Thinking About Thinking: How Do We Know What We Know?

Sunday 20 February 2011, 3.00pm – 4.30pm
AAAS 2011, Washington Convention Center

Synopsis

Humans have feelings of doubt and confidence, of certainty and uncertainty. You know if you do not know or remember – a perfect example of this is when something is on the tip of your tongue. This ability to evaluate and predict one's own mental performance is known as metacognition. It is one of our most sophisticated cognitive capacities and has even been thought to be uniquely human. Metacognition rivals language and tool use in its potential to reveal similarities and differences between human and animal minds.

This session presents a new, exciting collaboration across disciplines: philosophy, psychology and behavioural biology. It will explore how newly devised experimental paradigms, testing metacognition in dolphins and monkeys, show that it is not a uniquely human talent. Moreover, the same simple, non-verbal and perceptual tasks used to gauge animals can also be used to explore young children's earliest metacognitive achievements, something which has often been underestimated in existing verbal and introspective assessments.

The cutting-edge transfer of experimental paradigm from comparative to developmental psychology offers surprising insight into the nature of metacognition in humans and what this means for children's development and learning. The session will also expand on how this innovative research is profoundly affecting philosophers involved in the current debate on the theory of mind.

Moderator: Eva Hoogland, Senior Science Officer, European Science Foundation

Speakers

- **Recent Developments in the Study of Animal Metacognition**
  John David Smith, Professor, State University of New York at Buffalo, USA

- **Metacognition of Ignorance - What Can Animal Studies Teach Us?**
  Josef Perner, Professor of Psychology, University of Salzburg, Austria

- **Thinking About Thinking - Evolutionary, Developmental and Epistemological Aspects**
  Joëlle Proust, Director of Research, Institut Jean-Nicod, Paris, France

Presentation summaries

Recent Developments in the Study of Animal Metacognition
John David Smith, Professor, State University of New York at Buffalo, USA

Humans feel doubt and uncertainty. We know when we don't know or don't remember—a good example of this is the feeling that someone's name is "on the tip of your tongue." This sophisticated cognitive capacity to be aware of our own thinking is called metacognition and it is closely allied to humans' consciousness and self-awareness.

Pioneering research supported by the National Science Foundation and the National Institutes of Health is producing growing evidence that animals demonstrate important parallels with human conscious metacognition. They may share humans' ability to reflect upon, monitor, or regulate their states of mind. This research is reviewed in a recent issue of the journal Trends in Cognitive Sciences (Volume 13, Issue 9). The research has inaugurated a new and rapidly developing area of comparative-cognition research.

To find out whether non-human animals do have knowledge of their own cognitive states, the researchers work with macaques (Old World monkeys) and capuchins (New World monkeys) at the Language Research Center of Georgia State University, an internationally recognized center for comparative-cognition research. The researchers give the animal participants—who are all joystick-trained and who participate eagerly in computer-based cognitive tasks—difficult perceptual, cognitive, and memory problems. The difficulty creates uncertainty in the animals' minds. The researchers also give animals a response with which they can decline to complete any trials of their choosing. This uncertainty response lets animals report on, or cope with, the uncertainty and difficulty.

Macaques use this response identically to humans who say that their uncertainty responses are based on conscious uncertainty. Though macaques say nothing, several lines of converging research show that a higher-level cognitive interpretation of macaques' performance is required, too, perhaps at or near the level of consciousness.

In sharp and surprising contrast, capuchin monkeys barely express any capacity for metacognition, despite the fact that they are so highly responsive, adaptive, and intelligent in many domains that they are often called the poor-person's chimpanzee. This species difference testifies to the cognitive sophistication of the metacognitive data pattern when it is observed, and it also raises intriguing questions about the emergence of reflective mind within the primate lineage.

This research opens a new window on reflective mind in animals. It illuminates the phylogenetic roots of human metacognition, which might have implications for understanding how or why conscious cognitive regulation came to be such a crucial aspect of humans' cognitive system.

The research is also helping to reveal the developmental roots of human metacognition. The simple, nonverbal tasks that suit animals extend the range of paradigms available for developmental research. They are ideal for testing young humans and language delayed or autistic children.

