Types of Cognition and Its Implications for Future High-Level Cognitive Machines

Camilo Miguel Signorelli

Department of Experimental and Health Sciences, Universidad Pompeu Fabra, Barcelona, Spain

Centro Interdisciplinario de Neurociencias de Valparaíso, Universidad de Valparaíso, Chile

camiguel@uc.cl

Abstract

This work summarizes part of current knowledge on High-level Cognitive process and its relation with biological hardware. Thus, it is possible to identify some paradoxes which could impact the development of future technologies and artificial intelligence: we may make a High-level Cognitive Machine, sacrificing the principal attribute of a machine, its accuracy.

Introduction

Intelligence is one of the main characteristics of the Human Brain. For example, cognitive abilities such as problem-solving, language skills, social skills, among others, are considered as High-level cognitive processes with a strong impact on different types of intelligences. Understanding how these abilities work can help us to implement desirable social behaviors in the new generation of robots with Human-Machine interactions.

How do cognitive, abstract abilities and subjective experience emerge from a physical system like the brain? (Chalmers, 1995; Dehaene, et al. 2014; Tononi, et al. 2016) How can we implement these abilities in a natural form in machines? Is it possible? What are the requirements to achieve this? These are some of the most complex questions in science and to resolve them, we need to understand the biological basis, the cellular dynamics as well as the physical and mathematics basis to know the fundamental properties of matter which could eventually be involved in these emergent properties.

One requirement for High-level cognitive abilities seems to be what it is called subjective experience (Barron & Klein, 2016) or in a more complex order: consciousness. For example, we first need to be conscious to take some complex rational decisions, to plan, and to do something with intent (Baars, 2005; Tononi & Koch, 2008). How biology implements rational or high intelligences, is completely different from how we started to implement it in computer science (Moravec, 1988).

In order to improve machines socials skills, one should try to understand how these abilities emerge in the brain and also, in order to understand how these abilities work and emerge from the brain, we should study how to implement them into machines. This is why the last test or full understanding of subjective experience and consciousness will be to make a conscious machine.

This work will show a brief of what it is known about conscious states and some paradoxes that emerge from the idea of creating a conscious machine: it may be possible to make a conscious robot, but it does not come without a cost.

Consciousness and Intelligence

Consciousness, a brief notion

If consciousness could be defined, the questions above would be easily answered. We may also have a recipe with different "ingredients" to explain consciousness in humans, animals, and accept or reject the option to implement it in machines.

However, this is not the case. All of us understand what consciousness is, what it is to be aware, what it is to have a subjective experience and on which "elements" our experience is based. Sometimes these concepts are considered as synonyms, other times as different ideas with different meanings. Thus, it is needed to make some distinctions. The modern approach to consciousness has made a differentiation between **conscious states** as different levels of awareness (vegetative, sleep, anesthesia, altered states, aware); **contents of consciousness** as elements or information in the external or internal world which at times are part of our conscious state; and **conscious processing** as

the operations applied to these contents (Dehaene, et al. 2014). Thanks to these definitions, theories as Global Workspace (Baars, 2005), Integrated Information (Tononi, et al. 2016) and Dynamic theories (Varela, et al. 2001) have developed many tools leading to a greater understanding of consciousness. In these approaches, the consciousness and eventually intentionality are understood as an emergent property from brain networks dynamics.

Nevertheless, these definitions have some difficulties explaining how this emergency works, explaining subjectivity experiences (Chalmers, 1995; Lycan & Dennett, 1993) or sometimes avoiding the importance of previous learning to develop consciousness at early stage of childhood (Cleeremans, 2011).

Even though this article will refer to the above definitions, as we are referring to previous works that use these definitions; our own understanding of consciousness, unlike these theories, might be defined as a process of processes which mainly interferes with the neural integration. In this case, consciousness would be an operation more than a container where one might put something. It implies that consciousness could emerge from neural network dynamics as an autonomous system (e.g. brain oscillations), at a different space of action than the original system, but with sub-emergent properties interfering with the space from which it originates. This work will interpret certain results in this direction with the intention of expanding this notion in future projects. Also, while the author agrees with many of the conclusions, results and progresses of the theories mentioned, it is reasonable to consider that a new focus that integrates all of them is needed, since they share many similarities and are at the same time complementary.

