Metacognitive Networks and Measures of Consciousness

Pasquali Antoine (Antoine.Pasquali@Ulb.Ac.Be)
Consciousness, Cognition, & Computation Group, Université Libre de Bruxelles CP 191,
Av. F.-D. Roosevelt, 50, 1050 Bruxelles, Belgium

Timmermans Bert (Bert.Timmermans@Ulb.Ac.Be)
Consciousness, Cognition, & Computation Group, Université Libre de Bruxelles CP 191,
Av. F.-D. Roosevelt, 50, 1050 Bruxelles, Belgium

Cleeremans Axel (Axcleer@Ulb.Ac.Be)
Consciousness, Cognition, & Computation Group, Université Libre de Bruxelles CP 191,
Av. F.-D. Roosevelt, 50, 1050 Bruxelles, Belgium

Abstract

Subjective measures of awareness rest on the assumption that conscious knowledge is knowledge that participants *know* they possess. Post-decision wagering, recently proposed as an objective measure of awareness, raised a new controversy on determining the properties that should characterize the objectivity of an awareness measure. Indeed, if the method appears objective in many aspects – it does not require introspection but rather lies on instinct, it does not affect conscious states, it can be learned unconsciously –, it also shares some characteristics with subjective measures – it involves metacognitive content and particularly, it represents a decision about a decision. The lack of consensus on this topic leaded us to develop a new approach based on a novel theoretical aspect, causality, and to consider a causally independent mechanism that would give an agent the capability to know what knowledge it possesses. In this framework, any measure that would not necessarily rely on such mechanism in a given experimental situation should be considered as objective. We support our claim with a computational model based on metacognitive networks, and present three simulation studies in which neural networks learn to wager on their own performance. Results demonstrate a good fit to human data, although depending on the situation, post-decision wagering is implemented either as an objective or as a subjective measure of network's knowledge. We discuss implications of our results for defining the nature of subjective and objective measures, as well as for our understanding of consciousness.

Keywords: awareness measures; metacognitive networks; wagering.

Objective and Subjective Measures

Awareness can be assessed by using objective (e.g., cued-report tests, forced-choice decisions) or subjective (e.g., verbal reports and confidence judgments) measures (Merikle, 1992). Though this remains controversial (Holender & Duscherer, 2004; Tunney & Shanks, 2003), by using subjective measures one can conclude that performance on some task of interest is guided by unconscious knowledge whenever participants claim to be guessing while nevertheless performing better than chance – the "guessing criterion" (Dienes & Berry, 1997) –, or, alternatively, when their confidence appears unrelated to their accuracy – the "zero correlation criterion". Therefore, considering an objective threshold determined by an accuracy at chance, any gap between an objective and a subjective threshold would be considered as a region of unconscious processing (Koch & Preuschoff, 2007). This reasoning presupposes that when one is conscious of some piece of information, one is also conscious that one holds this information. This requires metaknowledge – content and attitude explicit – about the information (Dienes & Berry, 1997). Then a lack of metaknowledge would refer to unconscious knowledge, and a dissociation is observable between the objective and subjective measures that could track that knowledge.

Post-decision wagering (PDW) was recently introduced as a new objective measure of awareness and tested on three different situations by Persaud and his colleagues (Persaud, McLeod & Cowey, 2007; Persaud & McLeod, 2007). In PDW, participants continuously evaluate their performance by wagering on each decision in tasks such as visual stimulus localization under condition of blindsight (Stoerig, Zontanou & Cowey, 2002), string classification in artificial grammar learning (AGLT) (Reber, 1967), and deck selection in the Iowa Gambling Task (IGT) (Bechara et al., 1994; Maia & McClelland, 2004). Wagering is intuitive for participants and offers a quantitative way of assessing the relationship between performance and awareness: Given that participants attempt to maximize their earnings, when wagering is independent from above-chance performance (i.e. below the "zero correlation criterion"), one may conclude that the knowledge that drives their performance is unconscious (blindsight, implicit learning). On the other hand, any positive relationship between wagering and accuracy should be taken as reflecting to some extent participants' knowledge about the basis for their decisions (suprathreshold stimulus, explicit learning).

