EUROPEAN SCIENCE FOUNDATION EXPLORATORY WORKSHOP

METACOGNITION AND MENTAL STATE MONITORING

Maffliers, 7-9 December, 2006

Convened by :

Joëlle Proust (Institut Jean-Nicod, CNRS, Paris)

&

Bernard Renault (CNRS, UPR 640-LENA, Pitié-Salpétrière Hospital, Paris)

With the assistance of Sophie Bilardello, Research Assistant, Institut Jean-Nicod, CNRS, Paris

Keywords: metacognition, mind/brain dynamics, working memory, implicit learning, reflexivity

Abstract

The workshop aimed at understanding metacognitive processes in their evolutionary/developmental, dynamical and neurophysiological dimensions. Conceived as an interdisciplinary “brain-storming” workshop, the aim was to allow specialists from neuroscience, neuropsychology, comparative psychology, developmental end experimental psychology, robotics, mathematics, and philosophy of mind, to explore ways of developing future collaborations between European research units on this emergent topic.

1. Executive Summary

This workshop accomplished an important break-through in discussing the interdisciplinary relevance of metacognition, a capacity that seems to have emerged prior to mindreading, and possibly as a precursor for it. All the invited contributors attended the workshop, except for one participant from Ireland, Elena Magno, who was prevented at the last minute to come for personal reasons. Three additional participants were present : Anna Loussouarn, a doctoral student in charge of welcoming the participants and monitoring the sessions. Alexandre Billon, a postdoctoral researcher involved in the METACOGNITION CNCC Collective Research Project, and Prof. Louise Röska-Hardy, a philosopher from Mainz University, interested in mindreading and intentional attribution, who actively contributed to the discussion of the papers. Selecting a conference center located in Maffliers, a small village in the vicinity of Charles-de-Gaulle Airport, proved ideal for allowing rich exchanges both during and after the programmed sessions. The nice ambiance of an eighteenth century dinner-room, the tasty breaks,and the beautiful park all contributed to the success of the meeting. The alternance of short 30–mn presentations with 30-mn discussions was particularly effective to fuel scientific exchange. Discussions were always lively and prolonged during the breaks and the meals.

The conference was opened by a brief introduction by Joëlle Proust, who in turn introduced Dr. Gretty Mirdal, (Standing Committee for the Humanities) and Dr. Agnès Gruart (Standing Committee for the European Medical Research Councils), both representatives of the ESF. Gretty presented to the group the various networking and research activities sponsored by ESF.

The first session, « Phylogeny and ontogeny of metacognition, empirical and conceptual issues », was then opened by Josep Call (Max Planck Institute for Evolutionary Anthropology, Leipzig), with a stimulating talk entitled "Comparative metacognition: past and future challenges". Josep Call summarized past experiments on what chimpanzees know about what others can and cannot see, and presented future directions of research concerning metamemory in macaques and in chimpanzees (do they know when they have forgotten ?). This first talk and the subsequent discussion allowed registering the promises and difficulties raised by the existing paradigms of test declining as evidence for metacognition in non-human animals as well as ways of coping with them.

After a short break, Josef Perner, from the University of Salzburg, presented an interestingly controversial contribution on "Episodic memory and Theory-of-mind: The role of direct experience and mental imagery in development". Perner, who is one of the pionners and current leaders of mindreading research, argued that a metarepresentational understanding of the mind is a central precondition for episodic memory to develop in children. A lively discussion ensued, several participants objecting that such metarepresentation might not be necessary for episodic memory to occur. The next presentation, by Francisco Pons (University of Aalborg, Denmark) was entitled "Theory of Mind, Emotion Understanding, Language, and Working Memory in Children". Francisco convincingly defended the view that although mindreading is a necessary condition for young children to understanding emotion, it is not sufficient. His paper raised several important questions about the relations between executive and mentalizing capacities, and about the interaction between attention (metacognitive component) and emotional understanding (a mindreading component). After lunch break, Richard Breheny (University College London) presented an insightful paper on "The Mechanisms of Human Communication ». He defended a non-Gricean view of intention understanding in conversation based on the notion of a shared situation. Joint attention, on this view, rather than requiring a full-blown metarepresentational understanding of one's sharing information with another, is a matter of attending to the situation. A subsequent exciting discussion focussed on the relations between attending to a situation in human and non human primates. For technical reasons, the order of the next two speakers was switched. Esther Schlüter, (Sarland University Hospital) first addressed the issue of "The phylogeny of metacognitive processes: lessons from comparative neurophysiology of behavioural control in human and non-human species". This issue indeed is central for the evolution of metacognition. She argued that behavioral and brain imagery evidence suggests that working memory is subserved by two different systems ; an older system relies on sensory and spatial object features, while a newer system consists in a verbal rehearsal mechanism. This paper raised many questions as it questions the view of a central executive controlling cognition : metacognition rather depends on emergent properties of domain-specific processes. The existence of two kinds of executive stores might account for the discrepancies in reflexive control between humans and non human.