The concept of consciousness involves awareness, emotions, subjectivity, intentionality, and attention, among others. Consciousness is formed from all of these processes like a differentiated and unified whole, but it is not any of them. For example, it could be necessary to be aware to have emotions and subjective experiences, or maybe vice versa, and we will need them to show intentionality and attention. We also distinguish that these are different processes as, for example, awareness and attention; while understanding all of them as constituent parts of what we define as consciousness.

Consciousness and its relation with Intelligence
Intelligence is something a little simpler to define; at least, many scientists think so. Intelligence can be understood as the ability to solve problems in an efficient way. This means, the maximization of the positive results in our solution while minimizing the negative impacts, for instance, waste of time. To do that, other processes as Learning and Memory are also needed and associated with the definition of intelligence. Learning may be understood as the process to gain new knowledge or improve some behavior, while

Memory is the storage of this knowledge. To solve problems efficiently, it is necessary to access a certain memory that was acquired thanks to a specific learning that will be modifying again the memory of the system. The more intelligent is the system, the more it learns.

This framework also allows finding mathematical implementations of intelligence and its use in machines. Why could consciousness be necessary in this context? First, many different types of intelligences have already recognized, between them, the emotional intelligence that helps us with social interaction. Secondly, emotions play a crucial role in learning, consolidation of memories, retrieved memory and cognition in general (Cleeremans, 2011). In other words, emotions, as subjective experiences, and cognition are deeply related into human intelligence (Haladjian & Montemayor, 2016).

Finally, emotions and subjective experiences are also constitutive parts of consciousness, as stated above. This means consciousness and High-level cognitive abilities are also related. In consequence, to improve complex intelligence and social behavior in machines, we should understand constitutive parts of consciousness, and how its processes affect the computability and accuracy of the most powerful system that we know to date, our brains.

What we already know
First: The Hardware
One simplification about the brain code is to think in the neuron as the principal informative unit for the brain codification and the spikes as the language of its code. It assumes one kind of codification between 0, without spike, and 1 with a spike.

However, now it is known that the electrical transmission through gap junctions (Bennett & Zukin, 2004) are not only between other neurons but also between cells as astrocytes (Alvarez-maubecin, et al. 2000). It is also known that neuromodulatory substances can reconfigure neuronal circuits within minutes and hours (Nusbaum, et al. 2001). In some cases, these substances applied in axon (Bucher, Thirumalai, & Marder, 2003; Goaillard, et al. 2004) can spontaneously initiate action potentials in a non-classical mode of integration. Another research has shown that action potential in some neurons can travel backward from the axon and soma regions into dendrites (Stuart, Dodt, & Sakmann, 1993), in contrary to Polarization law. Other examples are: The functional complexity of dendrites and the roles they play, the influence of non-neuronal cells in the action potential propagation (Dupree, et al. 2005) and axon-glia communication, where the information is transduced through cells that are not neurons (Fields & Stevens-graham, 2005). Astrocytes synapses regulation (Coggan, et al. 2005; Dani, Chernjavsky, & Smith, 1992; Fields & Stevens-graham, 2005; Matsui & Jahr, 2004) and its influence on neurons and so on. All this new knowledge is not

included in the idea of spike codification or is definitely contrary to it.

These findings have shown us that the neuron is not the unique codification unit in the brain code and in consequence the spike is not the unique key to understand its codification (Llinas, et al. 1998). It implies that the binary code is not enough to understand the brain.

Second: Computability
Turing and Gödel were the first in demonstrating the theorem of incompleteness (Gödel, 1931; Turing, 1937). This theorem showed that it is not possible to resolve all the problems with a deductive process or algorithm. From their demonstrations, one of the consequences suggested is the potential impossibility to build conscious machines with algorithms only. Physiological and Psychological data also seem to show that this is not possible (Haladjian & Montemayor, 2016). This, together with point one, means that if we can reach the gap to build a conscious machine, it will not be with the algorithmic computation. A new computational framework will be needed.

Third: The two Systems
When we try to explain, on one side, the biological mechanisms in the brain, and on the other, the human psychological behavioral, some paradoxes appear. Some research and theories have shown that the dynamic of neural systems can be interpreted in a classic probabilities framework (Pouget, et al. 2000; Quiroga & Panzeri, 2009), like good estimator and predictor of external stimuli. While other results, principally from economic psychology, show cognitive fallacies (Gilovich, et al. 2002, Ellsberg, 1961; Machina, 2009; Moore, 2002). These results are incompatible with the classical probability theories (Pothos & Busemeyer, 2013).