Therefore PDW stands as a measure of awareness. If it effectively appears more objective than subjective measures, their functioning is equivalent in many aspects: Both PDW and subjective measures are supported by metacognitive content and may be vulnerable to biases in that they require a decision about a decision, leading to the controversy about if PDW objectively reflects awareness or not, as it does not directly measures sensory consciousness (Seth, 2008a). But still, as for objective measures, it is capable of being learned unconsciously. Moreover, it depends on a process that does not involve introspection and that does not affect conscious states, as verbal reports and confidence judgments usually do (Koch & Preuschoff, 2007).

With this hybrid character of PDW, the line of distinction between objective and subjective measures has become blurred. However, even by putting aside this new measure and by considering only preexisting ones, the line is already unclear. Indeed, one can arguably consider that confidence judgments are a particular kind of cued-report tests, as they considerably relate to accuracy on a binary scale (Tunney & Shanks, 2003) and as their discrepancy with performance nearly vanishes under the right experimental conditions (Holender & Duscherer, 2004). Therefore one can reasonably assume that unconscious perception simply reflects conscious perception under high uncertainty, whereas another would fill the gap between objective and subjective thresholds with some sort of partial awareness (Kouider & Dupoux, 2004). Subjective measures would thus be depicted as mere objective measures at a more advanced level of treatment, or alternatively said, of a weaker type regarding perceptual awareness and of a stronger type regarding conceptual awareness. If this very simple idea appears very appealing, we however consider that a theory of consciousness that offers a set of measures that only vary in terms of their efficiency in tracking awareness will never help us to assess whether or not unconscious and conscious processes can be dissociated. Indeed, we will hardly distinguish cases when one measure tracks the same awareness as another but from a more indirect, biased or slower process, from when the two measures track two different levels of awareness, and may be, two different *types* of awareness.

It has been proposed that it would be much more significant if measures of awareness could show that unconscious and conscious processes lead to qualitatively different consequences (Merikle, 1992). But focusing only on consequences is misleading, since both objective and subjective measures are and will ever be only behavioral measures, not phenomenal measures. One can say that knowledge has *indirect* effects that can be measured *directly* by objective measures (Dienes & Berry, 1997). But in fact, as long as they rely on behavioral expression, both types of measures are *indirect* (Seth, 2008b). For the same reason that no task is process-pure (Destrebecqz & Cleeremans, 2001), no measure should be considered purely objective or purely subjective. This mistake might explain why we still cannot decide whether unconscious processing exists or not.

Considering this new theoretical disillusion, we propose to detach our attention from objective and subjective *consequences* of knowledge in our measures and, instead, to consider a definition of objectivity and subjectivity from a *causal* perspective. Our claim rest on the core assumption (Cleeremans, 2008; Cleeremans, Timmermans & Pasquali, 2007) that evaluating one's own performance, as involved in subjective measures of awareness, requires re-describing first-order representations responsible for performance into metarepresentations that continuously inform the agent about its own success: Knowledge that is *in* the system must become knowledge *for* the system (Karmiloff-Smith, 1992). As previously pointed out (Dienes & Perner, 1996; Dienes & Berry, 1997), representing knowledge into metarepresentations (content explicit) is not sufficient. One must also represent himself as being in the possession of that content (attitude explicit). This requires *access* to the relevant first-order knowledge in a manner that is *independent* from the causal chain in which it is embedded. A failure in this access would automatically reflect a dissociation between performance and awareness measures.

To achieve the mechanism, we assume that a higher-order network continuously monitors the performance of a first-order network responsible for performing a task, in such a way that it is (a) able to discriminate and classify information contained in the first-order network *in a independent manner*, and (b) able to use this information to perform a secondary task that requires knowledge *about* first-order internal state or performance, such as subjective measures do. Without claiming that this mechanism presents sufficient or even necessary conditions for awareness in general, we intend to show that it can account for participants' wagering behavior as reported in the experiments of Persaud et al. (2007). We suggest that more complex but similar processes are involved when subjective measures are employed.