Ingar Brinck (Lund University) gave next a dense paper entitled "Attention-Based Metacognition". Ingar Brinck addressed the question of early preverbal metacognition , as being already manifest in pointing behavior. Context, and attention shifting, on her view, are the main cognitive ressources available to extend our human cognition. Metacognition is subserved by attention and intention selection. A discussion ensued as to what « meta » means and on whether protoconversation relies on metacognition.

On Friday december 8th, the second session on « The dynamics of metacognition: conceptual, empirical and formal viewpoints » opened with a presentation by Joëlle Proust (CNRS, Institut Jean-Nicod, Paris) entitled "Metacognition without Metarepresentation". She argued that metacognition has distinctive features that allow a sharp contrast to be made with metarepresentation, such as its predictive character, its normative dimension and the engagement it requires in self-simulation. A stimulating debate ensued with Josef Perner on the plausibility of this distinction from a developmental and comparative viewpoint. Jérôme Dokic (EHESS, Institut

Jean-Nicod, Paris) made a presentation entitled "Cognitive shortcuts: the case of epistemic feelings". He presented and discussed five models of the predictive value of the metacognitive feelings. Such models might each capture specific aspects of metacognition (see the difference between Tip of the tongue and feeling of knowing) . The discussion mainly addressed the issue of wh-questions, and whether they are available to non-linguistic agents.

After lunch, Hélène Frankowska (CREA, Ecole Polytechnique, Paris) presented her groundbreaking work on "A model of learning of sequences by basal ganglia: optimal control, viability theory and dimension reduction". She argued from a case study that Viability Theory allows reducing the dimensionality of a learning task, and offers a more parcimonious and effective way of representing subjects' performance. Questions to the speaker included the difference between her approach and classical statistical views, as well as how Viability Theory allows representing inverse models in a more tractable way than Kalman filters do. Peter Gärdenfors (Lund University), spoke next about the « Control theoretic aspects of intersubjectivity ». His presentation described how emulators generate hidden variables that explain causal mechanisms, whether physical or mental. Successive mental capacities can be understood in terms of emulators subserving them. The subsequent discussion emphasized that standard control theory does not tell us how the brain learns – a topic that brings us back to Frankowska's ideas on learning the viability kernel. Marius Usher (Birkbeck College, University of London), right after the break, addressed the issue of "Computational theories of metacognition". He suggested that total activation of units and Hopfield energy jointly provide good implicit measures for Feeling of Knowing and Feeling of warmth, which in turn suggests that no second order representation needs to develop for these feelings to affect behavior. The discussion examined applications of this innovative model to various forms of epistemic feelings and the predictions that can be made of metacognitive illusions based on certain patterns of energy thoughout the network. The last speaker of the session, Christian Balkenius (Lund University) dealt with "Anticipation & Monitoring in Robotics Systems". He described two robotic implementations of anticipatory mechanisms. The first learns to track dynamic objects and gradually develops a model of the environment through a Kalman-filter based tracking method. The second consists of a group of mobile robots with coinciding goals. The discussion concentrated on how these designs might or not apply to metacognition.

