



The Elephant in the Mirror: Bridging the Brain's Explanatory Gap of Consciousness

Jasmine A. Berry^{1*} and Alice C. Parker²

¹ Biomimetic Real-time Cortex Project, Computer Science Department, University of Southern California, Los Angeles, CA, USA, ² Biomimetic Real-time Cortex Project, Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, CA, USA

Keywords: self-recognition, self-awareness, consciousness, machine intelligence, hard problem

INTRODUCTION

The successes of the artificial retina and cochlea have lent encouragement to researchers in the general field of brain augmentation (Gantz et al., 2005; Dagnelie, 2012). However, in order for brain augmentation to progress beyond conventional sensory substitution to comprehensive augmentation of the human brain, we believe a better understanding of self-awareness and consciousness must be obtained, even if the “hard” problem of consciousness (Chalmers, 2007) remains elusive. Here we propose that forthcoming brain augmentation studies should insistently include investigations of its potential effects on self-awareness and consciousness. As a first step, it's imperative for comprehensive augmentation to include interfacing with the biological brain in a manner that either distinguishes self (biological brain) from other (augmentation circuitry), or incorporates both biological and electronic aspects into an integrated understanding of the meaning of self. This distinction poses not only psychological and physiological issues regarding the discrepancy of self and other, but raises ethical and philosophical issues when the brain augmentation is capable of introducing thoughts, emotions, memories and beliefs in such an integrated fashion that the wearer of such technology cannot distinguish his biological thoughts from thoughts introduced by the brain augmentation.

A consideration of self begins with the conventional mirror self-recognition test (MSR) (Gallop, 1970) that has been successfully executed with Eurasian magpies (Prior et al., 2008), bottlenose dolphins (Reiss and Marino, 2001), orca whales (Delfour and Marten, 2001), human infants typically between 18 and 24 months (Amsterdam, 1972; Rochat, 2003), and notably the Asian elephant (Plotnik et al., 2006). The only primate species reported to pass the Gallup Mirror Test, albeit controversially, were orangutans and chimpanzees (Suárez and Gallup, 1981). For years, MSR has been the designated litmus test for determining whether a species possesses self-awareness (SA), ultimately raising the question of whether the animal is then a conscious entity as a result of passing this test (De Veer and van den Bos, 1999). “Mirror self-recognition is an indicator of self-awareness,” proclaims Gallup et al. (2002). If indeed so, then the subsequent query to raise is whether self-awareness, the ability to differentiate oneself among others, is a precursor to or derivative of consciousness, and whether the mirror test is necessary and sufficient (Morin, 2011).

In light of brain research like the Blue Brain Project (Markram, 2006), BRAIN Initiative (Kandel et al., 2013), and development of neural prosthetics, the interest in consciousness is steadily growing. Here, we not only encourage the study of and suggest methods for addressing science's “elephant in the room,” which asserts consciousness is neither physical nor functional, but also place the *Elephas maximus* in our proverbial mirror to obtain a perspective toward forming a cohesive alliance between philosophical studies of consciousness and neural engineering's augmentative innovations. As MSR is purposed to grant the animal subject personal physical inspection from

OPEN ACCESS

Edited by:

Mikhail Lebedev,
Duke University, USA

Reviewed by:

Mazyar Fallah,
York University, Canada
Paul Manger,
University of the Witwatersrand,
South Africa
Alexander Fingelkurts,
BM-SCIENCE Brain & Mind
Technologies Research Centre,
Finland

*Correspondence:

Jasmine A. Berry
jasminab@usc.edu

Received: 08 April 2016

Accepted: 19 December 2016

Published: 06 January 2017

Citation:

Berry JA and Parker AC (2017) The
Elephant in the Mirror: Bridging the
Brain's Explanatory Gap of
Consciousness.
Front. Syst. Neurosci. 10:108.
doi: 10.3389/fnsys.2016.00108

an objective viewpoint, resulting in self-cognizance, so shall we take the approach to examine our modern scientific methods in conceptual mirrors, to appraise our consciousness dilemma and propose an assertion for progression in augmentative technologies. Following here is a succinct primer of consciousness and SA. We also issue a proposition as to how brain augmentation can influence the arrival of machine consciousness. Overall, we state our opinion for (1) why SA must be systematically examined in conjunction with brain augmentation approaches and (2) how such a merger could become a tool for investigating consciousness.

