Perspective taking studies

A. Questions
----

1. Which kinds of perspective taking are automatic (i.e. their occurrence is to an interesting extent independent of the subject's motives and task)?

2. Is there a difference between kinds of perspective taking with respect to the ages at infants first show evidence of perspective taking?

B. Conjectures

Conjecture 1S (Agi): (a) Infant mindreading is continuous with full-blown adult mindreading in the sense that it involves the same concepts and representations.
(b) Automatic mindreading in adult humans is full-blown mindreading.

Conjecture 2S (Steve): (a) Infant mindreading has signature limits distinct from those of full-blown adult mindreading. Quick and dirty version: these limits correspond roughly to a distinction between Level-1 and Level-2 perspective taking. More careful version: there is a minimal form of mindreading which involves representing relational mental states only, and does not involve representing representational mental states as such. Infant perspective taking is limited because it involves only this minimal form of mindreading.
(b) Automatic mindreading in adult humans is distinct from full-blown mindreading and subject to the same arbitrary limits that infant mindreading is.

We don't know for sure that infant mindreading is
automatic; so we can't assume in advance that 1S(a) is
incompatible with 2S(b) or conversely.

C. Background

Suppose that two people, A and B, are looking at a
pair of biscuits. The left biscuit is actually larger
than the right biscuit, and this is how it looks to A.
But B is subject to the Ebbinghaus illusion, so to her
the left biscuit looks smaller than the right biscuit.
To characterise the difference between A and B it is
most natural to appeal to a difference in the content
of their experiences. And it is hard to characterise
the difference in terms entirely of relational mental
states only. This is because:
    (i) there is a single pair of objects which A and
B experience*;
    (ii) with respect to the dimension of interest,
the relative size of two objects, no pair of objects
could simultaneously have two different values;
    (iii) this property does not actually change
throughout;
    (iv) person A experiences the pair of objects as
having one value on the dimension of interest; and
    (v) person B experiences the pair of objects as
having a different value on the dimension of interest.

Suppose that a mindreader knows (i)-(iv). Then in
order to track the way A's and B's experiences differ,
she cannot take an individual's encounters with how
things are to determine what that individual
experiences. (Strictly speaking, what matters is her
beliefs about how things are rather than how they
actually are.) Instead she needs to use a principle
equivalent to: how people experience the relative
sizes of things depends not on their actual relative
sizes but on how their relative sizes appear in

different contexts. In theory a principle along these
lines could be formulated with appeal to relational
mental states only. This could be done by defining
appearance profiles for objects which would capture
their apparent sizes in different contexts. So in
theory a defender of 2S could argue that even
abilities to track perceptual illusions would not be
evidence for more than relational mindreading. But
this would make `minimal mindreading' less minimal,
and it would impose a burden on proponents of 2S: they
would have to show that mindreaders really know about
the relevant appearance profiles and can apply that
knowledge in mindreading. Assuming they don't, we can
ignore this possibility and suppose that proponents of
2S agree that tracking differences in experience due
to the Ebbinghaus illusion is beyond the limits of
minimal mindreading abilities.

Conjecture 2S(a) predicts that infants might be able
to track both differences in experience which are due
to the simple occlusion of objects (`Level 1') but are
unable to track differences in experience due to
illusions (`Level 2'). So evidence that infants track
differences in experience due to illusions would
count against 2S(a) and in favour of 1S(a).

Conjecture 1S(a) predicts that the difference between
tracking experiences due to the simple occlusion of
objects (`Level 1') and experiences due to illusions
(`Level 2') should not be reflected in infants'
abilities, all things being equal. So, if we had
carefully matched tasks, evidence that turning a
simple occlusion task into an illusion task would
count against 1S(a) and in favour of 2S(a).

The two views make parallel predictions for
automaticity. Evidence for the automaticity of
abilities to track experiences due to illusions would
count against 2S(b) and in favour of 1S(b). Evidence
that automatic mindreading abilities can enable

tracking differences in experience due to occlusion
but not to illusion would count against 1S(b) and in
favour of 2S(b).