The third and last session of the workshop on « the psychopathology of metacognition: functional and philosophical issues » was opened on Saturday morning by Bernard Renault, (UPR640-LENA, Hôpital de la Salpétrière, Paris): on the topic : "Are event-related potentials relevant tools for the study of metacognition?" He argued that there are different P300 waves when the information is processed implicitly or explicitly. In prosopagnosia, for example, patients have a parietal P300 when they consciously recognize a familiar face, and a frontocentral P300 when they only have an implicit form of recognition. Discussion concentrated on the capacity to use this technique for judgments of uncertainty in blindsighters and for conscious visual perception of ambiguous forms. Nathalie Camille (University of Cambridge) presented in turn her work under the title "The Involvement of the Orbitofrontal Cortex in the Experience of Regret". The speaker constrasted regret – which originates in comparisons betwen factual decision outcome and counterfactual outcome – with disappointment – which only involves comparing an expected with an observed state

of the world. She emphasized the role of orbitofrontal cortex in regret. A lively discussion with Francisco Pons raised the question whether regret qualifies as an emotion. The session came to a close with Jean Lorenceau (CNRS, LENA): « Attending, processing and deciding: Psychophysics & Neuroscience ». Can one force oneself to look at the stimulus in a certain way ? To answer this question, Lorenceau presented his work on perceptual "illusions" and drew on signal detection theory to raise some metholodological questions that a metacognitive approach needs to answer. The discussion focused on the limitations that the brain encounters when having to take into account several probabilistic predictions in multidimensional tasks.

2. Scientific content of the event

The scientific content of our collective research can be summarized using the thematic structure of the workshop.

The phylogeny and ontogeny of metacognition : this workshop has shown that the case for metacognition as a precursor of mindreading is far from universally recognized at this point. Several promising directions of research, however, have been pointed out.

- One consists in improving controlled tests for judgments of uncertainty, metamemory and controlled information search in non-human primates. Some of these tests are currently performed by members of the workshop and/or by researchers participating in the METACOGNITION -CRP.
- Another consists in exploring metacognition in infancy and early childhood. Present evidence seems to be restricted to a few domains, such as : A) early pragmatic understanding of conversation in toddlers, predating by far access to mindreading capacities. B) judgments of source, usually taken to appear at the same time as mindreading. C) episodic memory and metamemory, also shown to coincide with mindreading. More studies need to be performed, involving *implici*t or *non-verbal* tests, and in particular *direct regulation* of (epistemic, motivational, emotional or conative) states rather than explicit verbal attributions.
- A last very promising approach will consist in studying the evolution of brain structures. Following Gruber's lead, one needs to collect comparative neurophysiological evidence of a functional distinction between two working memory stores, and of their possible cooperation/dissociation in humans in various metacognitive tasks.

The second session on the **dynamics of metacognition** demonstrated the crucial impact of modeling on theorizing. Three ideas emerged concerning how such modeling could fuel research on metacognition.

- A first basic, widely accepted view is that metacognition needs to be understood not only at the symbolic expressive level, but also as a set of procedures through which the mind monitors and controls its own informational processes in given motivational contexts. The accepted view contrasts a classical attributive view of mental states to an adaptive control view, and a static approach of mental content to a dynamic modeling

approach. The contrast does not aim at eliminating mental contents as verbally described, but rather at explaining their relations to a procedural non-linguistic level where contents are generated.

- A second type of modeling has been proposed using connectionist ideas, such as Hopfield energy in a neural network. This kind of modeling might help clarify some aspects of how self prediction can be performed.
- Metacognition at large should however benefit from a type of mathematics drawing directly on adaptive control theory. Viability Theory is based on the idea that the main goal of a coevolutionary system consists in finding the limits of its « kernel of viability ». Such an informational goal is supposed to drive learning, and allow reduce drastically the dimensionality of the task space. One of the main theoretical goals to which metacognitive research is confronted consists in developing further viability theory to help new models of metacognition to emerge, incorporating into traditional control views new ideas about how the mind represents the world as affording envelopes of viable trajectories.