The research furthers the development of animal models for metacognition, grounding the study of neurological substrates and neurochemical blocks and enhancements. Finally, the demonstrations of animal awareness emerging from the research have important implications regarding respectful, compassionate husbandry in all areas of animal research.

Metacognition of Ignorance – What Can Animal Studies Teach Us?
Josef Perner, Professor of Psychology, University of Salzburg, Austria

We are typically aware of what we know and what we don't know. We may sometimes underestimate our ignorance, but we can think about it. Being able to think (cognize) about one's own knowledge or ignorance is called "metacognition." When do children become aware of their own ignorance? Are animals able to reflect on their lack of knowledge? Why are we so sure that we can do so? ...Because we can talk about it.

There is surprisingly little direct research on when children develop this ability. There is quite a large body of literature on when children come to understand what other people know or don't know as part of their "theory of mind." When we ask children whether they know, for example, what is inside a box they have never seen before, they find it surprisingly difficult. They often answer with a guess but are then unable to say whether they are just guessing or really know before the age of four or even six. Do they have merely problems with the meaning of the word "know" or are they genuinely oblivious about what they know or don't know? Trying to circumvent the use of language with children puts us in the same position as researchers who study metacognition in animals.

Several recent studies have tried to show that monkeys and apes can distinguish between when they do know and when they do not know or are uncertain about what to do. Two basic procedures have been used: Information search and opt out. Chimpanzees, who have seen that a peanut was hidden in one of two pipes are more likely to check by peering through the pipes before choosing than when they have seen in which pipe the bait was put. Children as young as 2½ years do that too. Clearly apes and children behave differently when they know where something is than when they do not, but that does not mean that they have to be aware that they know or don't know.

The opt-out studies are more elaborate. Monkeys undergo training in a cage with initially two response keys and a display. They are trained to press key #1 if the display shows a certain number of points (2950 pixels) and key #2 if there are fewer points. If they press the correct key they get a peanut otherwise they get a mild disappointment in form of a delay before they can get to work again. Even when well trained, for any number of points close to the target of 2950 pixels the animals still make a lot of errors because they are uncertain about the correct response. At some point a third key is introduced. If they press this key they can opt out of the current trial, so they don't get the big reward but they don't get the annoying time out either but can continue with the next trial right away. Monkeys can learn this with enough training trials and it is tempting to conclude that they are able to learn to reflect on their uncertainty about pressing key #1 or #2 and then opt out by pressing key #3.

An alternative explanation for monkeys successful use of the opt-out key is that they learn to use it for the middle region of pixel density simply because it gives, on average, a better payoff than the other two keys in this region, i.e., without reflecting on their uncertainty about using key #1 or #2. Some data rule out this explanation. When given a completely new kind of display, e.g., shades of

colour, then monkeys tend to press key #3, even though they have never learned anything about what to press under these conditions. This supports again the use of metacognition. New conditions create uncertainty and if monkeys have learned to press key #3 when they are uncertain then they will prefer to press key #3 when a completely new display is used.

Unfortunately there is yet another possibility. Animals may not see all this as a task of which key is the correct one to press. They simply have tendencies to press a key (response tendencies) depending on what goes on in their cage. The strength of these tendencies varies with the likelihood of having been rewarded in the past. So if the display shows close to the maximum pixels the tendency to press key #1 is very strong. This tendency diminishes as pixels on the display become sparser but at the same time the tendency to press key #2 increases. In addition the animal has acquired a tendency to press key #3 regardless of what the display shows. So when the display shows pixels of great uncertainty the tendency to press keys #1 and #2 will be fairly low but the tendency to press #3 will be moderately high as for any value on the display. So, in the region of uncertainty, where the tendencies to press #1 and #2 are jointly at their lowest, the tendency to press #3 will dominate. Now, when given a completely new display (colour shades) there will be no tendency to press either key #1 or #2, only the tendency to press key #3, and that is what these animals do without necessarily having any metacognitive awareness of their uncertainty or ignorance about what to do.

This illustrates how difficult it is to find a measure that shows metacognitive awareness without the flexible use of language. But we should not give up because the question about whether animals and at what age children become able to reflect on their mental life is too interesting to stay unanswered. My purpose here is to highlight the problems of interpretation faced by existing methods in order to improve future methods. Some new ways of testing for fine conceptual distinctions have to be found. For instance, we need to find a way of distinguishing whether behaviour by an ignorant or uncertain animal is due to the animal simply being in that state or to the animal being reflectively aware that it is in that state. Similarly, we need to consider how the animal sees matters, and distinguish whether it responds because it feels uncertain (a metacognitive feeling) or whether it knows that it is in a difficult situation (a purely world oriented cognition).

