

More complicated proof-architecture:

In addition to perspective shifting machinery and belief formation principles, it also involves the Belief Manipulation rule

Perspectival Reasoning + Belief Formation + Belief Manipulation[3]

Thus, our logical analysis shows two stages in false-belief reasoning:

- First-order: Perspectival reasoning with modalized literals (technically, an indexed propositional logic)
- Second-order: The full machinery of the modal logic K

Thus, the second-order Sally-Anne test requires genuine modal reasoning, but the first-order version does not

[3] The distinction between Belief Formation and Belief Manipulation is adapted from Stenning and Van Lambalgen (2008)

Case 2

Part V

Comparing Sally-Anne to three other second-order false-belief tests

The experimental design in second-order Sally-Anne

A crucial role is played by a "principle of inertia" which says that an agent's belief is preserved over time unless the agent gets information to the contrary

Time t_0	Time t_1	Time t_2
Sally leaves after having put the marble in the basket Anne believes that Sally thinks that the marble is in the basket $B_{anne} B_{sally} \mathbf{basket}(t_0)$	*Anne moves the marble from the basket to the box* Sally sees through the keyhole that the marble is moved $B_{sally} \mathbf{box}(t_1)$ So $B_{sally} \neg \mathbf{basket}(t_1)$ and hence $\neg B_{sally} \mathbf{basket}(t_1)$	*Sally has returned* Correct answer: "Anne believes that Sally thinks that the marble is in the basket" $B_{anne} B_{sally} \mathbf{basket}(t_2)$ Derivable by inertia from t_0 as Anne does not know Sally's belief changed at t_1

The full information picture

Zero-order, first-order and second-order information in the second-order Sally-Anne test

Blue formulas are part of the experimental design

	Time t_0	Time t_1	Time t_2
Zero-order	$basket(t_0)$	$\neg basket(t_1)$	$\neg basket(t_2)$
First-order	$\mathbf{B_{sally}\, basket(t_0)}$	$\mathbf{B_{sally} \neg basket(t_1)}$	$\mathbf{B_{sally} \neg basket(t_2)}$
	$B_{anne}\, basket(t_0)$	$B_{anne} \neg basket(t_1)$	$B_{anne} \neg basket(t_2)$
Second-order	$\mathbf{B_{anne} B_{sally}\, basket(t_0)}$	$\mathbf{B_{anne} B_{sally}\, basket(t_1)}$	$\mathbf{B_{anne} B_{sally}\, basket(t_2)}$
	$B_{sally} B_{anne}\, basket(t_0)$	$B_{sally} B_{anne} \neg basket(t_1)$	$B_{sally} B_{anne} \neg basket(t_2)$

Note the asymmetry in Sally and Anne's second-order information:

From time t_1 on, Sally believes that Anne believes that the marble has been moved away from the basket, since Sally can see Anne moving the marble, **but Anne is not aware of this, hence, she is deceived**

The four standard second-order false belief tests

The range of second-order false belief tests is essentially[4] covered by the following:

- The second-order Sally-Anne test (described earlier)
- The ice-cream truck story
- The bake-sale story
- The puppy story

They all share the pattern that the correct answer is a formula $\boldsymbol{B_x B_y \phi}$ whose truth is preserved from t_0 to t_2 where the subformula $\boldsymbol{B_y \phi}$ becomes false at the intermediate stage t_1

[4] An exception is a second-order version of the Smarties test described in the unpublished manuscript Homer and Astington (2001)

Classification of the four second-order tests

The four tasks involve identical first-order information, but there are differences at the zero-order and second-order levels:

	Zero-order information	Second-order information
Second-order Sally-Anne	Change in world	Asymmetry (deception)
Ice-cream	Change in world	Symmetry
Puppy	No change in world	Asymmetry (deception)
Bake-sale	No change in world	Symmetry

Thus, the four standard second-order tasks covers all possible combinations of zero-order and second-order information