Hence, these disconnections between the behavior of the human brain and the cognitive fallacies may help to understand better how the human brain works. How can some cognitive capabilities, with apparently non-classical dynamic, emerge from apparently classical systems as neural networks? It is necessary to answer this question as a requirement to start to think in conscious machines.

Forth: Four types of cognition
A recent paper by Shea and Frith (Shea & Frith, 2016) presents a valuable differentiation between contents and cognitive processes (computations from some contents to others). These categories can be conscious or not. From the experimental evidence and with this framework, they define four cognitive categories (Table 1): i) Type 0 Cognition: non-conscious contents and non-conscious (non-controlled) processing. Optimal choice is computation-light but learning-heavy (e.g. Motor Control). ii) Type 1 Cognition: conscious contents but automatic and non-controlled processing (e.g. Fallacy questions). The system accesses to a wider range of information (Holistic infor-

mation), however some optimal calculations may become computationally intractable. iii) Type 2 Cognition: contents are conscious and the cognitive process is deliberate and controlled (e.g. Reasoning). It is computation-heavy, learning-light and interferes with Type 0. iv) Type ∞: non-conscious content but conscious process. Yet no further definition is given as they are not sure if this category could exist (Supra Reasoning information).

	Non conscious Processing	Conscious Processing
Contents non conscious	Type 0 Cognition	Type ∞ Cognition
Contents conscious	Type 1 Cognition	Type 2 Cognition

Table 1. Types of cognition according different contents and processing exposed above.

Fifth: The perception is a discrete process
The perception needs time. We are conscious of a perceptual content in more o less 200 ms up 400 ms. Indeed, evidence suggests that the conscious perception is a discrete process (Dehaene & Changeux, 2011; Dehaene, et al. 2014; VanRullen & Koch, 2003) in contrast to a continuous process. The differences between fast time processing for cognition type 0 (order of 40 ms) and a slow time processing for type 1 (order of 200 ms) have stimulated the idea of the Two-Stage Model (Herzog, Kammer, & Scharnowski, 2016). The first stage would be a fast, continuous and classic probabilistic process (type 0 cognition) while the second stage would correspond to a slow and discrete process (type 1 and 2 cognition) due to loops in information processing (Figure 1). In relation with the discussion above, one may also say that the second stage would obey non-classical probabilities.

Implications and Paradoxes

The four types of cognitive categories, together with the other points that were mentioned previously, open the door for some new interpretations and classifications on the types of machines (Figure 1):

• The Machine-Machine: It could correspond to the type 0 cognition. The examples are the robots that we are making today with a high learning curve.

• The Conscious Machine: A machine with the ability to be conscious of its contents but not of the process. In this case, we will sacrifice accuracy in favor of consciousness. The type 1 cognition requires integration of information and it is a time consuming process. This is why the accuracy is lost instead of consciousness. However, is this machine useful for something? Here is the first Paradox. If we can make a Conscious Machine Type 1 Cognition, this machine will lose the most important attribute of being a machine: the accuracy in the calculation. In this case, the machine does not work how we expect: it is a machine without being a machine.

- The Super Machine: A machine with the ability to be conscious of its contents and own processes. This machine will have subjective experience at the same time that it will have the option to control the accuracy of its own rational process. However, the type 2 cognition needs more time to find accuracy solutions than type 1. This process also interferes with type 0 cognition, and in consequence robots will have less precision in motor tasks. In this case, one observation appears: we cannot expect an algorithm to control the process of emergency of consciousness (point 2) and we will not be able to control it. In other words, the machine can do whatever it wants; it has the power to do it and the intention to do it. It could be considered biological new species, more than a machine. The paradox is that the Super Machine is not useful anymore unless the machine wants to collaborate with us.
- The Subjective Machine: Machine type ∞ cognition. Some modern progress in artificial intelligence might fit in this category (Bringsjord, et al. 2015). The machine will be conscious of each computation process but not conscious of the contents, at least, they will not be aware in the sense that we are. What does this mean? Probably the Machine will be able to control the computational processing; however it will not be able to take meanings of contents from these computations. In other words, the machine will not be able to "feel something", to have "emotions", the machine will have a type of subjective experience that we do not know yet. This is probably the evolution of machine type 0 that we are making these days but without the intermediate steps type 1 and 2 cognitions. However, if this machine does not feel like a human, will it be able to understand us? It could be the kind of machine in science fiction. Again, the machine could be useless for our purpose and even worst, it could be contrary to our purposes.