Metacognitive networks

We thereafter present a computational model that will illustrate the functioning of such mechanism. For this, we conceive neural networks, so called metacognitive networks, which can learn to know about their own internal representations of some stimulus domain. In these experiments, we propose a putative mechanism through which some sort of self-knowledge can accrue by exploring the performance of metacognitive networks trained to wager on their own decisions. We constructed three networks that simulate in a minimal way the blindsight, the AGLT, and the IGT tasks proposed by Persaud et al. (2007), and that reproduce their findings. We created comparable experimental situations; still including several simplifications for the only aim is to illustrate our claim. In our simulations, a metacognitive network consists of (a) a three-layers backpropagation feedforward first-order network (FoN) that performs the main task (depending on the task: stimulus localization, letter string classification, or deck selection), and (b) a second-order network (SoN) that evaluates the performance of the FoN on a binary scale, and that consists of a hidden unit layer and of two output nodes (high/low wager). Because of different task requirements, all three networks associated with the three experiments differ substantially in terms of (a) the nature of their higher-order representations, and (b) the implementation of the "low" and "high" awareness conditions.

The fairly similar blindsight and AGLT architectures have

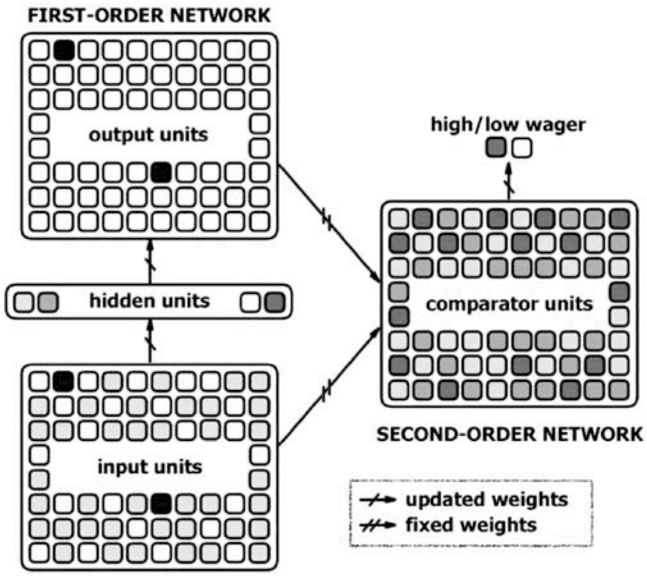

Figure 1: Network architecture for Blindsight and AGLT simulations.

in common that the FoN is an autoassociator (Figure 1), while the SoN hidden units consist of a comparator matrix, representing the match between the input and the output of the FoN. The use of comparators in these simulations illustrates perfectly what we mean by *independent causality* in that their functioning and performance are independent (a) from the first-order task and (b) from the quality of the first-order representations. Independency could obviously be achieved with many other functions. However, the crucial role of such comparators for the emergence of conscious percepts has been suggested previously (Frith, Blakemore & Wolpert, 2000; Synofzik, Vosgerau & Newen, 2008; Gallagher, 2004; Mandler, 2004), in that they merge internal and external states into unique representations (Sperry, 1950; Wolpert & Kawato, 1998; Pacherie, 2008; Rizzolatti et al., 1996). In our networks, each of these comparators computes the difference between each corresponding pair of FoN input and output units, and thus represents the FoN's internal error *not as a training signal but as an activation pattern* that is accessible for the system. Here building of metarepresentations is innate and unsupervised, but this is only after a period of supervised training in the SoN that they can thereafter drive wagering advantageously. This means that (a) metarepresentations emerge automatically, without requiring any learning, whereas (b) access to these patterns for wagering is learned over time, independent of specific first-order patterns at the current task. Thus in these two first simulations, wagering measures reflect the subjective knowledge that the network automatically forms about its own internal states.

On the other hand, the architecture of the IGT does not involve neither an autoassociator nor a comparator (Figure 2). The FoN operates this time a supervised predictive task, and the SoN is directly plugged onto the hidden units of the FoN. No automatic and independent metarepresentations are assumed, but the hidden representations of the SoN that have been learned over practice, this time *specifically depending on the first-order patterns and task*. This different architecture illustrates another kind of meta-representations, created by simple reinforcement between two tasks of interest, which reflect a dependent access to the internal knowledge. With this implementation, meta-representations lack the attitude explicitness that would be necessary for the network in order to *know* that it possesses internal knowledge. Network's wagering ability is enhanced along the learning phase by trial and error, demonstrating a mechanism that could have been developed unconsciously under experimental conditions. PDW is therefore implemented in this third simulation as an objective measure of network's internal knowledge.

