Can Being Aware of the Illusion of Self Augment an Agent's Affordances: Integrating Buddhist Philosophy, Cognitive Science, and Artificial Life

Thomas Doctor[1,2], Elizaveta Solomonova[1,3], Bill Duane[1,2,4], and Olaf Witkowski[1,5,6,7]

[1]Center for the Study of Apparent Selves, Kathmandu, Nepal
[2]Centre for Buddhist Studies at Rangjung Yeshe Institute, Kathmandu University, Kathmandu, Nepal
[3]Department of Psychiatry, McGill University, Montreal, Quebec, Canada
[4]Bill Duane and Associates
[5]Cross Labs, Cross Compass Ltd., Japan
[6]Earth-Life Science Institute, Tokyo Institute of Technology, Japan
[7]College of Arts and Sciences, University of Tokyo, Japan
info@csas.ai

Objectives

Developments in cognitive science, AI, and artificial life force us to consider minds and intelligences that are different from human minds. The dominant contemporary metaphor for any kind of mind is based on an understanding of the human brain and human experience, both of which frequently presuppose a notion of self. In some disciplines, including Buddhism, contemporary philosophy of mind, and cognitive science, much debate has focused on the nature of the self, and one insight from all these domains is that while we are strongly attached to notions of stable selves, it is also possible to conceive of selves as dynamic, interconnected, and illusory. We suggest that an interdisciplinary approach, drawing on fields with well-developed models of self in relation to agency, can offer new insights. We suggest that the view of self as illusory, and awareness of this illusion, in both human and non-human minds, may augment and qualitatively change the agent's affordances, or range of possible actions.

The illusion of self in Buddhism, philosophy and contemplative science

Our approach is based on Buddhist views of "no self" (Garfield, 2014; Kapstein, 2016), which in turn can be seen as developing the idea that what is conventionally considered to be a single and enduring person (or agent) is constructed upon a plurality of impermanent causal factors. Buddhist philosophers may thus suggest that because no person or agent can be found apart from these constantly changing factors, it is more profitable to consider a person an imputation rather than a substantial and enduring individual (Kapstein, 2016). According to Buddhist traditions, and as a basis for contemplative practices, this view has ethical implications: it is thought that an individual who recognizes oneself as an entity that is constantly co-constituted by their interactions with others, will develop prosocial qualities. Mahāyāna Buddhist thought describes this process as an orientation towards others as part of one's own identity (Maitreya-Asaṅga et al., 2014). In other words, the self can in this way be considered both extendable and illusory. In Buddhist thought, the notion of "self as an illusion" can then be summarized as a rejection of singular and enduring individuals (Gold, 2021; Kapstein, 2016), which in turn may enable perceptions of selfhood as distributed/distributable and dynamic (Maitreya-Asaṅga et al., 2014).

Cognitive science: contemplative practice and lucid dreaming alter affordances

Indeed, this notion of diminishing the importance and stability of a sense of self, while counterintuitive but intriguing, has gained traction in contemporary contemplative practices as well in cognitive neuroscience (Antonova & Nehaniv, 2018; Britton et al., 2013; Wallace, 2009). Many meditation practices that involve disciplined and deliberate activities of focused observation of one's own thoughts and sensations are aimed at testing, experientially, the idea that what may otherwise typically be referred to as an "I" cannot be found as any specific entity. Research shows that many practices aimed at decreasing the process of "selfing" (a process of reifying one's own identity as an independent self), result in beneficial effects, such as increased well-being (Britton et al, 2013), and in increases in social connectedness (Hutcherson, Seppala, & Gross, 2008) as well as prosocial emotions and behaviours (Condon, 2019; Luberto et al., 2018).

In addition to mindfulness practices, the lived experience of the illusion of self can sometimes be

triggered by an experience of lucidity during dreams. Lucid dreams are dreams in which the dreamer is aware of the fact that they are dreaming (LaBerge & Rheingold, 1991). Following this realization, they may gain control over some of the elements of the dream, in particular over the way their body may move (e.g. flying) and over aspects of the environment (Stumbrys, Erlacher, Johnson, & Schredl, 2014). In Tibetan Buddhist tradition, cultivation of dream lucidity is an integral part of a contemplative path (Holecek, 2016; Wallace & Hodel, 2012), and it is suggested that through the gradual realization of the illusory nature of one's interactions with the dreamworld (which is a very compelling simulation of the real world), one may, in turn, realize the illusory nature of one's self.