D. `Take the biscuit': Ebbinghaus (/Titchener) Task,
`Level-2 version'

There are two circular biscuits, one just visibly
larger than the other. There is also a special screen
with two circular holes. The biscuits can be placed
in the holes. One side of the screen is blank, so
that inserting the biscuits into the holes and looking
at them from this side has no affect on their apparent
sizes. The other side of the screen---call it the
Illusory Side---has circles painted surrounding the
holes in such a way that an Ebbinghaus illusion is
generated. So when the biscuits are appropriately
placed into the holes in the screen, seen from the
Illusory Side of the screen the larger biscuit appears
to be smaller than the small biscuit.

Background requirements: (1) If subjects are prompted
to take a biscuit, they will spontaneously take what
visually appears to be the larger of the two biscuits.
(2) If subjects attempt to take the larger biscuit, if
they are naive about the screen, having only ever seen
it from one side, and if they have only ever seen
these biscuits in the screen, then they will
mistakenly take the smaller biscuit.

Task: Subjects are familiarised with the screen and
its effect on the apparent sizes of biscuits. They
see it from different sides. The biscuits are then
placed into the screen. Subjects sit on one side of
the screen. A Protagonist now spontaneously appears.
In the Congruent Condition, a Protagonist sits beside
to the subject, so that both subject and Protagonist
experience the biscuits in the same way. In the

Incongruent Condition, a Protagonist sits opposite the subject, so that Protagonist sees the other of the screen. In this condition the protagonist and subject have conflicting visual experiences* concerning the relative sizes of the biscuits (conflicting in the sense that they cannot both be veridical). Subjects are then prompted to take a biscuit. We measure which biscuit they choose, movement onset time, total time taken to retrieve a biscuit, the trajectory of their reach, which biscuit they look at first, and maybe other things.

Are subjects' influenced by the protagonists' perspective? A difference in any measure between the Congruent and Incongruent conditions would be evidence that they are. Ideally we would find something like this for proportion of correct choices:

Subject sees:>	Plain Side	Illusory Side
Protagonist sees:		
Plain Side	.9	.6
Illusory Side	.6	.3

I.e. when the subject faces an illusion, the Protagonist's veridical experience facilitates; when the subject's experience is veridical, the Protagonist's illusory experience interferes. But strictly we only need either half of this.

Since subjects' task is only ever to take a biscuit and there is no incentive for them to consider the Protagonist's experience, it seems reasonably safe to say that modulation by the Protagonist's experience is evidence for automatic perspective taking. (Is this right?)

Any such evidence supports 1S(b) over 2S(b). And if we got this pattern with infants, we'd have to accept 1S(a) and reject 2S(a).

To find evidence for 2S we need a `Level-1' version of Take the Biscuit that is as similar as possible to the above. We might achieve this by having two pairs of biscuits. So, ideally, all that changes is this: we put different pairs of biscuits either side of the screen, instead of having a single pair of biscuits visible from both sides of the screen. Otherwise we have the same Congruent and Incongruent conditions. If the difference in condition (Congruent vs Incongruent) influences subject's performance in this Level-1 version but not in the Level-2 version of Take the Biscuit, this would be evidence for 2S over 1S.

E. Potential problems with Take the Biscuit

(i) in this version of Take the Biscuit, the Incongruent Condition always involves the Protagonist sitting opposite, so that left and right are reversed. This might make the task unnecessarily difficult. An alternative might be to have the subject and Protagonist wear colour filtering glasses, red vs. green. The painted circles on the barrier could be coloured so that they cannot be seen when wearing the red glasses. (And in the matching `Level-1' version, there could be a `biscuit' that cannot be seen when wearing the red glasses.) In this version the Protagonist and the subject can always sit side-by-side: the Incongruent Condition works by having them wear different glasses. In this version we might find that any effect of condition is be modulated by whether the Protagonist near to the subject or sitting far away. (This might be tell us something about how automatic mindreading mechanisms are triggered.)

