On the practical nature of artificial qualia

Raúl Arrabales[1], Agapito Ledezma[1] and Araceli Sanchis[1]

Abstract. Can machines ever have qualia? Can we build robots with inner worlds of subjective experience? Will qualia experienced by robots be comparable to subjective human experience? Is the young field of Machine Consciousness (MC) ready to answer these questions? In this paper, rather than trying to answer these questions directly, we argue that a formal definition, or at least a functional characterization, of artificial qualia is required in order to establish valid engineering principles for synthetic phenomenology (SP). Understanding what might be the differences, if any, between natural and artificial qualia is one of the first questions to be answered. Furthermore, if an interim and less ambitious definition of artificial qualia can be outlined, the corresponding model can be implemented and used to shed some light on the very nature of consciousness. In this work we explore current trends in MC and SP from the perspective of artificial qualia, attempting to identify key features that could contribute to a practical characterization of this concept. We focus specifically on potential implementations of artificial qualia as a means to provide a new interdisciplinary tool for research on natural and artificial cognition.

1 INTRODUCTION

It is well known that human conscious perception is not directly based on data acquired by senses, but heavily biased by psychological aspects like, for instance, cognitive context, subject history, and expectations. Additionally, context and subjective history is in turn shaped by the specific way in which stimuli are consciously perceived. Although we tend to believe that we perceive the reality, the fact is that qualia generated in our brains are far from being a truthful representation of real world. Nevertheless, generally our conscious experience of the world proves to be highly reliable in terms of everyday tasks. In short, the world is interpreted by the subject in such a way that is advantageous for his or her goals.

The former intuitive definition of conscious perception does not seem to be sufficient to build a comprehensive model of a conscious machine that includes phenomenal aspects (see for instance [1] for a discussion about the grand illusion of consciousness and perceptual phenomenology). We do not seem to have a satisfactory model or theory about what exactly qualia could be (either in conscious machines, humans or other animals). In other words, the phenomenal dimension of consciousness, both in natural and artificial creatures, remains elusive to scientific study. Additionally, as pointed out by Sloman [2], many times we might be discussing bogus concepts due to the use of misleading contexts and ill-defined terms. Although we are proposing here a set of new definitions, it is not our aim to contribute to the existing confusion in the field of MC, but to help clarifying concepts from the engineering point of view. As argued by Sloman [2], directing the basic questions about consciousness to machines with different designs can help to figure out what really needs to be explained.

While many Artificial Intelligence (AI) implementations cover some aspects of the broader picture of cognition, we have not yet reached the point where human-like conscious machines are possible [3]. One of the greatest challenges that still needs to be addressed is the design of computational models of qualia, i.e. models of artificial qualia. MC designers are in the need of a practical definition of what could be an artificial conscious mind.

In this work, we assume that a complete and scientifically established definition of what artificial qualia could be does not yet exist. Therefore, we propose to circumvent this problem by looking for alternative, interim, and partial definitions that could contribute to the development of MC research field.

This approach does not necessarily lead to better implementations in terms of performance, but the whole exercise might at least provide more insight about what conscious machines could be like, what design strategies appear to be more promising, and how neuroscience could benefit from the application of computational models of qualia. As Chrisley has suggested [4], one of the aims of SP should be to characterize the phenomenal states possessed or modelled in MC implementations.

In the remainder of this paper we attempt to provide a decomposition of the artificial phenomenology problem into more tractable and recognizable steps. It is our hypothesis that the partial definitions of artificial qualia that we outline here may become useful conceptual tools in the domain of SP. Initially, in Section 2, we analyze the context of this research and introduce some of the constraints and difficulties that we have to deal with. Relevant related work is briefly discussed in Section 3. Section 4 covers the specific characterization of artificial qualia that we put forward. Then, in Section 5, the proposed approach is considered from the point of view of practical scientific applications. We illustrate this topic by suggesting that qualia could be studied in machines with relatively simple experiments involving, for instance, the apparent motion effects in visual experience. Finally, conclusions are presented in Section 6.

2 TACKLING THE PROBLEM

Taking into account the complexity of the problem described in the previous section, we have no other option but to bear with some of the well-known but controversial issues of the scientific study of consciousness. In relation with the specification of phenomenal states, it is convenient to distinguish between different components of such an intricate concept as consciousness. Additionally, the way in which phenomenology can be studied satisfactorily has to be analyzed. In particular, the problem of private first-person observations is discussed.