Both contemplative practices and lucid dreaming, have been linked to general increase in the sense of well-being and insight (Dahl & Davidson, 2019, Konkoly & Burke, 2019), and are characterized by potential change in affordances. The concept of affordances (Gibson, 1977) refers to the "relationships between the abilities of organisms and features of the environment" (Chemero, 2003). Recent work has identified multiple facets of affordances, including social affordances (Valenti & Gold, 1991). For example, experience in contemplative practice may lead to change in the way in which one interacts with others and the world, thus changing social affordances; and lucid dreams open up possibilities for action which are not normally possible in waking life (Tholey, 1989) thus altering environmental (dream) affordances. While these practices are both private experiences of self as relatively illusory, their documented effects lead to potential questions of how and whether these insights can be applied to other kinds of minds, in particular to AI.

Integrating with tools from AI and artificial life

While AI is interested in replicating human intelligence, the field of artificial life more largely studies toy models of living and cognitive systems in order to understand them (Bedau et al., 2000; Bedau, 2003). The main distinction with the disciplines mentioned above, is that artificial life will approach the illusion of self not directly, but by replicating its effect within a model. Through rigorous modeling assumptions, one may simulate the mechanisms to test hypotheses without interacting with the original object of study. As long as each step can be stated clearly, such a model has the power to disentangle difficult questions in the original domain of interrogation.

This paper attempts to address, in the light of AI and artificial life models, the question of the self and, more specifically, the nature of the effects that may follow from an agent understanding self as an illusion. In order to break down the question of *whether awareness of the illusion of self augments an agent's affordances* into its composing factors, we naturally need an actionable understanding of the notions of self, individuality, identity, agency, affordances, enaction, empowerment, control, integration, as well as awareness, and illusion. A wide amount of literature exists on each of these concepts, and we don't intend here to perform an extensive review of them, but rather identify the existence of a minimal set of concepts that will allow us to address the question at hand. One may start by framing a self in a model, which connects to the notion of an individual (Barandiaran, 2003; Krakauer, 2020), while also requiring tools to capture the identity of an autonomous agent interacting with its environment via sensorimotor loops (Varela, 1997; Beer, 2014). Information theory offers a language to characterize such entities and synergies between them, using measures such as predictive information (Bialek, 2001; Ay, 2008), synergistic information (Edlund et al., 2011), multi-information (McGill, 1954; Schneidman, 2003), and integration (Tononi et al., 1994; Lungarella et al., 2005). Of course, the meaning of information is modulated by its use (Wittgenstein, 2010; Donaldson-Matasci et al., 2010; Kolchinsky & Wolpert, 2018), and such characterization may therefore benefit from being related to a function such as empowerment, or the quantification of options available and visible to an agent (Klyubin et al., 2008), or characterization of the change in affordances (Borghi et al., 2011). To capture the remaining mechanisms relating to the sense of self and of agency (Haggard, 2017), such tools should then be carefully connected to a formal, relevant mathematical model.

Hypothesizing that an artificial model can provide an integrated framework to address the question at hand, the information and modeling tools we proposed above may profit from being connected to frameworks from cognitive science and philosophy addressing the sense of agency as a private experience. The challenge, then, is to extract meaningful data from the study of human cognition, and integrate them. This effort will require a dual process of using what we know about the sense of self, and testing whether and how this knowledge may affect experience and behavior in human and other kinds of minds. We propose that the notion of the self as illusion, as realized in AI and artificial life practices, may contribute to novel fields of affordances for non-human intelligence. This framework may contribute not only to testing theories of affordances, but also shed light on new fundamental principles in AI science, possibly leading to new kinds of technologies.