Ineffable Consciousness

The first pitfall encountered with consciousness is the inability to derive a functional explanation for what it means to *experience*. Chalmers (2007) lists the “easy” problems of consciousness as “the ability to discriminate, categorize, and react to environmental stimuli; the integration of information by a cognitive system; the reportability of mental states; the ability of a system to access its own internal states; the focus of attention; the deliberate control of behavior; the difference between wakefulness and sleep.” These phenomena are relatively feasible to exploit and can be described in computational model terms and neural operation derivations. Chalmers then counteracts them with the “hard” problem of lacking competency to explain why and how we have phenomenal experiences when being entertained by a movie, exhibiting a sensation toward classical music, or having feelings when watching a sunset. Explaining how the brain processes visual and auditory signals is trivial in comparison to how those same signals translate to *qualia*, subjective phenomenal experiences.

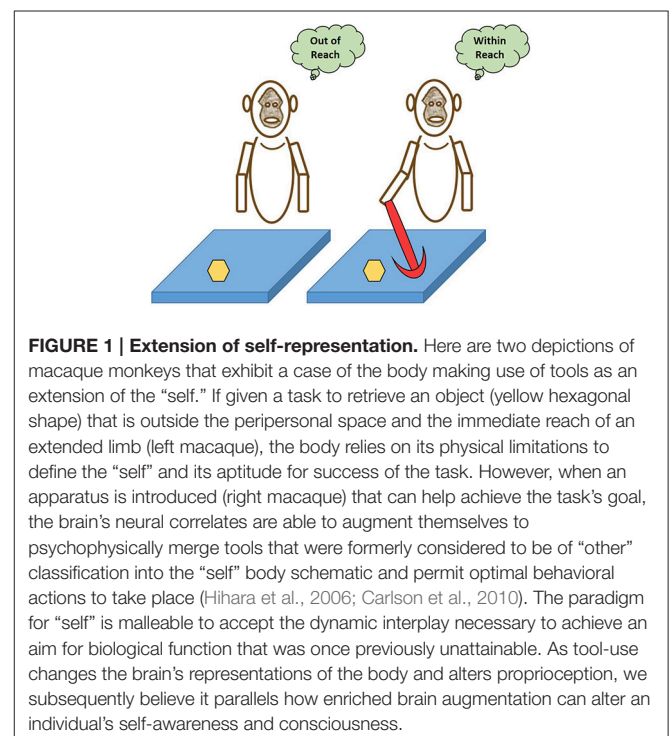
Explanatory Gap Dilemma

The term *explanatory gap*, coined by philosopher Joseph Levine (1983), notes our inability to connect physiological functions with psychological experience, thus creating the *gap*. Although Levine synonymizes consciousness with subjective feelings, the explanatory gap additionally alludes to reasoning, desires, memory, perception, beliefs, emotion, intentions, and human behavior/action. Correlating physical brain substrates to thoughts and feelings is the base of dispute between two parties: materialist reductionists and non-reductionists (Sawyer, 2002). Materialists’ chief view, representative of most neuroengineers, on the matter involves the belief that “when the brain shuts off, the mind shuts off” and the brain is the sole causative driver for consciousness. However, non-reductionists (typically philosophers) embrace a holism approach of mandating that the brain’s cortical components are insufficient in capturing consciousness, undertaking the possibility of supernatural properties. It’s an inquiry of necessity and sufficiency. The brain may be necessary for mental functions, but is it sufficient? Earlier analytical inspections on conscious experience have implied that an exclusive reductive justification is not satisfactory in delineating its emergence (Churchland, 1988; Kim, 2005; Clayton, 2006; Feinberg, 2012). A novel approach is needed to explain such experience. Our explanatory gap needs an explanatory bridge.