[1] Dept. of Computer Science, Carlos III University of Madrid, 28911. Madrid. Spain. Email: rarrabal@inf.uc3m.es.

2.1 Decomposing a complex concept

Many problems in the science of consciousness are rooted in the fact that the term consciousness can be used to refer to multiple related aspects [5]. In other words, consciousness can be seen as different phenomena according to the perspective of the observer. If we distinguish between phenomenal and functional dimensions of consciousness as suggested by Block [6], the brief description of the perception process outlined in Section 1 can be seen from two different perspectives:

a) *Conscious perception is the set of phenomenal experiences of which our inner life is made of (the 'what is it like' to have experiential mental states* [7]*), e.g. the redness of red.*

b) *Conscious perception is the set of functional representations or internal models of the world adapted to our needs, which are made available for use in reasoning and action, e.g. the neuronal encoding of color in the brain (see* [8] *for details).*

These two views (phenomenal consciousness and access consciousness respectively) should not be considered exclusive or contradictory, but complementary aspects of the same complex process. In fact, natural cognition shares these properties: the conscious contents of our minds are both experiential mental states and also functional representations accessible for reasoning and action. As Haikonen has suggested [9], qualia are the direct products of the perception process, and without qualia there is no consciousness. Therefore, qualia cannot be neglected in the study of consciousness, especially if we want our MC computational models to be of any use to the quest for the understanding of human cognition.

While qualia are usually associated with the first view (phenomenal consciousness), most of the work done in the domain of artificial cognitive systems is exclusively related to the second view (access consciousness). One of the reasons for this bias is the poor comprehension of phenomenal aspects of consciousness. Another significant reason is that plenty of work still needs to be done on machines that apparently do not require qualia for successful performance. A related problem is to determine why some machines might need qualia. These issues are briefly covered in the next subsections.

2.2 Avoiding the first-person/third-person problem

The problem of considering both personal and access views at the same time is that phenomenal consciousness is only available to the first-person observer, i.e. it is a private property [10]. Inner experience in human counterparts is usually inferred by third-person observers based on the similarity to their own case: if I feel pain when I get hurt, I infer other humans will likely feel the same in the same situations (because they have a nervous system like mine). However, when it comes to detecting the presence of phenomenal states in machines we do not even have the similarity argument as a factor to take into account in the inference process.

Being phenomenal qualia inherently private, how could we determine if a machine is experiencing any inner life? (See [11] for a comprehensive discussion about the problem of measuring machine consciousness) This issue is quite related to the so-called hard problem of consciousness [12], which seems to remain an open problem. Essentially, we do not yet have any convincing explanation of phenomenal consciousness (at least, we do not have a theory that could be translated into a computational model).

Does this mean that any attempt to create artificial qualia will be futile? Or should we, instead of giving up the challenge, try to explore machine qualia as a means to shed light on the nature of consciousness? Do we really need to understand the very nature of consciousness in order to reproduce it in machines? Is there a real lack of scientific tools to address the problem of consciousness? Can we develop a model of phenomenal consciousness exclusively based on third-person approaches? Perhaps these questions cannot be answered yet; however, we believe third-person approaches can be successfully applied to make progress in the field of artificial cognitive systems and their application in AI-inspired biology.

Typically, applying third-person approached involved looking exclusively at behaviour, including forms of accurate report [13]. However, external observers are also able to inspect the architecture and inner machinery of the creature. The inner inspection and monitoring of biological living organisms, including humans, is much more problematic than the inspection of working implementation of artificial cognitive systems. Therefore, the analysis of correlations between observed behaviour and internal inspection of MC implementations has to be exploited as it could provide valuable information about the models being tested (without the existing limitations in analogous experiments with biological creatures).

We argue that, following this line of research, a limited definition of artificial qualia can be made in such a way that only third-person approaches are used. This partial definition might not explain phenomenal qualia as it is present in humans, but it could be used to create computational models, and subsequent implementations, which then could be used to enhance our understanding of "natural consciousness".

2.3 The function of qualia

Understanding what is the function of qualia and why they emerged as part of biological evolution is an essential part of the challenge of the scientific study of consciousness. As usual, the research interplay between natural and artificial sciences can be seen both ways: on one hand, a comprehensive understanding of qualia, as they manifest in biological creatures, might make possible the building of conscious machines; on the other hand, the path to a complete understanding of qualia in biology might lay through the research on new computational models focused on phenomenology.