Part VI

The four false belief tests: Empirical results

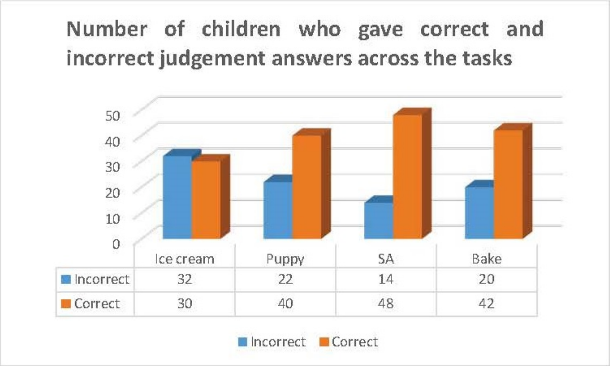
Participants and materials of study

62 Danish children with ASD, age between 7 and 15

- Battery of tests, key ones being the four second-order false belief tasks
- Plus Recursive Embedding Tool (RET)
- Plus training component

BUCLD 2017 paper and Irina's PhD thesis (March 2019)

Judgement scoring: Is the answer correct?



Note: Ice-cream task hard whereas Sally-Anne easy

All tasks are positively correlated, but what more can be said?

Latent Class Analysis (LCA)

Aim: Identify patterns of responses to the tasks, in particular patterns involving *combinations* of tasks.

Given the four task scores as observed variables, LCA splits the subjects into latent classes such that

- subjects in the same class give similar answers to the tasks, and
- subjects across classes give different answers to the tasks.

Latent classes correspond to variables that are not measured directly

The LCA with two latent classes gave the best fit

Table: Probability of passing a task for each class

Task	**Class 1**	**Class 2**
Ice-cream	.9884	.0000
Puppy	.8000	.4966
Sally-Anne	.9666	.5896
Bake-sale	.8332	.5280

Table: Probability of passing a task for each class

Task	Class 1	Class 2
Ice-cream	.9884	.0000
Puppy	.8000	.4966
Sally-Anne	.9666	.5896
Bake-sale	.8332	.5280

A child belongs to Class 1 iff the child passes the IS task

If a child belongs to Class 1 then the child passes the SA task (but not vice versa)

Table: Probability of passing a task for each class

Task	Class 1	Class 2
Ice-cream	.9884	.0000
Puppy	.8000	.4966
Sally-Anne	.9666	.5896
Bake-sale	.8332	.5280

A child belongs to Class 1 iff the child passes the IS task

If a child belongs to Class 1 then the child passes the SA task (but not vice versa)

Thus, the set of subjects passing IS is included in the set of subjects passing SA

Much stronger than the observation that more subjects passes SA (48 out of 62) than IS (32)

Fits logical classification

The second-order information in IS is symmetric (no deception), but in SA it is asymmetric (deception)

But IS and SA involves the same zero-order information ("change in world")

Thus, our logical analysis shows that moving from IS to SA corresponds to adding deception and our empirical study shows that moving from IS to SA makes the task easier in a very strong sense

In line with deception being known to have a facilitative effect when included in first-order false belief tasks

More information

T. Braüner, P. Blackburn and I. Polyanskaya. Being Deceived: Information Asymmetry in Second-Order False Belief Tasks, *Topics in Cognitive Science*, to appear

I. Polyanskaya, T. Braüner and P. Blackburn. Second-order false beliefs and recursive complements in children with Autism Spectrum Disorder, *BUCLD 42: Proceedings of the 42nd annual Boston University Conference on Language Development*, 2018

T. Braüner, P. Blackburn and I. Polyanskaya. Recursive belief manipulation and second-order false-beliefs, *Proceedings of the 38th Annual Meeting of the Cognitive Science Society*, 2016

T. Braüner. Hybrid-Logical Reasoning in the Smarties and Sally-Anne Tasks: What Goes Wrong When Incorrect Responses are Given?, *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*, 2015

T. Braüner. Hybrid-Logical Reasoning in the Smarties and Sally-Anne Tasks, *Journal of Logic, Language and Information*, volume 23, 2014