The last session on **the psychopathology of metacognition** suffered from the absence of Dr. Magno, as well as from the difficulty of neuropsychologists and psychiatrists to accept the 3-day format of our workshop. This situation is particulary unfortunate, given the relevance of neuropsychology to explore metacognition in its different facets through naturally occurring dissociations. In spite of this, several important results were scored. One is a confirmation of a dissociation found in patients with schizophrenia, between a maintained or disturbed capacity to monitor one's cognitive states and a disturbed or maintained capacity to adjust control to monitoring. For example, it was found that orbitofrontal patients may experience regret while not adjusting their strategy as a consequence of their felt regret. Another important finding in this domain was that EEG may prove to be a powerful way of investigating implicit forms of metacognition. This kind of method might be used to ascertain the existence of implicit perceptual metacognition in patients with blindsight.

3. Assessment of the results, contribution to the future direction of the field, outcome

On the basis of these new results, work will be conducted within the ESF CNCC funded METACOGNITION CRP as well as in other projects related to metacognition in order to take advantage of them. Clearly, theorizing on metacognition can only benefit from expanding comparative research and neurophysiological exploration, as well as from a more systematic exploration of the metacognitive deficits in human and non human-animals. In parallel, dynamic modeling on metacognition should be encouraged using either classical connectionist or adaptive control formalisms. Most of all, the workshop reveals the need for developing a networking activity targetted to neuropsychology and cognitive psychophysiology, where metacognition is highly relevant to understand the force of delusion or anosognosia, the value of insight, and the possibility of confabulation.

4. Final programme

Wednesday 6 December 2006

Evening Arrival and dinner (8:30 pm)

Thursday 7 December 2006

Session A: Phylogeny and ontogeny of metacognition, empirical and conceptual issues

08:45-09:00 Welcome and Introduction

09:00-09:15 Presentation of the European Science Foundation (ESF) Agnès Gruart (Standing Committee for the European Medical Research Councils) and Gretty Mirdal (Standing Committee for the Humanities)

09:15-09:45 Josep Call (Max Planck Institute for Evolutionary Anthropology, Leipzig): "Comparative metacognition: past and future challenges"

09:45-10:15 Discussion

10:15-10:30 Pause

10:30-11:00 Josef Perner (University of Salzburg): "Episodic memory and Theory-of-mind: The role of direct experience and mental imagery in development"

11:00-11h30 Discussion

11:30-12:00 Francisco Pons (University of Aalborg, Denmark): "Theory of Mind, Emotion Understanding, Language, and Working Memory in Children"

12:00-12:30 Discussion

12:30-14:00 Lunch

14:00-14:30 Richard Breheny (University College London): "The Mechanisms of Human Communication"

14:30-15:00 Discussion

15:00-15:30 Ingar Brinck (Lund University, SE): "Attention-Based Metacognition"

15:30-16:00 Discussion

16:00-16:15 Pause

16:15-16:45 Esther Schlüter (Sarland University Hospital): "The phylogeny of metacognitive processes: lessons from comparative neurophysiology of behavioural control in human and non-human species”

16:45-17:15 Discussion

Evening Dinner: 19:30

Friday 8 December 2006

Session B: The dynamics of metacognition: conceptual, empirical and formal viewpoints

10:00-10:30 Joëlle Proust (CNRS, Institut Jean-Nicod, Paris): "Metacognition without Metarepresentation"

10:30-11:00 Discussion

11:00-11:15 Pause

11:15-11:45 Jérôme Dokic (EHESS, Institut Jean-Nicod, Paris): "Cognitive shortcuts: the case of epistemic feelings"

11:45-12:15 Discussion

12:15-14:00 Lunch

14:00-14:30 Hélène Frankowska (CREA, Ecole Polytechnique, Paris): "A model of learning of sequences by basal ganglia: optimal control, viability theory and dimension reduction"

14:30-15:00 Discussion

15:00-15:30 Peter Gärdenfors (Lund University, SE): "Control-theoretic aspects of intersubjectivity - Mind-reading as control theory"

15:30-16:00 Discussion

16:00-16:15 Pause

16:15-16:45 Marius Usher (Birkbeck College, University of London): "Computational theories of metacognition"

16:45-17:15 Discussion

17:15-17:45 Christian Balkenius (Lund University, SE): "Anticipation & Monitoring in Robotics Systems"

17:45-18:15 Discussion

Evening: Dinner: 20:00

Saturday 9 December 2006

Session C: The psychopathology of metacognition: functional and philosophical issues

09:30-10:00 Bernard Renault (UPR640-LENA, Hôpital de la Salpétrière, Paris): "Are event-related potentials relevant tools for the study of metacognition?"