(ii) Subjects have to recognise that they are seeing the same pair of biscuits from different sides of the

screen. If they somehow ended up confused into
thinking that there are two pairs of biscuits, the
task would collapse into `Level-1'. We should be able
to avoid this line of objection with training.

(iii) Someone might object that having the Protagonist
sit on the opposite side of the screen (or wear the
Incongruent glasses) merely makes salient the
illusion. So (according to the objection) it's not
any fact about the Protagonist's experience that
influences the subject's responses, only the varying
salience of the illusion at the point of decision.
(*How to get around this?)

(iv) This might not work with infants and younger
children because they might not be deceived by the
Ebbinghaus illusion, as Doherty et al 2010 suggest.
(This is part of the motivation for having subjects
take a biscuit instead of the perhaps more explicit
pointing response those authors required; might be
interesting if we could show that infants are
susceptible to the illusion when measured in this
way.)

F. Feed the Duck-Rabbit

There is a wooden shape. In isolation it is a
duck-rabbit ambiguous figure. It can be placed into a
hole in a screen. On one side of the screen, the
surround is painted in a way that biases seeing the
figure as a rabbit (the Rabbit Side); on the other
side, the surround is painted so as to bias seeing the
figure as a duck (the Duck Side). (Alternative: we
only use a single side of the screen; the features are
painted onto the screen in such a way that wearing
green glasses provides visual information which biases
seeing the figure as a rabbit, and whereas wearing red
glasses (or no glasses) biases seeing a duck.)

Background requirements: (1) If subjects are prompted to choose whether feed the thing carrots or seeds, all things being equal they will normally choose the carrots when looking at the Rabbit Side and the seeds when looking at the Duck Side.

Task: Subjects are familiarised with the screen and its effect on the thing. They see it from different sides. The duck-rabbit is then placed into the screen. Subjects sit on one side of the screen. A Protagonist appears for the first time. In the Congruent Condition, a Protagonist sits beside to the subject, so that both subject and Protagonist would normally see duck-rabbit in the same way (as the rabbit, say). In the Incongruent Condition, the Protagonist sits opposite the subject, so that Protagonist and subject would normally see the duck-rabbit differently (one as the rabbit, the other as the duck). Subjects are then prompted to feed the thing. We measure whether they choose carrots or seeds, movement onset time, total time taken to retrieve a food, the trajectory of their reach, which food they look at first, and maybe other things.

Are subjects' influenced by the protagonists' perspective? Just as in Take the Biscuit, a difference in any measure between the Congruent and Incongruent conditions would be evidence that they are. One ideal pattern of movement onset times (ideal for proponents of 1S, anyway) would be:

Subject sees:>	Duck Side	Rabbit Side
Protagonist sees:		
Duck Side	fast	slow
Rabbit Side	slow	fast

Again, since subject's task doesn't involve the Protagonist (they just have to feed the thing), it seems safe to say that modulation by how the

Protagonist would normally see the thing is evidence for automatic perspective taking. (?)

The theoretical link between this task and the 1S vs. 2S conjectures is not quite the same as Take the Biscuit. Take the Biscuit centers on a dimension---the relative size of two objects---on which no pair of objects could simultaneously have two different values. If a single thing can simultaneously depict or represent two different things (a duck and a rabbit), results from this task can't be interpreted in exactly the same way as Take the Biscuit. Unless the perhaps subjects think of the things being a duck as excluding its also being a rabbit. So maybe there is a way to argue that 2S predicts failure in tracking differences in experience in this task. (?)

If we need a `Level-1' version of Feed the Duck-Rabbit, we could just change the setup so that the duck and rabbit are distinct objects, and each lives on one side of the barrier.

Footnotes

[*] There are philosophers who claim that all experiences are relational and would deny that any experiences would be veridical. They attempt to get around the problem of illusion by appeal to something like an `appearance profile'. I'm hoping that we can ignore this complication.