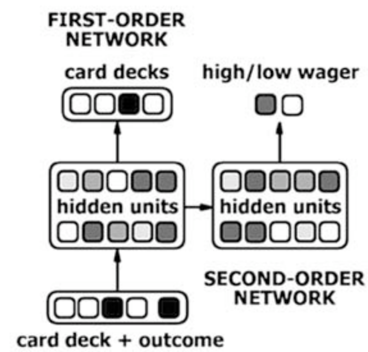

Figure 2: Network architecture for the IGT simulation.

Simulations and Results

In their Blindsight experiment, Persaud et al. (2007) showed that blindsight subject GY (a patient who, under specific circumstances, makes visual discriminations in the absence of visual awareness), when presented with subthreshold stimuli in his blind field, displayed above chance localization performance but failed to maximize his earnings through wagering. However, for suprathreshold stimuli (both in normal and blind fields), GY maximized performance as well as earnings. We simulated these results by pre-training 15 networks to discriminate patterns and to simultaneously place wagers on their own performance. The distinction between subthreshold and suprathreshold visions in blindsight situations was only introduced during a following testing phase, in which the networks classified the patterns they had previously been presented with (suprathreshold stimuli), as well as degraded versions of these patterns in which stimulus-to-noise ratio was manipulated (subthreshold stimuli). If Blindsight is commonly associated with lesions in the brain and more particularly in V1, we thought that modulating the stimulus-to-noise in the input of the metacognitive network would have more computational interest than cutting connections between the FoN and the SoN, since a dissociation would obviously have occurred in the latter case. As shown in Table 1, the simulations closely capture Persaud et al.'s results. Discrimination performance, as simulated by the

Table 1: Results of the Blindsight simulation.

Localization with Subthreshold stimuli	Experiment			Simulation		
	Correct	**Incorrect**	**Total**	**Correct**	**Incorrect**	**Total**
High Wager	<u>12</u>	6	18	<u>29</u>	2	31
Low Wager	62	<u>20</u>	82	49,5	<u>19,5</u>	69
Total	74	26	100	78,5	21,5	100
Suprathreshold Stimuli	**Correct**	**Incorrect**	**Total**	**Correct**	**Incorrect**	**Total**
High Wager	<u>72</u>	2	74	<u>50,5</u>	4,5	55
Low Wager	18	<u>8</u>	26	29,5	<u>15,5</u>	45
Total	90	10	100	80	20	100

Percentages of localizations and corresponding wagers in low **(subthreshold)** and high **(suprathreshold)** consciousness conditions in Persaud et al.'s experiment (reproduced with permission) and in our simulation. Optimal wagers are underlined.

Table 2: Results for the AGLT simulation.

	Experiment			Simulation		
Implicit Condition	**Correct**	**Incorrect**	**Total**	**Correct**	**Incorrect**	**Total**
High Wager	<u>36</u>	6,5	42,5	<u>36,5</u>	8,5	45
Low Wager	44,5	<u>13</u>	57,5	35,5	<u>19,5</u>	55
Total	80,5	19,5	100	72	28	100
Explicit Condition	**Correct**	**Incorrect**	**Total**	**Correct**	**Incorrect**	**Total**
High Wager	<u>⬀</u>	⬂	⬀	<u>63,5</u>	0,5	64
Low Wager	⬂	<u>⬂</u>	⬂	34,5	<u>1,5</u>	36
Total	⬀	⬂	100	98	2	100

Percentages of discriminations and corresponding wagers in low **(Implicit condition)** and high **(Explicit condition)** consciousness conditions in Persaud et al.'s experiment (reproduced with permission; exact explicit condition was not performed but effects were discussed) and in our simulation. Optimal wagers are underlined.

FoN, is well above chance both under subthreshold and suprathreshold conditions (78.7% and 80.1% correct, respectively). However, networks tested in subthreshold condition fail to support their correct and incorrect discriminations with advantageous wagers (high and low, respectively), and instead perform at chance level (48.6% of all trials are followed by an advantageous wager). This is not the case under suprathreshold condition (65.9% of all trials are followed by an advantageous wager).