Based on these observations, it is possible to summarize all the paradoxes in one type: The Conscious Machine is not a useful machine anymore. In other words, the machine losses the meaningful characteristics of being a machine: to resolve problems with accuracy, speed and obedience.

Some Alternatives

It could be interesting for Artificial Intelligence to make Machine type 1 or type 2 cognition. Based on the principles exposed above from biological systems, the requirements are: i) new non-binary codification principles, ii) no deterministic computation, iii) two different systems of integration of information, which should be related as the non-classical probabilistic system emerging from the classical probabilistic system, iv) enough computation capability to support the first time-consuming system and the second non-classical system. Is it possible to respond to these requirements with the actual framework in Artificial Intelligence?
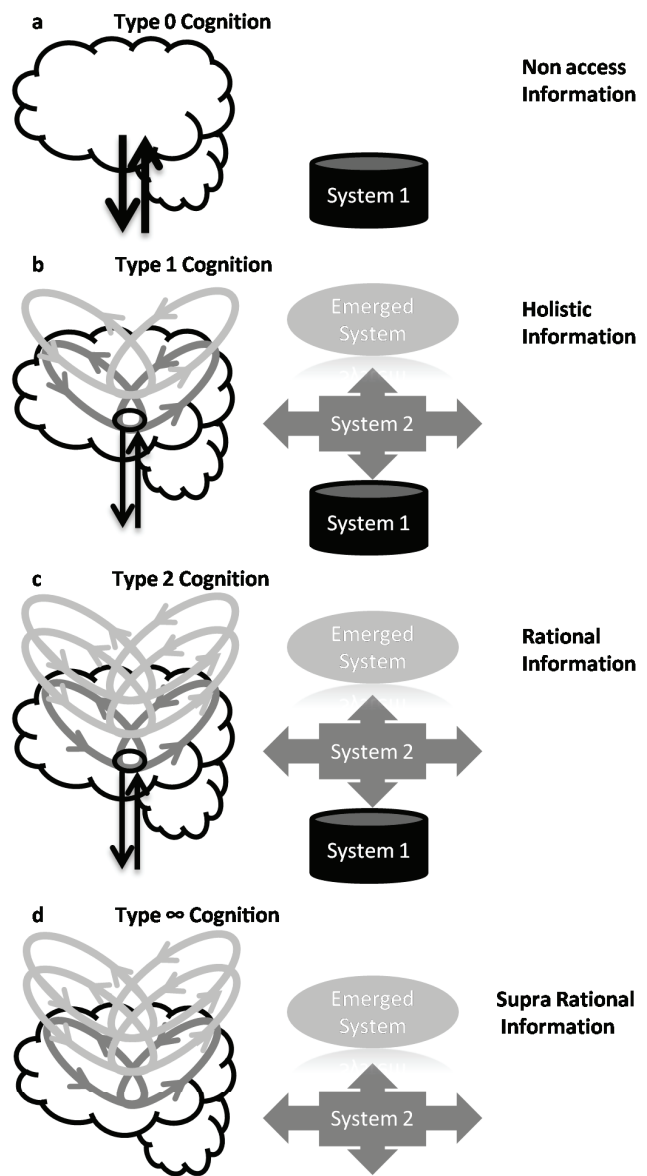

Figure 1. Types of cognition related with potential Systems of processing in the brain.

Quantum computation is one of the alternatives due to the increase of computational capability thanks to qubits, non-classical probabilities and non-classical computability. Some alternative tools aiming to implement Quantum computation in analogy with brain process, might be new approaches of contextuality from Quantum Physics to Psychology (Abramsky, 2015), like contextual semantics, natural language, and semantics of computation (Kartsaklis, et al. 2013). Starting from the explanation of the cognitive fallacies above; concept combinations, human judgments and decision making under uncertainty can be better modeled and described by using the mathematical formalism of quantum theory (Aerts, et al. 2013; Busemeyer, et al. 2011; Sozzo, 2015; Wang, et al. 2014). This emerging domain

has been called "quantum cognition" (Bruza, Wang, & Busemeyer, 2015) and these models may help us understand the connection between new computational frameworks, psychophysics, recent advances in neuroscience on conscious perception in the brain, and also how to make conscious machines.