10:00-10:30 Discussion

10:30-10:45 Pause

10:45-11:15 Nathalie Camille (University of Cambridge): "The Involvement of the Orbitofrontal Cortex in the Experience of Regret".

11:15-11:45 Discussion

11:45-12:15 Jean Lorenceau (CNRS, LENA): "Attending, processing and deciding: Psychophysics & Neuroscience".

12:15-12:45 Discussion and concluding remarks

13:00-14:30 Lunch

5. List of participants

1. Joëlle Proust,
Institut Jean-Nicod (CNRS), Philosophie, EHESS-ENS Paris, (Philosophy of mind; neurophilosophy; schizophrenia; animal cognition)

2. Bernard Renault,
CNRS, UPR 640-LENA, Cognitive Neurosciences and Brain Imaging, Pitié-Salpétrière Hospital, Paris (cognitive neuroscience, covert and overt recognition, event related potentials, EEG and MEG spatio-temporal analysis.)

3. Jérôme Dokic,
EHESS, Institut Jean-Nicod (CNRS), Paris, (Philosophy of mind; mental dynamics; reflexivity)

4. Hélène Frankowska,
CREA, Ecole Polytechnique, Paris (Mathematics for Dynamic systems; viability theory; adaptive control.)

5. Jean Lorenceau
CNRS, UPR 640-LENA. Pitié-Salpétrière Hospital, Paris, Equipe: Liages Dynamiques: Formes, Mouvement, Actions.

6. Anna Loussouarn, master student, Philosophy, Institut Jean-Nicod, Paris

7. Alexandre Billon, postdoctoral researcher, Philosophy, Institut Jean-Nicod, Paris

GERMANY

8. Josep Call,
Co-director Wolfgang Köhler Primate Research Center, Max Planck Institute for Evolutionary Anthropology, Leipzig.

9. Esther Schlüter,
Klinik und Poliklinik für Psychiatrie und Psychotherapie, Universität Göttingen`

10. Louise Röska-Hardy, Universität Mainz.

SWEDEN

11. Peter Gärdenfors,
Lund University, Professor of Philosophy, Chairman of the Cognitive Science Program. (evolution of mind; model of revision of belief; feedforward and feedback modelling)

12. Christian Balkenius,
Associate Professor, Cognitive robotics, Lund University.

13. Ingar Brinck,
Associate Professor, Lund University (philosophy of mind, evolution of cognition, joint attention).

AUSTRIA

14. Josef Perner,
Developmental psychology, department of psychology, University of Salzburg (episodic memory in children; theory of mind).

DENMARK

15. Fancisco Pons,
Professor of Developmental Psychology, Faculty of Humanities, University of Aalborg, Denmark (emotion understanding)

UNITED KINGDOM

16. Nathalie Camille,
Behavioural and Clinical Neurosciences Institute, Department of Experimental Psychology, University of Cambridge (Cognitive neuroscience, experience of regret. Orbitofrontal cortex. Counterfactual thinking. Decision.)

17. Richard Breheny,
Department of Linguistics, University of College London

18. Marius Usher,
Institute of Cognitive Neuroscience, Dept. of Psychology, Birkbeck College (metacognition, feeling of knowing, dynamics of prediction).

6. Statistical information on participants (age bracket, countries of origin, etc.)

- Junior researchers (postdoc level or less) : 4/18 = 22,22%

- Midcareer researchers : 9/18 = 50%

- Senior researchers : 5/18 = 27,78 %

- Austria : 1 = 5 ,5%
- Denmark : 1 = 5,5%
- France : 7 = 39% ; (speakers : 5 = 27,78 %)
- Germany : 3 = 16,66 %
- United Kingdom : 3 = 16,66 %
- Sweden : 3 = 16,66 %