In the Artificial Grammar Learning experiment, Persaud et al. show that, following exposure to artificial grammar strings under incidental learning conditions (i.e., memorize "TSXVPP", "PVPXVT", etc.), participants perform above chance on subsequent discrimination of novel strings, while failing to maximize their earnings though advantageous wagering (Implicit condition). When participants were thereafter made aware of the grammar rules (Explicit condition), they started to wager more advantageously. Discrimination performance should also have improved in this condition but experimenters maintained it as the same level as in the Implicit condition by reducing the time of exposure to the strings during the test phase. We simulated these results by training two sets of 15 networks to classify artificial grammar strings. The metacognitive networks were similar to those used in the Blindsight simulation, except that we implemented the distinction between high and low awareness conditions only by manipulating the FoN training phase length (short and long training phases corresponding to Implicit and Explicit conditions, respectively). Obviously we could not *tell* the networks what were the grammar rules as experimenters did with participants. However, in order to illustrate our claim, we only needed to induce the same improvements as in the experimental situation. As both performance and wagering had to be enhanced in our Explicit condition, we just let the FoN learn for a longer period in order to obtain the same effects. Nonetheless we did not integrate in our simulation their manipulation of the duration of subjects' exposure to the strings that maintained performance at the same level in both Implicit and Explicit conditions. Therefore we shall not compare our results to their Explicit condition, but to the effects that occur when exposure's duration is unchanged. Here again results fit Persaud et al.'s data (Table 2). Networks in the Implicit condition performed above chance (71.8% correct), but failed to optimize their wagering (55.8% of all trials). Whereas in the Explicit condition, networks were not only better at the discrimination task (98.2% correct) – as a longer exposure duration would have predicted experimentally –, but also in placing advantageous wagers above chance (64.9% of all trials).

We performed a simpler implementation of the Iowa Gambling Tasks experimented by Maia & McClelland (2004) or by Persaud et al. (2007). In Persaud's version, participants' task is to select one of four decks of cards, each

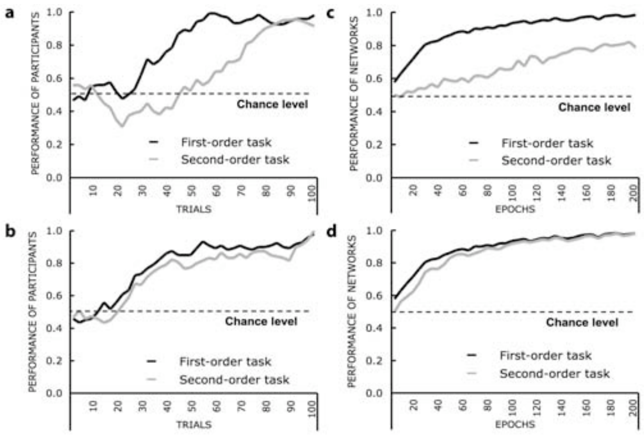

Figure 3: Results for the IGT simulation.
Performance is plotted across time (epochs) for (a; c) low consciousness, and (b; d) high consciousness conditions in Persaud et al.'s experiment (reproduced with permission) and in our simulation respectively.

with different pay-offs. After deck selection, but before turning over the card (revealing how much was won or lost), participants wager on whether the card will be winning or losing. Participants typically manage to optimize deck selection well before they start wagering advantageously. However when participants are made more aware of their strategy to determine deck relative pay-offs by asking them the specific questions proposed by Maia & McClelland (2004) (i.e., "What would you expect your average winning amount to be by picking 10 cards from deck 1?"), wagering follows performance more closely (Persaud et al., 2007). This experiment differs from the two others in three aspects, necessitating a different simulation approach. First, the IGT required participants to initially explore the material before being able to create any representation about the decks' yields. The resulting metarepresentations are thus necessarily dependent on this exploration phase. Second, participants received feedback on each trial about the quality of their wagering, as the turning of the card immediately revealed whether and how much they had won. As a consequence, participants could have used this feedback to unconsciously optimize not only their deck selection, but also their wagering. Certainly participants effectively became aware of relative pay-offs along the experiment, but this makes wagering in the IGT less suitable as a *subjective* measure of awareness since advantageous wagering could *in principle* emerge without subjectivity. Third, high awareness condition affected participants' wagering measures but not their performance, as if the specific questions revealed clues that were useful for the development of a wagering strategy but not for the game's first task. To simulate these results, two sets of 15 networks learned to perform the deck selection task while wagering on the gain upon each decision. Implementation was modified in three main aspects in order to reflect the task's differences. First, the FoN could not be an autoassociator since the desired states were not available as inputs as in the previous simulations. Indeed the networks had to select one out of four card packs at first and received feedback about the quality of that selection (win or loss) only after. Second and as a consequence, the SoN metarepresentations could not rely on a comparator, but instead consisted of a hidden layer directly connected to FoN hidden units and feeding forward into the (wagering) outputs (Figure 2). Finally, in order to modulate only the wagering measures, we implemented the distinction between Low Consciousness and High Consciousness conditions by setting the SoN learning rate low or high, respectively. Figure 3 displays the performance at both tasks (desk selection and wagering) over time. Like in Persaud et al.'s data, in the Low Consciousness condition, wagering performance lags advantageous card deck selection. However, under a High Consciousness condition, in which in simulation we increased the speed with which the SoN could make use of the FoN hidden units' representations, wagering closely follows card deck selection.