Thinking About Thinking – Evolutionary, Developmental and Epistemological Aspects
Joëlle Proust, Director of Research, Institut Jean-Nicod, Paris, France

Humans have long been thought to be the only creatures who can be conscious of themselves. A classical explanation for this supposed exclusivity is that we humans have the ability to read our own minds: the ability to interpret behavior as being caused by mental states like beliefs and desires. The notion of “thinking about thinking”, for a non-linguistic animal, seems preposterous. But what if there are ways of thinking about thinking that do not involve reasoning about one’s own mental states? It is possible that a primitive form of accessing to one’s cognitive dispositions does not require the capacity to represent them as mental states. The functions and evolutionary history of metacognition and mind reading are explored in this presentation through evidence from animals and children. By examining the role of metacognition in controlled actions, it is possible to understand why metacognition may have evolved in a species.

Controlled actions, in contrast to automatic ones, need additional information in order to prepare the particular sequence of operations and monitor progress towards the goal, because the internal model of the action needs to be adjusted to the requirements of the present context. For example, to cross a ditch you have to predict whether your present bodily condition allows you to jump that far, and in order to remember the contents of a forgotten shopping list, you have to predict whether you have a reasonable chance of succeeding. So it is not enough to have knowledge about what you can generally do. You need, in addition, to evaluate how well you will do in a specific context, in order to make the best decision. Once the action is performed, you also need to evaluate how well you have done.

While standard actions aim to control the world, cognitive actions aim to control your knowledge state: for instance, they aim to acquire accurate or exhaustive information by adjusting attention resources accordingly. Returning to the example of remembering the forgotten shopping list, you don't just want to reconstruct it, you want to retrieve the list favoring accuracy, accepting some misses, or exhaustively, accepting some false items. This flexibility of cognitive adjustment does not happen in standard actions. The ability to act cognitively also requires sensitivity to one's informational needs, and to the constraints associated with these needs, such as accuracy, exhaustivity and ease of processing. An individual perceives the output of an evaluation of feasibility for a cognitive action as specific emotional states correlating to the subject's subjective uncertainty. Such feelings also guide individuals in recognizing the validity of the cognitive action once it is performed. Thus metacognition does not require individuals to be able to read their own minds. It may only require attention to the kinds of cognitive feedback that will regulate successful commands in given cognitive tasks. Not every species, however, seems endowed with this type of attention. No doubt evolutionary hypotheses can be devised to explain why, for example, such sensitivity has been adaptive for rhesus monkeys, in contrast with capuchin monkeys. Given the existence of such sensitivity, metacognition is used as a "normative compass"; in non-humans, it predicts adequacy in cognitive performance. In humans, it can be extended to more diversified norms of evaluation, such as informativeness, coherence, consensus, or relevance.

The function of metacognition is to monitor feasibility and success in specific cases where individuals have to perform a cognitive, rather than a standard, action. Metacognitive monitoring can be based on associative rules (heuristics) rather than on conceptual knowledge. Some types of metacognitive evaluation, however, might not be available without the capacity to have beliefs about the mind, whether it's one's own or another's. This might be the case, for example, when evaluating whether something said in a conversational exchange was informative, or relevant.

So there are two sources of information for evaluating one's own cognitive competences. Therefore, thinking about thinking, in non-humans, can occur without mindreading. It is still arguable that in humans, mindreading takes precedence over metacognition. Knowledge made available to oneself through mindreading might be automatically used in self-evaluation and thus replace associative rules and feelings; reciprocally, no metacognitive evaluation, in humans, could be conducted without using concepts of mental states. This assumption is wrong. Metacognitive self-evaluation, in humans, does not always lead to the same decision as one inspired by a theory about cognitive capacities. Predicting learning in others or in oneself does not result in the same judgment, when the prediction is based on monitoring oneself or on a folk theory about learning. This suggests that we share with non-humans a primary ability to form predictions of cognitive success on the basis of experience rather than of a theory about 'what is a mind'.