These ideas about qualia are not free of controversies. While some argue that qualia are mere epiphenomena (e.g. [14]), we believe phenomenal consciousness appeared as an evolutionary advantage. One way to prove it would be to compare the performance of phenomenally conscious machines versus unconscious machines, both confronted to complex tasks in unstructured environments. Given that such an experiment is not realizable nowadays, we will focus on both evidence from the biological world and current computational models.

There are many features related with consciousness that we know are useful because they contribute to survival (for instance, Theory of Mind [15], to name an illustrative one). These cognitive capabilities have a function and that is the reason why they have been selected by evolution. But, what about phenomenal aspects of consciousness? Do they have a clear functional role?

Qualia or subjective experience should not be seen as an additional component of the complex notion of consciousness, but a process that is present in relation with cognitive features. In the context of the proposed functional role, qualia are experienced by a creature when it is able to introspect some of its perceptual processes and use that introspection to generate a meta-representation which in turn is used to modulate the whole system. Qualia are indeed the output of the perception process [9], which in some cases are made explicit thanks to a transparent access to the perception process outcome (sensory system response to stimuli). In short, when we are aware of a red object in our field of view, we do not perceive the colour red, but the redness quale, which is the reaction of our perceptual system to the red colour stimulus.

The role of qualia described above could be studied in artificial systems. The generation of artificial qualia along these lines could provide insight about consciousness applicable to biological creatures. As argued by Sloman and Chrisley [16], a machine could even develop private ontologies for referring to its own private perceptual contents and states. The use of this ontology for modulating system processes is the function of qualia. A system with qualia is a system with *meta-management* capabilities (combination of introspection and active control based on self-monitoring). Making a serious effort to design and build such systems will contribute to the confirmation or refutation of these hypotheses about the role of qualia in biological creatures.

3 RELATED WORK

Neglecting the hard problem of consciousness has been one of the most common approaches in the practical design of artificial consciousness models. We think this position should not be criticized; at least when the phenomenal dimension is ignored explicitly, because tackling initially some of the so-called *easy problems* (see [12]) seems to be a plausible engineering strategy to begin with. However, as we pursue more ambitious objectives, like the ones outlined in the previous section, the field of MC should tend to adopt more tools based on the concept of Synthetic Phenomenology [4]. This aspect could be essential for the progress of existing research lines in the direction of human-like machines.

The analysis of the computational correlates of artificial qualia done by Chella and Gaglio is a remarkable example in which the authors have directly addressed the concept of artificial phenomenology [17]. In this work, based on the cognitive architecture developed by the same researchers [18], an active process integrating internal and external flows of information is designed to build a two-dimensional, viewer dependent, reconstruction of the subjective scene as perceived by the robot. It is argued that the matching process between this reconstruction and internal perceptual data represents the artificial qualia of the system.

Haikonen's cognitive architecture is another example of a computational model that takes into account the generation of artificial qualia [9, 19]. It is argued that a realization of the conscious machine proposed by Haikonen would be endowed with an inherent mechanism to produce qualia with grounded meanings. Additionally, the machine would be able to report these qualia via secondary symbols such as uttered words.

Even though other MC implementations have not been initially designed to have phenomenal consciousness, their capacity for sustaining phenomenal states can be explored. This is for instance the case of LIDA [20], which is based on the Global Workspace Theory [21], and has been designed to be functionally (but not phenomenally) conscious. What mechanisms would need to be added to an implementation like LIDA in order to render her phenomenally conscious (i.e., in order to have her generating and reporting artificial qualia)? As pointed out by her creators [22], implementing a mechanism for the generation of a stable, coherent perceptual world, along the lines discussed by Merker [23, 24] (i.e., suppressing apparent motion produced by the movement of sensory receptors) might contribute to the design of phenomenally conscious machines.

Another significant work towards the definition of human-like information processing architectures is H-CogAff [25], where a meta-management layer based on reflective processes is added. This particular case of CogAff [26] is aimed at specifying a minimal architecture for a human-like system. In this work the issue of qualia is addressed as an architecture-based concept [16]. More specifically, virtual machine architectures are considered as a suitable domain for current MC analyses and experimentation. The virtual machine functionalism provides an account for phenomenology in which virtual machines, although not being physical, are real machines that can affect and be affected by the physical environment (see [16] for a thorough discussion on this point). Moving the discussion about qualia to the level of description of virtual machines – which is not definable in terms of physical sciences – might help clarifying what qualia actually are.