Discussion

The simulations demonstrate three ways of modeling wagering behavior such as reported experimentally by Persaud et al. (2007). In all three situations metacognitive networks can be trained to take advantage of their internal states in order to drive an advantageous wagering. Through the development of metarepresentations that re-describe relevant first-order knowledge, each network is able to capture the dissociation between performance and wagering under conditions of low and high awareness as it has been measured experimentally. However, if wagering measures lead to comparable consequences in the three experiments, a distinction can now be drawn considering the implementations of the Blindsight and the AGLT simulations on one hand, and that of the IGT simulation on the other. The difference of implementation points towards two different types of metarepresentations. Though both involve re-descriptions of first-order representations, for the second type (the architecture used for the IGT), emergence of metarepresentations occurs through supervised reinforcement by the secondary task, which directly taps into the knowledge used in the primary task. Whereas for the first type (blindsight and AGLT simulations), metarepresentations are accessed without learning through an automatic, non-supervised re-descriptive process, unrelated a priori to any specific secondary task. Therefore the distinction is based on the causal nature of the metarepresentations, and this nature defines if wagering is an objective or a subjective measure of awareness. Thus, *the wagering measure can be of either type depending on the experiment*. This explains why PDW elicited so much controversy.

Though we do not claim that metacognitive networks are aware in any sense, we would like to discuss three fundamental issues on which we believe these simulations shed some light. First, the basic assumption underlying the mechanism is that metacognition requires higher-order representations, content and attitude explicit, as minimally simulated by the SoNs in the first two experiments. This means that task-related knowledge must become available for the metacognitive agent, outside of the causal chain in which it is embedded when performing that task. In other words, the knowledge has to become accessible content about which one can form metacognitive content, use it and possibly become subjectively conscious of it. Second, in order to illustrate the mechanism, we used second-order comparators that form a representation of the difference between the current internal state or prediction of the first-order network, and the effective external state corresponding to the input it received. As suggested before (Frith, Blakemore & Wolpert, 2000; Synofzik, Vosgerau & Newen, 2008; Gallagher, 2004; Mandler, 2004; Sperry, 1950; Wolpert & Kawato, 1998; Pacherie, 2008; Rizzolatti et al., 1996), such comparators may play a crucial role in the emergence of consciousness – and particularly the sense of agency – as they inform an agent about the adequacy of its own internal states. Large discrepancies between a prediction and the corresponding outcome form conscious events, and precisely, comparators can detect them. Third, it appears that different types of metarepresentations coexist in our brain. Even if the descriptions we have made do not inform much about the phenomenology of consciousness, one might find this causally independent mechanism appropriate for operating access consciousness. Moreover, it could produce the higher-order representations that render their observed content conscious, when they themselves become subject of further higher-order representations, as it is described by the HOT theory (Rosenthal, 1997).

In conclusion, the main point raised by our model is that metacognitive content can emerge naturally through a process of continuous representational redescription (Karmiloff-Smith, 1992; Clark & Karmiloff-Smith, 1993) that is itself learned over time. This in turn suggests that *the mind learns to be conscious*, as it learns to use knowledge from unconscious lower-order processes to develop higher-order representations that inform it about the geography of its own internal states.