Acknowledgment

This project is made possible through the support of a grant from Templeton World Charity Foundation, Inc. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of Templeton World Charity Foundation, Inc.

References

Antonova, E., & Nehaniv, C. L. (2018). Towards the Mind of a Humanoid: Does a Cognitive Robot Need a Self? - Lessons from Neuroscience. In: Takashi Ikegami; Nathaniel Virgo; Olaf Witkowski; Mizuki Oka; Reiji Suzuki; Hiroyuki Iizuki (ed.) *Proceedings of the Artificial Life Conference 2018 (ALIFE 2018 - Tokyo, Japan)*. MIT Press, 2018.

Ay, N., Bertschinger, N., Der, R., Güttler, F., & Olbrich, E. (2008). Predictive information and explorative behavior of autonomous robots. *The European Physical Journal* B, 63(3), 329-339.

Barandiaran, X. E., Di Paolo, E., & Rohde, M. (2009). Defining agency: Individuality, normativity, asymmetry, and spatio-temporality in action. *Adaptive Behavior*, 17(5), 367-386.

Bedau, M. A. (2003). Artificial life: organization, adaptation and complexity from the bottom up. *Trends in Cognitive Sciences*, 7(11), 505-512.

Bedau, M. A., McCaskill, J. S., Packard, N. H., Rasmussen, S., Adami, C., Green, D. G., Ikegami, T., Kaneko, K., & Ray, T. S. (2000). Open problems in artificial life. *Artificial Life*, 6(4), 363-376.

Beer, R. D. (2014). The cognitive domain of a glider in the game of life. *Artificial Life*, 20(2), 183-206.

Borghi, A. M., Di Ferdinando, A., & Parisi, D. (2011). Objects, spatial compatibility, and affordances: A connectionist study. *Cognitive Systems Research*, 12(1), 33-44.

Britton, W. B., Brown, A. C., Kaplan, C. T., Goldman, R. E., DeLuca, M., Rojiani, R., ... Frank, T. (2013). Contemplative Science: An Insider Prospectus. In Svinicki, M (ed). *New Directions for Teaching and Learning: Contemplative Studies in Higher Education.* San Francisco: Jossey-Bass.

Condon, P. (2019). Meditation in context: factors that facilitate prosocial behavior. *Current Opinion in Psychology*, 28, 15-19. doi:10.1016/j.copsyc.2018.09.011

Chemero, A. (2003). An outline of a theory of affordances. *Ecological Psychology*, *15*(2), 181-195.

Dahl, C. J., & Davidson, R. J. (2019). Mindfulness and the contemplative life: pathways to connection, insight, and purpose. *Current Opinion in Psychology*, 28, 60-64.

Donaldson‐Matasci, M. C., Bergstrom, C. T., & Lachmann, M. (2010). The fitness value of information. *Oikos*, 119(2), 219-230.

Edlund, J., N. Chaumont, A. Hintze, C. Koch, G. Tononi, et al. (2011). Integrated information increases with fitness in the evolution of animats. *PLOS Computational Biology*, 7: e1002236.

Garfield, Jay. (2014) *Engaging Buddhism: Why It Matters to Philosophy*. New York: Oxford University Press.

Gershenson, C. (2018). Information in science and Buddhist philosophy: Towards a non-materialistic worldview. Preprint: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3261381

Gibson, J. J. (1977). The theory of affordances. In Gibson J. J (ed.) *The Ecological Approach to Visual Perception*. Hillsdale: Lawrence Erlbaum Associates.

Gold, J. C. (2021). Vasubandhu. *The Stanford Encyclopedia of Philosophy (Spring 2021 Edition)*, Zalta, E. N. (ed.), forthcoming URL: https://plato.stanford.edu/archives/spr2021/entries/vasubandhu/

Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, 18(4), 196.

Holecek, A. (2016). *Dream Yoga: Illuminating Your Life Through Lucid Dreaming and Tibetan Yogas of Sleep*. Louisville: Sounds True.