In relation with the problem of testing for the presence of artificial qualia, Gamez has proposed a scale that attempts to estimate the probability of a machine being able to sustain phenomenal states [27]. This scale makes no assumptions as to what specific mechanisms could be used to generate qualia. It is, instead, based on the similarity to human brain in terms of complexity.

More elaborated measures have been proposed to detect the presence of phenomenal states in neural systems. According to Tononi's theory of information integration [28], phenomenal consciousness can be detected looking at the level of dynamical complexity of a system, i.e. the combination of integration (unity of conscious experience) and differentiation (ability to discriminate conscious experiences amongst a vast repertoire of possible scenes). According to this, an artificial system would be able to experience qualia only if it is able to sustain a high level of information integration.

Apart to its theoretical application to biological organisms, Aleksander and Gamez have used Tononi's information integration theory to analyze how different artificial neural network topologies generate different effective information levels [29]. This work is focused in the discriminative or informational value of qualia. Although information integration can be considered a fundamental aspect in relation with qualia, other properties, like the ones discussed above, have to be taken into consideration [11].

4 CHARACTERIZING ARTIFICIAL QUALIA

In the domain of MC research, designers have to deal with the concept of qualia in order to develop implementations that could be claimed to be conscious. One way to mitigate the complexity involved in this task is to conceptually decompose the notion of qualia into different aspects that can be analyzed separately. It is out hypothesis that working with these partial views might provide some insight to the whole picture. What follows is a specific proposal of partial but complementary definitions of artificial qualia.

4.1 Partial definitions of artificial qualia

We propose to distinguish between three different stages or characterizations of the development of mechanisms that support qualia in machines (see Fig. 1 for an illustration of the relation between these stages; see Section 5 for an explanation of the content represented in Fig. 1):

- ***Stage 1. Perceptual Content Representation.*** At this stage the information acquired by the perceptual system of the machine is integrated and interpreted, generating a subjective representation. This content is built as a result of the combination of exteroceptive and proprioceptive sensing subsystems, hence giving place to an inherently subjective content representation. The process that generates this perceptual content involves a continuous checking for consistency. In other words, a number of possible partial reconstructions of the world compete for being integrated into the final consistent match. This match, or inner world final reconstruction, is achieved by a coherent integration between what is being perceived from the external world and what it currently represented as inner depiction (see Dennett's Multiple Draft Model [30], for a more metaphorical description of these sorts of competitive/collaborative content creation process).

- ***Stage 2. Introspective Perceptual Meta-Representation.*** This stage refers to the monitoring of the processes mentioned in stage 1, and also the creation of derived meta-representations. Observing how machine's own perceptual content is created can potentially result in a private ontology (meta-representation) about what is it like for the machine to experience subjective perceptual contents.

- ***Stage 3. Self-Modulation and Reportability.*** In the case of machine being able to achieve stages 1 and 2, meta-representations from stage 2, or introspective ontologies [2], could be used to modulate the way all perceptual systems work (including stages 1 and 2). This constitutes a self-regulation loop that has clear functional implications; i.e. qualia as defined here is part of a causal process. Additionally, introspective ontologies created in stage 2 could be used to report the artificial qualia of the machine.