UNRAVELING SELF-AWARENESS TOWARD AUGMENTATION

Although many facets of consciousness are difficult to investigate, the development of objective tests for SA could be utilized for brain augmented technologies. With SA comes the sense of *agency*. Agency imparts a sense of who is the owner of an action/trait, the self, and who represents any entities excluding self, the other(s). Self-other dichotomy processing in the brain is essential to consciousness due to the necessary implications the embodiment of “self” must have to form body ownership. Once an agent gains the ability to discern when its own body is the source of sensory perceptions, it will form body awareness that entails proprioceptive information. We can look to working experiments that attempt to showcase how the brain augments the “self” when necessary to complete a task (Figure 1). Perceptual parametric information builds a premeditated awareness of (1) body part locations and (2) the manipulation of those same parts in space. Body awareness was demonstrated by a machine via Gold and Scassellati (2007) who built a robot named Nico that successfully distinguished its own “self” from “other.” Nico observably achieved self-recognition by completing mirror-aided tasks expending inverse kinematics. Nevertheless, it’s believed Nico lacked consciousness.

Before the sense of agency becomes fully refined through experiences over time, there must be a repertoire built for perceptions and actions. Whether, action and perception are interdependent or each fundamentally isolated has been the focus of another ongoing debate. It’s not yet concretely understood how the representation of self is formed during the initial stages of life. Either an agent first uses perception to motivate their



actions in the world or it first directs their actions to help drive perception of the sensory world or both occur simultaneously. In either method bodily awareness is eventually acquired which contributes to defining subjective cognitive attributes. Two opposing views attempt to solve this problem: the action-oriented theory of visual perception, which suggests that perception results from sensorimotor dynamics in an *acting* observer (Gibson, 1966; Noë, 2004; Mandik, 2005), and the dual-visual systems hypothesis, which advocates independent streams of perception and action (Schneider, 1969; Goodale and Milner, 1992; Jacob and Jeannerod, 2003; Milner and Goodale, 2006). Self-awareness uses expectation of impending perceptions and actions to gauge the assimilation of inner experience and external reality. Building a self-aware framework in augmentative technologies requires integration of an *expectancy intuition*, which is the capability to critique on the basis of differences between reality and internal experience. This is our tactic for creating systems with faculties for using perception and action to make predictions of self-sensory states, become self-adaptable to new environmental stimuli, and set objectives for self-improvement.

Crucial for understanding agency is determining how the embodied senses fuse to form self-referential experience (Fingelkurts et al., 2016a,b). It's our opinion that future advances of brain augmentation hinges on the application of such knowledge. Once we bridge this gap of the unknown we'll be challenged to use computational intelligence to create consciousness artificially and to integrate synthetic qualia with that produced in the brain. Presently, artificial devices can create various aspects of consciousness. Artificial perception is made available via cochlear, retinal, and tactile implants. But they simply work alone as replacements for sensory organs with consciousness and SA arriving later in the brain's neural processing. Applications for augmenting consciousness would contribute to studies relating to emotions, attention, supplementing memory capacity, personality alteration, experience enrichment, sensory perception enhancement, and hypernormal brain plasticity for self-repair.

