

## Alejo Salles <alejosalles@gmail.com>

## Fwd: A n HBM modeling question

1 message

Alison Gopnik <gopnik@berkeley.edu>
To: Alejo Salles <alejosalles@gmail.com>

Cc: Caren M Walker <caren.walker@berkeley.edu>

Wed, Nov 12, 2014 at 7:53 PM

Dear Alejo, I talked to Tom who said you are working on something like this for Fei, and of course, it immediately occurred to me that you are the right person to ask about the issue in the thread below! Any thoughts? Maybe you and I and Caren should meet? I'm attaching the paper draft so that you can see more of the details. Alison

Alison Gopnik
Professor of Psychology and Affiliate Professor of Philosophy
Tolman Hall
University of California at Berkeley
Berkeley, CA, 94720
gopnik@berkeley.edu
510-642-2752
www.alisongopnik.com

## Begin forwarded message:

From: Alison Gopnik <gopnik@berkeley.edu> Subject: A n HBM modeling question Date: November 12, 2014 at 2:36:16 PM PST

To: Caren M Walker <caren.walker@berkeley.edu>, Tom Griffiths <tom\_griffiths@berkeley.edu>

Dear Tom,

Caren and I have been working on the relational match to sample data and are finding an interesting pattern. As you may recall the three year olds do worse than the 18 month olds on both the SAME task (two objects that are the same make it go, two that are different don't) and the DIFFERENT task (the reverse) Our hypothesis is that this is (partly) because they develop a stronger prior for the over-hypothesis that kinds of objects, rather than relations between them, make the machine go as they grow older. To test this we gave them evidence against this over-hypothesis - they see that individual objects fail to make the detector go and then that combinations do make it go. The result is that this improves performance significantly in the "same" condition but not in the "different" condition, though there is a slight increase. (see figure below). Getting them to explain leads to equal improvements in both conditions.

Thinking through this, it seems to me that the data pattern in the "same" condition has a lower likelihood given the "same kind" over-hypothesis than the pattern in the "different condition". Assuming that the blocks are randomly sampled and that some fixed proportion of block types activate the machine, then the probability that the machine activates on any trial should be higher when two different kinds of blocks are on the machine, that is where there are two potential activators, than when the same kind of blocks are on

the machine, where there is only one potential activator - this is true in the "DIFFERENT" condition but not the "SAME" condition. So the "SAME" data pattern potentially offers counter-evidence to the "same kind" over-hypotheses, in a way that the "DIFFERENT" pattern does not.

In contrast, on the "relational" over hypothesis - that the relations between the blocks lead to activation-- both patterns are equally likely.

I think that this could explain the data pattern, the younger kids have less of a prior towards the kind hypothesis, so there is less of an asymmetry between same and different. In the negative evidence condition, there is stronger intrinsic evidence against the over hypothesis in SAME than DIFFERENT, in addition to the negative evidence we provide and so the kids are more likely to switch to the relational over-hypothesis. In the explanation condition, on our account, there is no difference in evidence, but instead a privileging of more abstract/general hypothesis which might lead kids to prefer the relational hypothesis in both conditions.

Does this sound plausible to you given an HBM approach? Would it be worth trying to do some more explicit modeling? Would one of your students be interested in taking this on perhaps or perhaps we could meet and discuss? A

Alison Gopnik
Professor of Psychology and Affiliate Professor of Philosophy
Tolman Hall
University of California at Berkeley
Berkeley, CA, 94720
gopnik@berkeley.edu
510-642-2752
www.alisongopnik.com

## 2 attachments



RMTS-UDE\_PNAS.docx 261K



PastedGraphic-1.pdf 42K