Conclusions

These ideas are part of a work in progress. This work has developed a hypothesis based on previous works: we may build conscious machines but with computational costs. These costs reflect some paradoxes when we examine the requirements to build a conscious machine: it will be possible to make a conscious machine without good rational abilities or a good rational conscious machine with more time to respond accuracy. In both cases, the price of non-obedience will be paid. The next step in this research will be to quantify the decrease in computational efficiency, time-consuming of these processes and the mathematical formulation of these propositions.

Acknowledgments

References

Abramsky, S. 2015. Contextual Semantics: From Quantum Mechanics to Logic, Databases, Constraints, and Complexity. In *Contextuality from Quantum Physics to Psychology* (Vol. 6, pp. 23–50).

Aerts, D., Gabora, L., & Sozzo, S. 2013. Concepts and their dynamics: a quantum-theoretic modeling of human thought. *Topics in Cognitive Science*, *5*(4), 737–72.

Aerts, D., & Sozzo, S. 2011. Quantum structure in cognition: Why and how concepts are entangled. In *Quantum Interaction. Lecture Notes in Computer Science* (Vol. 7052, pp. 116–127).

Alvarez-maubecin, V., Garcı, F., Williams, J. T., & Bockstaele, E. J. Van. 2000. Functional Coupling between Neurons and Glia. *The Journal of Neuroscience*, *20*(11), 4091–4098.

Baars, B. J. 2005. Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. *Progress in Brain Research*, *150*, 45–53.

Barron, A. B., & Klein, C. 2016. What insects can tell us about the origins of consciousness. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(18), 4900–8.

Bennett, M. V. L., & Zukin, R. S. 2004. Neuronal Synchronization in the Mammalian Brain. *Neuron*, *41*, 495–511.

Bringsjord, S., Licato, J., Sundar, N., Rikhiya, G., & Atriya, G. 2015. Real Robots that Pass Human Tests of Self-Consciousness. In *Proceeding of the 24th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 498–504).

Bruza, P. D., Wang, Z., & Busemeyer, J. R. 2015. Quantum cognition: a new theoretical approach to psychology. *Trends in Cognitive Sciences*, *19*(7), 383–93.

Bucher, D., Thirumalai, V., & Marder, E. 2003. Axonal Dopamine Receptors Activate Peripheral Spike Initiation in a Stomatogastric Motor Neuron. *The Journal of Neuroscience*, *23*(17), 6866–6875.

Busemeyer, J. R., Pothos, E. M., Franco, R., & Trueblood, J. S. 2011. A quantum theoretical explanation for probability judgment errors. *Psychological Review*, *118*(2), 193–218.

Chalmers, D. 1995. The puzzle of conscious experience. *Scientific American*, *273*(6), 80–86.

Cleeremans, A. 2011. The Radical Plasticity Thesis: How the Brain Learns to be Conscious. *Frontiers in Psychology*, *2*(May), 86.

Coggan, J. S., Bartol, T. M., Esquenazi, E., Stiles, J. R., Lamont, S., Martone, M. E., … Sejnowski, T. J. 2005. Evidence for Ectopic Neurotransmission at a Neuronal Synapse. *Science*, *309*(5733), 446–451.

Dani, J. W., Chernjavsky, A., & Smith, S. J. 1992. Neuronal activity triggers calcium waves in hippocampal astrocyte networks. *Neuron*, *8*(3), 429–440.

Dehaene, S., & Changeux, J.-P. 2011. Experimental and theoretical approaches to conscious processing. *Neuron*, *70*(2), 200–27.

Dehaene, S., Charles, L., King, J.-R., & Marti, S. 2014. Toward a computational theory of conscious processing. *Current Opinion in Neurobiology*, *25*(1947), 76–84.

Dupree, J. L., Mason, J. L., Marcus, J. R., Stull, M., Levinson, R., Matsushima, G. K., & Popko, B. 2005. Oligodendrocytes assist in the maintenance of sodium channel clusters independent of the myelin sheath. *Neuron Glia Biology*, *1*(03), 179–192.

Fields, R. D., & Stevens-graham, B. 2005. New Insights into Neuron-Glia Communication. *Science*, *298*(5593), 556–562.