PROPOSED TRANSITION TO MACHINE CONSCIOUSNESS

The marvel of human intelligence is its ability to eclipse physical limitations and overcome our biological constraints to form an ever-evolving existence (Jerison, 1973). One primary goal for reverse-engineering the human brain is to recreate the same functional mechanisms that underlie human consciousness in our software infrastructures, neurobotic agents, and computational systems. However, prosthetic memory, sensory implants, neurofeedback (EEG Biofeedback) and brain computer interfaces (BCIs) are all working examples of fusing such "intelligent" systems with the brain, leading to conceivable prospects for consciousness-altering devices. Although BCIs commonly target disability treatments and brain function recovery from lesion, the amalgamation of computational devices with the cortical brain itself (Fingelkurts et al., 2012) may even prompt increasing developments of an operational "exobrain" (Bonaci et al., 2014) for the purposes of better understanding

how our brain works. For example, in a scenario where a split-brain condition is present within a subject, we now have the option to look toward interfacing artificial exobrains with the cerebrum; such an interface can either serve as a replacement for neurological issues or supplement features the brain does not naturally comprise. If these exobrains have a modicum of manipulability, then we can explore the plausibility of mind transfer from device to organ and vice versa; thus, providing speculation for a conscious machine that can affect how we can perceive, act, express emotion, feel, and adapt. This poses ethical concerns as it opens the door for alterations of an individual's SA when augmentation is capable of modifying reasoning skills and subjective judgment. Successful augmentation of the sort might render the individual powerless in discriminating actual characteristics and thoughts from those that are mock and introduced artificially outside the cortex. Combining the precision and information processing speed of a computer with the intrinsic non-computational attributes of a human may provoke discoveries of the mind (e.g., consciousness) that we as humans are currently incapable of resolving. We suggest efforts made toward an augmentative interface between brain and machine that prompts the human mind to think beyond its unknown limits for the construction of our explanatory bridge.

CHALLENGES MOVING FORWARD

Many people view an in-depth exploration into consciousness and its emergence as a gamble, considering decades already spent on the matter with a void of consensus (Dennett, 1991; Jibu and Yasue, 1995; Hameroff and Penrose, 1996; Stapp, 2001; Crick and Koch, 2003; Tononi, 2008; De Sousa, 2013; Seager, 2016). Before we attempt to create another hypothesis, our approach needs to change; it's our suggestion to further refine the constructs and emergence of SA and to use brain augmentation as an instrument for inspection. We need to define an objective test for determining whether an entity is a sentient being. This test in addition to advances in neural engineering provide optimism that disputes within the consciousness field can be resolved. Augmentation has a promising future as an enhancement to our brains and will hopefully influence our centuries-old methods of thinking about consciousness toward an answer for science's greatest mystery.

AUTHOR CONTRIBUTIONS

JB: This author was responsible for providing the overall opinion of the topic and crafting the outline of the article. JB contributed most of the text and references supplied in the document. JB also primarily conducted the necessary research to develop this article. JB supplied the latter sections of the text with original ideas to add to the discussion of the Brain Augmentation research topic. AP: This author assisted in crafting the thesis of the article and provided additional topic of discussion to implement within the text. AP provided introductory material for the article and also made revisions to each section to ensure a coherent message about our opinion was made. AP also restructured the format of the article to enhance comprehension to an audience with varying backgrounds.

