Emond, B., Ferres, L. (2001) Modeling the False-Belief Task : An ACT-R Implementation of Wimmer & Perner (1983). Second Bisontine Conference for Conceptual and Linguistic Development in the Child Aged from 1 to 6 Years. Besançon : France, March 21-23

Modeling the False-Belief Task:

An ACT-R Implementation of Wimmer & Perner (1983)

Bruno Emond
Université du Quebec à Hull, Hull, Quebec, Canada
Email: bruno.emond@nrc.ca

Leo Ferres
Carleton University, Ottawa, Ontario, Canada
Email: lferres@ccs.carleton.ca

According to empirical evidence, it is not until around three years of age that children master some of the concepts possessed by adults in their mature theory of mind. This developmental staging was made obvious by the now famous false-belief task, devised by Heinz Wimmer and Joseph Perner in 1983, (later called the Sally-Anne test (henceforth, S-AT)). Details aside, the S-AT goes as follows: the subject is introduced to two puppets: Sally and Anne. While playing, Sally puts a marble into a basket and then goes outside (the puppet disappears under the table, for example). When Sally is not around, naughty Anne changes the location of the marble. She takes it out of the basket and puts it in a box. Some time later, Sally comes back and wants to play with her marble. Subjects are then asked the critical question: where will Sally look for her marble? Older subjects, whose ages are just under five years, will know that Sally does *not* know that the marble has been moved from A to B and will therefore correctly predict that she will look for it in the basket, and not in the box. In turn, the younger group will fail to answer this question correctly. When asked, they predict that Sally will look for her marble in the last location they saw it in; that is, in the box, not where Sally left it first (the basket).

There are actually two seemingly mutually exclusive explanations for this empirical finding. On the one hand, there is what I will call the *conceptual explanation*: the above mentioned results may suggest that children younger than three years of age do not have an understanding of folk psychology (Davies & Stone, 1995). Specifically, young children would seem to lack the *concept* of belief (or, alternatively, that of false-belief); i.e. the ability to understand people as capable of entertaining beliefs that are different from the child's own (see, most notably, Wellman (in press)). On the other hand, there is what I will call the *computational explanation*: the results above may suggest that children younger than three years do not have access to the computational mechanisms needed to predict the behavior of the characters involced in the story. Fodor (1992) claimed that very young children (even those who do not pass the S-AT) do have what he calls a 'very simple theory of mind' (for Fodor, 'theory of mind' is, in fact, an innate modularized database). He concludes then that three-year-old children *have* the concept of false-belief, and, by logical necessity, that of belief. In Fodor's view, the problem is that such young children do not

have access to the computational mechanisms required to predict the behavior of others, given their false beliefs.

In order to shed light on these issues, we will attempt to model all the stages involved in the S-AT by using the Act-R modeling environment (Anderson & Lebiere, 1998). ACT-R is particularly useful in the modeling of this situation for an important reason: there are, essentially, three kinds of memories associated with the flow of information in ACT-R: the first is a *goal stack*, which prioritizes the intentions that guide the behavior of the system (in the S-AT, the intention would be to answer the question posed by the experimenter); the second kind of memory in ACT-R is called *declarative memory*, which contains 'chunks' of information that the system 'knows'. An English translation of a possible chunk in the declarative memory of the system during performance in the S-AT could be the following: "Sally put [+past] the marble in box A". Lastly, the last kind of memory available in ACT-R is called *procedural memory*, containing the production rules that will eventually help the system come up with a particular behavior. In the S-AT, we might expect a production rule of the following sort: "IF I know that Sally put the marble in box A AND the question is 'where did Sally put the marble' THEN answer 'Box A'".

In trying to model the false-belief task, we encountered what seem to be important implications for the debate between the 'computational' and the 'conceptual' explanations of performance in the S-AT. The purpose of this paper is, then, to show how some of the models would work and how certain assumptions lead to favoring one or the other explanation.

REFERENCES

Anderson, J. & Lebiere, C. (1998). *The Atomic Components of Thought*. Mahwah, NJ: Lawrence Erlbaum.

Davies, M. & Stone, T (Eds.). (1995). *Folk psychology*. Oxford, UK: Blackwell.

Fodor, J. (1992). A theory of the child's theory of mind. *Cognition*, ***44***, 283-296.

Wellman, H. M., Cross, D. & Watson, J. (in press). Meta-analysis of theory of mind development: The truth about false belief. Child Development.

Wimmer, H. & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, ***13***, 41-68