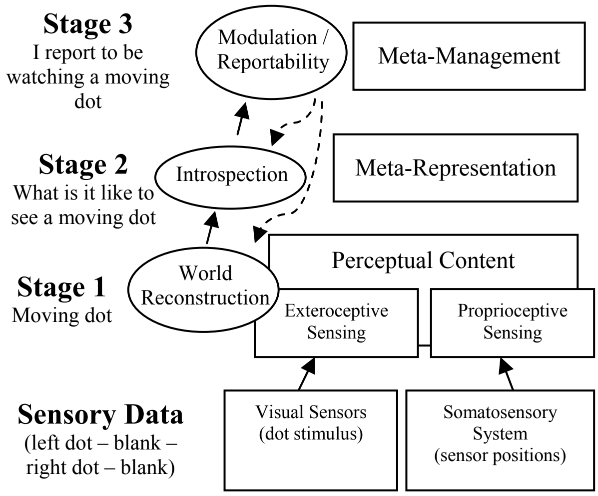

Figure 1. Stages in the development of artificial qualia

Stages described above make no claims about the qualities or modalities of the specific contents of the artificial mind, or to what extent they could resemble human subjective experience. The presence of different sensory modalities (laser ranging, for instance) and different mechanisms for cognition will produce different conscious contents and associated qualities. Nevertheless, we argue that the abstract mechanisms described above resemble those that generate qualia in biological organisms. We suggest that this claim could be supported by the application of the proposed framework in computational models, thus enabling a comparative analysis between qualia generated by humans and qualia generated in machines.

The proposed stages are just the components of a conceptual framework or guideline for the design of MC architectures. The subsequent implementations and experiments are expected to clarify some aspects of the nature of phenomenal experience and its impact in cognitive abilities.

Self-consciousness is not specifically addressed in the proposed definition because it is not considered a requirement for phenomenal consciousness. Nevertheless, self-consciousness could be explained in the context of the proposed framework having a model of the body as part of the perceptual content representation. The concept of self would be expected to arise as a meta-representation in stage 2. Then, references to the self could be found in accurate reports generated in stage 3.

4.2 Detecting the presence of qualia

The definition of qualia that we have described represents a hypothesis to be tested. To be precise, it should not be taken as a valid and established knowledge. That would lead us to think that the presence of qualia can be scientifically tested just by detecting the proposed mechanisms through inspection. In contrast, what we suggest is to use this approximation to what qualia might be as a working hypothesis. Thus approach calls for the experimentation with MC implementations designed following these assumptions.

The results of this testing process are expected to prove whether or not the original hypothesis was correct. One of the benefits of this kind of approach is that the firs-person problem mentioned above can be circumvented. Nevertheless, more work should be done in order to design meaningful experiments combining both behavioural outcome and architecture inspection (see for instance [31] for a particular discussion of misleading experiment design about self-consciousness using the mirror test). Also, identification of hallmarks of qualia, like bistable perception [32], should be put in the context of the present discussion.

5 VISUAL EXPERIENCE AND APPARENT MOTION

In order to illustrate the proposed segmented characterization of artificial qualia, the conscious perception of apparent motion is analyzed. Humans can perceive motion not only from real moving objects, but also from series of images containing spatially segregated visual stimuli [33]. Simple experiments to test this consist of two stationary blinking stimuli which are presented using different spatial and temporal parameters (see Fig. 2). At certain rates, subjects perceive motion (apparent motion).

Figure 2. Sequence of images used to generate apparent motion qualia in humans

The sequence of images depicted in Fig. 2 are used to generate apparent motion qualia in humans, and presumably also in phenomenally conscious machines. Note that a blank inter-stimulus interval (ISI) is inserted after every dot stimulus (the looping sequence is as follows: left dot stimulus – blank – right dot stimulus – blank).

This experiment, which is usually carried out with human subjects, could be carried out using a machine as subject. Putting this experiment in the context of the former characterization of qualia, let us consider a robot with a visual perceptual system modelled after the human visual cortex. The basic content at each stage could be described as:

- Stage 1: *"moving dot"*.
- Stage 2: *"what is it like to see a moving dot"*.
- Stage 3: *"I report to be watching a moving dot"*.

Fig. 1 represents these different levels of content in each of the stages that we have defined for the development of artificial qualia.

The perception process in the robot would follow these steps: first of all, visual sensors, let's say a stereo visual system composed of two digital cameras, acquires the images using their light detection sensors. At the same time, robot's somatosensory system acquires the relative position of the cameras, their orientation, and their foveation. The combination of sensory data from exteroceptive sensors (pixel maps from cameras) and proprioceptive sensors (relative coordinates from camera position and focus sensors) is then used to form depictions of percepts along the lines described by Aleksander and Dunmall [34].

As the sequence described in Fig. 2 is presented to the robot, single depictive percepts are created to represent the appearance of dots and their relative positions. Subsequently, robot motion detectors (this detectors could be built for instance with artificial neural networks), fed with the stream of dot percepts, will eventually create new motion percepts depending, amongst other things, on the duration of ISI frames. These moving dot percepts (or moving dot representations) are the contents of *Stage 1* ("moving dot").