REFERENCES

- Amsterdam, B. (1972). Mirror self-image reactions before age two. *Dev. Psychobiol.* 5, 297–305. doi: 10.1002/dev.420050403
- Bonaci, T., Herron, J., Matlack, C., and Chizeck, H. J. (2014). “Securing the exocortex: a twenty-first century cybernetics challenge,” in *Norbert Wiener in the 21st Century (21CW)* (Boston, MA), 1–8.
- Carlson, T. A., Alvarez, G., Wu, D. A., and Verstraten, F. A. (2010). Rapid assimilation of external objects into the body schema. *Psychol. Sci.* 21, 1000–1005. doi: 10.1177/0956797610371962
- Chalmers, D. (2007). “The Hard Problem of Consciousness,” in *The Blackwell Companion to Consciousness*, ed M. Velmans and S. Schneider (Malden, MA: Blackwell Publishing), 225–235.
- Churchland, P. S. (1988). “Reduction and the neurobiological basis of consciousness,” in *Consciousness in Contemporary Science*, ed A. J. Marcel and E. Bisiach (New York, NY: Oxford University Press), 273–304.
- Clayton, P. (2006). *Mind and Emergence: From Quantum to Consciousness*. New York, NY: Oxford University Press.
- Crick, F., and Koch, C. (2003). A framework for consciousness. *Nat. Neurosci.* 6, 119–126. doi: 10.1038/nn0203-119
- Dagnelie, G. (2012). Retinal implants: emergence of a multidisciplinary field. *Curr. Opin. Neurol.* 25, 67–75. doi: 10.1097/WCO.0b013e32834f02c3
- Delfour, F., and Marten, K. (2001). Mirror image processing in three marine mammal species: killer whales (*Orcinus orca*), false killer whales (*Pseudorca crassidens*) and California sea lions (*Zalophus californianus*). *Behav. Processes* 53, 181–190. doi: 10.1016/S0376-6357(01)00134-6
- Dennett, D. C. (1991). *Consciousness Explained*. Boston, MA: Little, Brown and Company.
- De Sousa, A. (2013). Towards an integrative theory of consciousness: part 1 (neurobiological and cognitive models). *Mens Sana Monogr.* 11, 100–150. doi: 10.4103/0973-1229.109335
- De Veer, M. W., and van den Bos, R. (1999). A critical review of methodology and interpretation of mirror self-recognition research in nonhuman primates. *Anim. Behav.* 58, 459–468. doi: 10.1006/anbe.1999.1166
- Feinberg, T. E. (2012). Neuroontology, neurobiological naturalism, and consciousness: a challenge to scientific reduction and a solution. *Phys. Life Rev.* 9, 13–34. doi: 10.1016/j.plrev.2011.10.019
- Fingelkurts, A. A., Fingelkurts, A. A., and Kallio-Tamminen, T. (2016a). Long-term meditation training induced changes in the operational synchrony of default mode network modules during a resting state. *Cogn. Processes* 17, 27–37. doi: 10.1007/s10339-015-0743-4
- Fingelkurts, A. A., Fingelkurts, A. A., and Kallio-Tamminen, T. (2016b). Trait lasting alteration of the brain default mode network in experienced meditators and the experiential selfhood. *Self Identity* 15, 381–393. doi: 10.1080/15298868.2015.1136351
- Fingelkurts, A. A., Fingelkurts, A. A., and Neves, C. F. (2012). “Machine” consciousness and “artificial” thought: an operational architectonics model guided approach. *Brain Res.* 1428, 80–92. doi: 10.1016/j.brainres.2010.11.079
- Gallo, G. G. Jr. (1970). Chimpanzees: self recognition. *Science* 167, 86–87. doi: 10.1126/science.167.3914.86
- Gallup, G. G. Jr., Anderson, J. L., and Shillito, D. P. (2002). “The mirror test,” in *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition*, eds M. Bekoff, C. Allen, and G. M. Burghardt (Cambridge, MA: MIT Press), 325–334.
- Gantz, B. J., Turner, C., Gfeller, K. E., and Lowder, M. W. (2005). Preservation of hearing in cochlear implant surgery: advantages of combined electrical and acoustical speech processing. *Laryngoscope* 115, 796–802. doi: 10.1097/01.MLG.0000157695.07536.D2
- Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems*. Boston, MA: Houghton Mifflin.
- Gold, K., and Scassellati, B. (2007). “A Bayesian robot that distinguishes “self” from “other,”” in *Proceedings of the 29th Annual Meeting of the Cognitive Science Society* (Nashville, TN), 1037–1042.