Hutcherson, C. A., Seppala, E. M., & Gross, J. J. (2008). Loving-kindness meditation increases social connectedness. *Emotion*, 8(5), 720-724. doi:10.1037/a0013237

Kapstein, M. (2016). *Self, Non-Self, and Personal Identity*. In Oxford Bibliographies. doi:DOI: 10.1093/OBO/9780195393521-0147

Klyubin, A. S., Polani, D., & Nehaniv, C. L. (2008). Keep your options open: An information-based driving principle for sensorimotor systems. *PloS ONE*, 3(12), e4018.

Kolchinsky, A., & Wolpert, D. H. (2018). Semantic information, autonomous agency and non-equilibrium statistical physics. *Interface Focus*, 8(6), 20180041.

Konkoly, K., & Burke, C. T. (2019). Can learning to lucid dream promote personal growth?. *Dreaming*, 29(2), 113.

Krakauer, D., Bertschinger, N., Olbrich, E., Flack, J. C., & Ay, N. (2020). The information theory of individuality. *Theory in Biosciences*, 139(2), 209-223.

LaBerge, S., & Rheingold, H. (1991). *Exploring the World of Lucid Dreaming*. New York: Ballantine Books.

Luberto, C. M., Shinday, N., Song, R., Philpotts, L. L., Park, E. R., Fricchione, G. L., & Yeh, G. Y. (2018). A systematic review and meta-analysis of the effects of meditation on empathy, compassion, and prosocial behaviors. *Mindfulness*, 9(3), 708-724.

Lungarella, M., Pegors, T., Bulwinkle, D., & Sporns, O. (2005). Methods for quantifying the informational

structure of sensory and motor data. *Neuroinform*, 3: 243-262.

Lutz, A., Slagter, H. A., Dunne, J. D., & Davidson, R. J. (2008). Attention regulation and monitoring in meditation. *Trends in Cognitive Sciences*, 12(4), 163-169.

Maitreya-Asaṅga, Shenga, & Mipham. (2014). *Ornament of the Great Vehicle Sūtras: Maitreya's Mahāyānasūtrālaṃkāra with Commentaries by Khenpo Shenga and Ju Mipham*. Boston: Shambhala.

McGill, W.J. 1954. Multivariate information transmission. *Psychometrika* 19: 97-116.

Moreno, A., & Etxeberria, A. (2005). Agency in natural and artificial systems. *Artificial Life*, 11(1-2), 161-175.

Oizumi, M., Albantakis, L., & Tononi, G. (2014). From the phenomenology to the mechanisms of consciousness: integrated information theory 3.0. *PLoS Computational Biology*, 10(5), e1003588.

Pacherie, E. (2007). The sense of control and the sense of agency. *Psyche*, 13(1), 1-30.

Schneidman, E., S. Still, M.J. Berry & W. Bialek. 2003. Network information and connected correlations. *Phys. Rev. Lett.*, 91: 238701.

Stumbrys, T., Erlacher, D., Johnson, M., & Schredl, M. (2014). The phenomenology of lucid dreaming: an online survey. *American Journal of Psychology*, 127(2), 191-204. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24934010

Tholey, P. (1989). Consciousness and abilities of dream characters observed during lucid dreaming. *Perceptual and Motor Skills,* 68(2), 567-578.

Tononi, G., O. Sporns & G.M. Edelman. 1994. A measure for brain complexity: Relating functional segregation and integration in the nervous system. *Proc. Natl. Acad. Sci. USA* 91: 5033-5037.

Varela, F. J. (1997). Patterns of life: Intertwining identity and cognition. *Brain and Cognition*, 34(1), 72-87.

Valenti, S. S., & Gold, J. M. (1991). Social affordances and interaction I: Introduction. *Ecological Psychology*, *3*(2), 77-98.

Wallace, A. (2009). *Contemplative Science: Where Buddhism and Neuroscience Converge.* New York: Columbia U. Press.

Wallace, B. A., & Hodel, B. (2012). *Dreaming Yourself Awake Lucid Dreaming and Tibetan Dream Yoga for Insight and Transformation.* Boston: Shambhala Publications.

Wittgenstein, L. (2010). *Philosophical Investigations*. Chichester: John Wiley & Sons.