The presence of the motion percepts will in turn trigger a set of reactions in the system. For instance, if the robot is designed to keep track of some moving objects, or detect some types of trajectories, the associated detectors will be activated. Also, affective evaluations of percepts could be invoked (the robot could be designed, or could have learnt, that moving dots have to be evaluated positively, and therefore maintain bonds with them). If the robot were endowed with a mechanism to represent these reactions, he would generate meta-representations of "what is like" for the robot to see a moving dot. This content corresponds to our *Stage 2* definition.

Finally, if Stage 2 introspective content is used both for self-regulating global perception-action processes, and also for reporting purposes, then the robot would be able to reason explicitly about what does it mean to him to see a moving dot. Provided with the necessary linguistic skills, the robot will also be able to report his mental content using his own ontology (*Stage 3* content).

6 CONCLUSIONS AND FUTURE WORK

A developmental view of qualia based on former work by other authors has been defined as an attempt to provide a conceptual framework for the creation of new MC models. The definition of qualia presented in this paper advocates for a functional role of phenomenal consciousness. Furthermore, self-modulation and integration of perceptual systems is appointed as a process being driven by the construction of introspective meta-representations. Additionally, reportability is assumed to be based on the same meta-representations.

It is expected that implementations based on the proposed conceptual framework are able to incorporate the phenomenal dimension of consciousness into their models. That is not to say that they will become phenomenally conscious just because the proposed conceptual definition is considered. Nevertheless, we argue that exploring the design space in the proposed direction might shed some light to the problem of production of qualia, both in natural and artificial creatures.

We are currently working in the development of visual illusion experiments, like the one described in the previous section, using the cognitive architecture CERA/CRANIUM [35]. We expect to draw some significant conclusions in the domain of synthetic phenomenology from these sorts of experiments.

Phenomenology is typically one of the fields where advancement in natural sciences is more difficult and we think this discipline could benefit more from the research with artificial systems. It is especially in more challenging aspect of natural sciences where we think that using AI-inspired biology can sometimes provide some useful clues to boost research.

ACKNOWLEDGEMENTS

This work was supported in part by the Spanish Ministry of Education under CICYT grant TRA2007-67374-C02-02.

REFERENCES

[1] A. Noë, "Is the Visual World a Grand Illusion?" *Journal of Consciousness Studies,* vol. 9, pp. 1-12, 2002.

[2] A. Sloman, "Why some machines may need qualia and how they can have them: Including a demanding new turing test for robot philosophers." in AI and Consciousness: Theoretical Foundations and Current Approaches - AAAI Fall Simposia 2007, pp. 9-16.

[3] R. Arrabales, A. Ledezma and A. Sanchis. Establishing a roadmap and metrics for conscious machines development. In Proceedings of the 8th IEEE International Confenrence on Cognitive Informatics. G. Baciu, Y. Wang, Y.Y. Yao, W. Kinsner, K. Chan & L.A. Zadeh (Eds.). Hong Kong. pp. 94-101. June 2009.

[4] R. Chrisley, "Synthetic Phenomenology," *International Journal of Machine Consciousness,* vol. 1, pp. 53-70, 2009.

[5] N. Block, "Some Concepts of Consciousness," 9, Oct. 20, 2001. 2001. Available: http://www.nyu.edu/gsas/dept/philo/faculty/block/

[6] N. Block, "On a Confusion about a Function of Consciousness," *Behav. Brain Sci.,* pp. 227-287, 1995.

[7] T. Nagel, "What Is It Like To Be a Bat?" *The Philosophical Review,* vol. 83, pp. 435-450, 1974.

[8] P. Lennie. "The physiology of color vision," in *The Science of Color (Second Edition)* Steven K. Shevell, Ed. 2003.

[9] P. O. A. Haikonen. "Qualia and conscious machines". *International Journal of Machine Consciousness 1(2),* pp. 225-234. 2009.

[10] D. C. Dennett. "Quining qualia," *Consciousness in Modern Science.* 1988. Available: http://cogprints.org/254/

[11] R. Arrabales, A. Ledezma and A. Sanchis, "Strategies for Measuring Machine Consciousness," *International Journal of Machine Consciousness,* vol. 1, pp. 193-201, 2009.

[12] D. Chalmers, "Facing Up to the Problem of Consciousness," *Journal of Consciousness Studies,* vol. 2, pp. 200-219, 1995.

