THE HIERARCHICAL ORGANIZATION OF THE BRAIN AS A KEY TO THE STUDY OF CONSCIOUSNESS IN HUMAN AND NON-HUMAN ANIMALS: PHYLOGENETIC IMPLICATIONS

R Van den Bos

Animal Welfare Centre, Utrecht University, Yalelaan 17, NL-3584 CL Utrecht, The Netherlands

Recently I discussed a framework to study consciousness in animals (Van den Bos 2000). This framework comprised:

i) a hierarchical negative feedback model of the central nervous system in which behaviour is the control of input (perception) expressed as transitions of behavioural patterns (= motion) and as transitions of mental states (= consciousness; Van den Bos 1997). Consciousness is hypothesized to be a property of neuronal networks of self-organizing systems dedicated to dealing with rapidly-changing environments affording flexibility of behavioural patterning;

ii) mental states – mental representations – as intentional states as described in folk psychology (feelings, desires, beliefs, insights, etc);

iii) the separation of mental states into two components, an invariant part ('the presence as such'; *that*) and a variant part ('the way it looks'; *how*), which are separately related to the organization of the central nervous system, ie a 'neuronal network' and 'momentary active connections within the neuronal network determined by input and output of this neuronal network' respectively;

iv) the idea that every mental state (psychological trait) – whether feelings, desires or beliefs – can be viewed conceptually as the product of a specialized neuronal network and information encoded in the network (neuronal input-output relations) which is expressed as a specific series of spontaneous behavioural patterns or in specific tests; and

v) a phylogeny based on the invariant part, ie mental states, and the presence of a neuronal network.

For the study of evolution of mental states, this implies: i) a hypothesis and reconstruction of the phylogeny of a neuronal network; ii) understanding the kind of information encoded in the neuronal network; and iii) a species-specific behavioural analysis of the expression of the trait in a sufficient number of normally-reared individuals.

I have recently applied this approach to the question of theory of mind (ToM) in human and non-human primates (Van den Bos *et al* 1998). The concept of ToM (second-order intentionality or mindreading ability) stands for the ability to have insight into one's own mental states and those of others and know that they are (causally) related to observable series of behavioural patterns. Critical review of data on ToM (especially one of its initial stages, mirror self-recognition) revealed inter- and intra-specific differences which may be due to: i) hardware differences (neuronal structures); ii) methodology; or iii) rearing conditions (De Veer & Van den Bos 1999; Van den Bos 1999). We have first concentrated on inter-specific differences in relation to hardware differences. The question of which putative neuronal network might underlie ToM was answered by analysing reported brain deficits in autistic children and schizophrenic adults (Baron-Cohen 1995; Frith 1998). A candidate network appeared to consist of orbitofrontal cortex, medial frontal cortex, amygdala and (parts of the) temporal cortex. Formal phylogenetic reconstruction of this

neuronal network showed that differences in parts of these neuronal structures in different species might underlie differences in the presence of ToM in these species.

References

Baron-Cohen S 1995 *Mindblindness: an Essay on Autism and Theory of Mind*. MIT Press: Cambridge, USA

De Veer M W and Van den Bos R 1999 A critical review of methodology and interpretation of mirror self-recognition research in nonhuman primates. *Animal Behaviour 58:* 459-468

Frith C 1998 The role of the prefrontal cortex in selfconsciousness: the case of auditory hallucinations. In: Roberts A, Robbins T and Weiskrantz L (eds) *The Prefrontal Cortex: Executive and Cognitive Functions* pp 181-194. Oxford University Press: Oxford, UK

Van den Bos R 1997 Reflections on the organisation of mind, brain and behaviour. In: Dol M, Kasanmoentalib S, Lijmbach S, Rivas E and Van den Bos R (eds) *Animal Consciouness and Animal Ethics: Perspectives from the Netherlands* pp 144-166. Van Gorcum: Assen, The Netherlands

Van den Bos R 1999 Reflections on self-recognition in nonhuman primates. *Animal Behaviour 58:* F1-9

Van den Bos R 2000 General organizational principles of the brain as key to the study of animal consciousness. *Psyche 6* http://psyche.cs.monash.edu.au/v6/psyche-6-05-vandenbos.html

Van den Bos R, Nolten C and Van Veller M 1998 Phylogenetic reconstruction of a neural network underlying Theory of Mind in primates. *European Journal of Neuroscience 10:* 197