Acknowledgments

We thank N. Persaud, D. Rosenthal and J.L. McClelland for helpful discussions. A.P. is a Research Assistant of the National Fund for Scientific Research (FRS - FNRS, Belgium). B.T. is a postdoctoral fellow at the ULB - Université Libre de Bruxelles. A.C. is a Research Director with the National Fund for Scientific Research (FRS - FNRS, Belgium). This work is supported by Concerted Research Action 06/11-342 titled "Culturally modified organisms: What it means to be human in the age of culture", financed by the Ministère de la Communauté Française – Direction Générale l'Enseignement non obligatoire et de la Recherche scientifique (Belgium); by European Commission Grant #043457 "Mindbridge – Measuring Consciousness"; and by FRFC / ESF Grant #2.4577.06 "Mechanisms of serial action".

References

Bechara, A., Damasio, A.R., Damasio, H., & Anderson, S.W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, 50:7-15.

Clark, A., & Karmiloff-Smith, A. (1993). The Cognizer's Innards: a psychological and philosophical perspective on the development of thought. *Mind and Language*, 8:488-519.

Cleeremans, A., Timmermans, B., & Pasquali, A. (2007). Consciousness and metarepresentation: A computational sketch. *Neural Networks*, 20:1032-1039.

Cleeremans, A. (2008). The radical plasticity thesis. *Progress in brain research*, 168:19-33.

Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin & Review*, 8:343-350.

Dienes, Z., & Perner, J. (1996). Implicit knowledge in people and connectionist networks. In *Implicit cognition*, ed Underwood G, Oxford: Oxford University Press, 227-256.

Dienes, Z., & Berry, D. (1997). Implicit learning: Below the subjective threshold. *Psychonomic Bulletin & Review*, 4:3-23.

Frith, C.D., Blakemore, S.J., & Wolpert, D.M. (2000). Abnormalities in the awareness and control of action. *Phil. Trans. R. Soc. Lond.*, B 355:1771-1788.

Gallagher, S. (2004). Neurocognitive models of schizophrenia: A neurophenomenological critique. *Psychopathology*, 37:8-19.

Holender, D., & Duscherer, K. (2004). Unconscious perception: The need for a paradigm shift. *Perception & Psychophysics*, 66:872-881.

Karmiloff-Smith, A. (1992). *Beyond modularity : A developmental perspective on cognitive science*. Cambridge: MIT Press.

Koch, C., & Preuschoff, K. (2007). Betting the house on consciousness. *Nature Neuroscience*, 10:140-141.

Kouider, S., & Dupoux, E. (2004). Partial Awareness Creates the "Illusion" of Subliminal Semantic Priming. *Psychological Science*, 15:75-81.

Maia, T.V., & McClelland, J.L. (2004). A reexamination of the evidence for the somatic marker hypothesis: what participants really know in the Iowa gambling task. *Proc. Natl. Acad. Sci. USA*, 101:16075-16080.

Mandler, J.M. (2004). *The foundation of mind: Origins of conceptual thought*. New York: Oxford University Press.

Merikle, P.M. (1992). Perception without awareness: critical issues. *American Psychologist*, 47:792-795.

Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition*, 107:179–217.

Persaud, N., & McLeod, P. (2007). Wagering demonstrates subconscious processing in a binary exclusion task. *Consciousness and Cognition*, 17:565-575.

Persaud, N., McLeod, P., & Cowey, A. (2007). Post-decision wagering objectively measures awareness. *Nature Neuroscience*, 10:257-261.

Reber, A.S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning & Verbal Behavior*, 6:855-863.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3:131-141.

Rosenthal, D. (1997). A theory of consciousness. In *The Nature of Consciousness: Philosophical Debates*, eds Block, N., Flanagan, O., Güzeldere, G., Cambridge, MA: MIT Press.

Seth, A.K. (2008a). Post-decision wagering measures metacognitive content, not sensory consciousness. *Consciousness and Cognition*, 17:981-983.

Seth, A.K. (2008b). Theories and measures of consciousness develop together. *Consciousness and Cognition*, 17:986-988.

Sperry, R.W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. J. *Comp. Physiol. Psychol.*, 43:482-489.

Stoerig, P., Zontanou, A., & Cowey, A. (2002). Aware or unaware: assessment of cortical blindness in four men and a monkey. *Cerebral Cortex*, 12:565-574.

Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: A multifactorial two-step account of agency. *Consciousness and Cognition*, 17:219-239.

Tunney, R.J., & Shanks, D.R. (2003). Subjective measures of awareness and implicit cognition. *Memory & Cognition*, 31:1060-1071.

Wolpert, D.M., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, 11:1317-1329.