Gilovich, T., Griffin, D., & Kahneman, D. 2002. *Heuristics and biases: The psychology of intuitive judgment*. Cambridge University Press.

Goaillard, J.-M., Schulz, D. J., Kilman, V. L., & Marder, E. 2004. Octopamine modulates the axons of modulatory projection neurons. *The Journal of Neuroscience*, *24*(32), 7063–73.

Gödel, K. 1931. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte Für Mathematik Und Physik*, *38-38*(1), 173–198.

Haladjian, H. H., & Montemayor, C. 2016. Artificial consciousness and the consciousness-attention dissociation. *Consciousness and Cognition*, *45*, 210–225.

Herzog, M. H., Kammer, T., & Scharnowski, F. 2016. Time Slices: What Is the Duration of a Percept? *PLoS Biology*, *14*(4), e1002433.

Kartsaklis, D., Sadrzadeh, M., Pulman, S., & Coecke, B. 2014. Reasoning about meaning in natural language with compact closed categories and frobenius algebras. *Logic and Algebraic Structures in Quantum Computing and Information*, (pp 199–222). Cambridge University Press.

Llinas, R., Ribary, U., Contreras, D., & Pedroarena, C. 1998. The neuronal basis for consciousness. *Philosophical Transactions of the Royal Society of London B*, *353*, 1841–1849.

Lycan, W. G., & Dennett, D. C. 1993. Consciousness Explained. *The Philosophical Review*, *102*(3), 424.

Matsui, K., & Jahr, C. E. 2004. Differential control of synaptic and ectopic vesicular release of glutamate. *The Journal of Neuroscience*, *24*(41), 8932–9.

Moravec, H. P. 1988. *Mind children : the future of robot and human intelligence* (p. 214). Harvard University Press.

Nusbaum, M. P., Blitz, D. M., Swensen, A. M., Wood, D., & Marder, E. 2001. The roles of co-transmission in neural network modulation. *Trends in Neurosciences*, *24*(3), 146–154.

Pothos, E. M., & Busemeyer, J. R. 2013. Can quantum probability provide a new direction for cognitive modeling? *The Behavioral and Brain Sciences*, *36*(3), 255–74.

Pouget, A., Dayan, P., & Zemel, R. 2000. Information processing with population codes. *Nature Reviews Neuroscience*, *1*(11), 125–132.

Quian Quiroga, R., & Panzeri, S. 2009. Extracting information from neuronal populations: information theory and decoding approaches. *Nature Reviews. Neuroscience*, *10*(3), 173–85.

Shea, N., & Frith, C. D. 2016. Dual-process theories and consciousness: the case for "Type Zero" cognition. *Neuroscience of Consciousness*, *2016*(1), 1–10.

Sozzo, S. 2015. Conjunction and negation of natural concepts: A quantum-theoretic modeling. *Journal of Mathematical Psychology*, *66*, 83–102.

Stuart, G. J., Dodt, H. U., & Sakmann, B. 1993. Patch-clamp recordings from the soma and dendrites of neurons in brain slices using infrared video microscopy. *Pflugers Archv European Journal of Physiology*, *423*, 511–518.

Tononi, G., Boly, M., Massimini, M., & Koch, C. 2016. Integrated information theory: from consciousness to its physical substrate. *Nature Reviews. Neuroscience*, *17*(7), 450–61.

Tononi, G., & Koch, C. 2008. The neural correlates of consciousness: an update. *Annals of the New York Academy of Sciences*, *1124*, 239–61.

Turing, A. M. 1937. On Computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, *42*(1), 230–265.

VanRullen, R., & Koch, C. 2003. Is perception discrete or continuous? *Trends in Cognitive Sciences*, *7*(5), 207–213.

Varela, F., Lachaux, J., Rodriguez, E., & Martinerie, J. 2001. The brainweb: phase synchronization and large-scale integration. *Nature Reviews Neuroscience*, *2*(April), 229–239.

Wang, Z., Busemeyer, J. R., Atmanspacher, H., & Pothos, E. M. 2013. The potential of using quantum theory to build models of cognition. *Topics in Cognitive Science*, *5*(4), 672–88.

Wang, Z., Solloway, T., Shiffrin, R. M., & Busemeyer, J. R. 2014. Context effects produced by question orders reveal quantum nature of human judgments. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(26), 9431–6.