[13] A. Seth, B. Baars and D. Edelman, "Criteria for consciousness in humans and other mammals," *Consciousness and Cognition,* vol. 14, pp. 119-139, 3. 2005.

[14] F. Jackson, "Epiphenomenal Qualia," *The Philosophical Quarterly,* pp. 127-136, 1982.

[15] L. S. Vygotsky. "*Mind in Society: The Development of Higher Psychological Processes." 1980.*

[16] A. Sloman and R. Chrisley "Virtual machines and consciousness". *Journal of Consciousness Studies, 10.* pp. 133-172. 2003.

[17] A. Chella and S. Gaglio, "In search of computational correlates of artificial qualia," in Advances in Intelligent Systems Research (AGI-09), May 2009.

[18] A. Chella, M. Frixione and S. Gaglio. "A cognitive architecture for robot self-consciousness," *Artif. Intell. Med. 44(2),* pp. 147-154. 2008.

[19] P. O. A. Haikonen, *Robot Brains. Circuits and Systems for Conscious Machines.* UK: John Wiley & Sons, 2007.

[20] S. Franklin, U. Ramamurthy, S. K. DiMello, L. McCauley, A. Negatu, R. L. Silva and V. Datla, "LIDA: A computational model of global workspace theory and developmental learning," in *AAAI Fall Symposium on AI and Consciousness: Theoretical Foundations and Current Approaches,* 2007, pp. 61-66.

[21] B. J. Baars, *A Cognitive Theory of Consciousness.* Cambridge: Cambridge University Press, 1988,

[22] U. Ramamurthy, "Might a LIDA controlled robot be phenomenally conscious?" in *Proceedings of the Nokia Workshop on Machine Consciousness,* 2008, pp. 32-33.

[23] B. Merker. The liabilities of mobility: A selection pressure for the transition to consciousness in animal evolution. *Conscious. Cogn. 14(1),* 2005. pp. 89-114.

[24] S. Franklin, "Evolutionary pressures and a stable world for animals and robots: A commentary on Merker," *Conscious. Cogn.,* vol. 14, pp. 115-118, 2005.

[25] A. Sloman and M. Scheutz. A framework for comparing agent architectures. Presented at In UKCI'02: Proceedings of the UK Workshop on Computational Intelligence. 2002.

[26] A. Sloman. Varieties of affect and the CogAff architecture schema. Presented at Proceedings Symposium on Emotion, Cognition, and Affective Computing AISB'01 Convention. 2001.

[27] D. Gamez, "An ordinal probability scale for synthetic phenomenology," in Proceedings of the AISB05 Symposium on Next Generation approaches to Machine Consciousness .2005, pp. 85-94.

[28] G. Tononi. "An information integration theory of consciousness." *BMC Neuroscience 5(1),* 2004. pp. 42.

[29] I. Aleksander and D. Gamez, "Iconic training and effective information: Evaluating meaning in discrete neural networks," in Biologically Inspired Cognitive Architectures II. Papers from the AAAI Fall Symposium. AAAI Technical Report FS-09-01 2009, pp. 2-10.

[30] D. C. Dennett, *Consciousness Explained.* Boston: Little, Brown and Co, 1991,

[31] P. O. A. Haikonen, "Reflections of consciousness: The mirror test," in *Proceedings of the 2007 AAAI Fall Symposium on Consciousness and Artificial Intelligence,* 2007, pp. 67-71.

[32] N. Fürstenau. "A computational model of bistable perception-attention dynamics with long range correlations". *KI 2007: Advances in Artificial Intelligence* pp. 251-263.

[33] L. Muckli, N. Kriegeskorte, H. Lanfermann, F. E. Zanella, W. Singer and R. Goebel. (2002, May 1). Apparent motion: Event-related functional magnetic resonance imaging of perceptual switches and states. *J. Neurosci. 22(9),* pp. 219.

[34] I. Aleksander and B. Dunmall, "Axioms and Tests for the Presence of Minimal Consciousness in Agents," *Journal of Consciousness Studies,* vol. 10, 2003.

[35] R. Arrabales, A. Ledezma and A. Sanchis, "CERA-CRANIUM: A test bed for machine consciousness research," in International Workshop on Machine Consciousness 2009. Hong Kong. June 2009. pp. 105.