- Goodale, M. A., and Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends Neurosci.* 15, 20–25. doi: 10.1016/0166-2236(92)90344-8
- Hameroff, S., and Penrose, R. (1996). Orchestrated reduction of quantum coherence in brain microtubules: a model for consciousness. *Math. Comput. Simul.* 40, 453–480. doi: 10.1016/0378-4754(96)80476-9
- Hihara, S., Notoya, T., Tanaka, M., Ichinose, S., Ojima, H., Obayashi, S., et al. (2006). Extension of corticocortical afferents into the anterior bank of the intraparietal sulcus by tool-use training in adult monkeys. *Neuropsychologia* 44, 2636–2646. doi: 10.1016/j.neuropsychologia.2005.11.020
- Jacob, P., and Jeannerod, M. (2003). *Ways of Seeing: The Scope and Limits of Visual Cognition*. Oxford: Oxford University Press.
- Jerison, H. J. (1973). *Evolution of the Brain and Intelligence*. New York, NY: Academic Press.
- Jibu, M., and Yasue, K. (1995). *Quantum Brain Dynamics and Consciousness*. Amsterdam; Philadelphia, PA: John Benjamins Publishing.
- Kandel, E. R., Markram, H., Matthews, P. M., Yuste, R., and Koch, C. (2013). Neuroscience thinks big (and collaboratively). *Nat. Rev. Neurosci.* 14, 659–664. doi: 10.1038/nrn3578
- Kim, J. (2005). *Physicalism, or Something Near Enough*. Princeton, NJ: Princeton University Press.
- Levine, J. (1983). Materialism and qualia: the explanatory gap. *Pac. Philos. Q.* 64, 354–361.
- Mandik, P. (2005). “Action-Oriented Representation,” in *Cognition and the Brain: The Philosophy and Neuroscience Movement*, eds K. Akins and A. Brook (Cambridge: Cambridge University Press), 284–305.
- Markram, H. (2006). The blue brain project. *Nat. Rev. Neurosci.* 7, 153–160. doi: 10.1038/nrn1848
- Milner, A. D., and Goodale, M. A. (2006). *The Visual Brain in Action*. Oxford: Oxford University Press.
- Morin, A. (2011). Self-recognition, theory-of-mind, and self-awareness: what side are you on? *Laterality* 16, 367–383. doi: 10.1080/13576501003702648
- Noë, A. (2004). *Action in Perception*. Cambridge, MA: MIT press.
- Plotnik, J. M., de Waal, F. B. M., and Reiss, D. (2006). Self-recognition in an Asian elephant. *Proc. Natl. Acad. Sci. U.S.A.* 103, 17053–17057. doi: 10.1073/pnas.0608062103
- Prior, H., Schwarz, A., and Gunturkun, O. (2008). Mirror-induced behavior in the magpie (*pica pica*): evidence of self-recognition. *PLoS Biol.* 6:e202. doi: 10.1371/journal.pbio.0060202
- Reiss, D., and Marino, L. (2001). Mirror self-recognition in the bottlenose dolphin: a case of cognitive convergence. *Proc. Natl. Acad. Sci. U.S.A.* 98, 5937–5942. doi: 10.1073/pnas.101086398
- Rochat, P. (2003). Five levels of self-recognition as they unfold early in life. *Conscious. Cogn.* 12, 717–773. doi: 10.1016/S1053-8100(03)00081-3
- Sawyer, R. K. (2002). Emergence in psychology: lessons from the history of non-reductionist science. *Hum. Dev.* 45, 2–28. doi: 10.1159/000048148
- Schneider, G. E. (1969). Two visual systems: brain mechanisms for localization and discrimination are dissociated by tectal and cortical lesions. *Science* 163, 895–902. doi: 10.1126/science.163.3870.895
- Seager, W. (2016). *Theories of Consciousness: An Introduction and Assessment*. New York, NY: Routledge.
- Stapp, H. P. (2001). Quantum theory and the role of mind in nature. *Found. Phys.* 31, 1465–1499. doi: 10.1023/A:1012682413597
- Suárez, S. D., and Gallup, G. G. (1981). Self-recognition in chimpanzees and orangutans, but not gorillas. *J. Hum. Evol.* 10, 175–188. doi: 10.1016/S0047-2484(81)80016-4
- Tononi, G. (2008). Consciousness as integrated information: a provisional manifesto. *Biol. Bull.* 215, 216–242. doi: 10.2307/25470707

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Berry and Parker. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.