The Major Transitions in Cognitive Evolution

PSA 2020 Symposium Proposal

Participants and Affiliations

- Andrew Barron (andrew.barron@mq.edu.au). Professor of Neuroethology, Department of Biological Sciences, Macquarie University; Australian Research Council Future Fellow.
- Rachael Brown (rachael.brown@anu.edu.au). Lecturer, School of Philosophy; Director, Centre for Philosophy of the Sciences, The Australian National University.
- Marta Halina (mh801@cam.ac.uk). University Lecturer, Department of History and Philosophy of Science, University of Cambridge; Director, Kinds of Intelligence Program, Leverhulme Centre for the Future of Intelligence.
- Colin Klein (colin.klein@anu.edu.au). Associate Professor, School of Philosophy; Foundations Lead, Humanising Machine Intelligence Project, The Australian National University.

Descriptive Summary

As we come to appreciate the wealth and diversity of intelligences, the challenge becomes how to make sense of this complexity. How can we comprehend it in a systematic way? In this symposium, we explore the idea that one important piece of the puzzle involves treating the evolution of cognition as a series of major transitions. Each transition involved a qualitative change in information flow within nervous systems. Each transition opened up new cognitive capacities, while transforming the power and scope of existing cognitive functions. We consider several aspects of the major transitions framework: the role of learning, brain size, and resource constraints on the evolution of nervous systems; the relationship between major transitions and behavioral innovation; and the implications of major transitions on methods in psychology. This symposium contributes to contemporary debates in general philosophy of science, philosophy of biology and philosophy of cognitive science while drawing on work in evolutionary biology, computational neuroscience and comparative psychology.

Description and Justification

We explore the idea that the evolution of different kinds of cognitive abilities can be modelled in part in terms of major transitions in cognitive evolution. Each transition represents a qualitative shift in how information flows within an organism's nervous system. These shifts had profound consequences for what was computationally feasible for organisms with the new capacities for information flow. Transitions define types of systems that differ in both cognitive potential and evolvability. Recognizing major transitions in cognitive evolution will help us comprehend the diversity of biological intelligences, as well as give us a potential handle on other types and sources of intelligence.

The term 'major transitions' is borrowed from Smith and Szathmáry (1995), who introduced the idea into evolutionary biology. Smith and Szathmáry identified several major evolutionary transitions—such as the origins of eukaryotes and multicellular organisms—which made possible radically different forms of life. We hypothesize that the evolution of cognition can be modelled in a similar way. In this symposium, we evaluate this hypothesis, considering constraints on cognitive evolution and how and why major transitions might occur. We also examine the implications of the major cognitive transitions framework on traditional topics in philosophy of science, such as measurement, explanation, and parsimony.

Philosophers of biology have long grappled with the question of why major evolutionary transitions occur, leading to the development of several new frameworks for understanding biological innovation and evolvability (Godfrey-Smith, 2006; Trestman, 2013). The transitions framework raises a number of important philosophical questions at the interface between philosophy of biology and cognitive science.

Barron will examine the significance of brain size and major transitions. It is widely recognised that brain size does not correlate in a straightforward way with cognitive abilities. Work in computational neuroscience, artificial intelligence and comparative psychology also demonstrates impressive computational feats by small neural networks. Barron will show how a better understanding of the energy and computational tradeoffs of high neuron number provides a picture of the constraints on major cognitive transitions and when we should expect such transitions to map onto brain size. Klein will further explore the effects of resource constraints (such as time, space and energy) on brain evolution, while advocating for the importance of resource explanations in the philosophy of neuroscience.

Brown will evaluate the claim that major transitions transform the behavioral innovativeness of species. She will do this by advancing a multidimensional framework for measuring behavioral innovation and considering the implications of this framework on our understanding of the evolution and distribution of behavioral innovation. Halina will focus on the recent proposal by Eva Jablonka, Simona Ginsburg and colleagues that unlimited associative learning constitutes a major cognitive transition. She will draw on this thesis to inform debates in philosophy of science on null models, arguing that the features of unlimited associative learning qua major transition make it a compelling candidate for a null model in comparative psychology.

This symposium will bring together early career researchers and established figures working at the intersection of philosophy and the evolution of cognition. Andrew Barron is a world-renowned neuroethologist, specializing in insect cognition. His research is highly transdisciplinary and integrative with projects spanning neuroscience, cognitive science, ethology, computational biology and philosophy. Rachael Brown is a distinguished early career philosopher of biology working on evolvability, behavioral novelty and animal minds. Colin Klein is a world-leading expert in philosophy of cognitive science. His collaborative, cross-disciplinary research has been featured in global media such as the New York Times and Wired. Marta Halina works at the intersection of philosophy, biology and comparative psychology and has led high-impact research initiatives such as the Kinds of Intelligence Program at the University of Cambridge and the Animal-AI Olympics.

Paper Titles and Abstracts

Cognition with Large and Small Brains
Andrew Barron

An assumption that a bigger brain is better is extremely pervasive in animal behaviour studies and comparative neuroscience, but presently we have no clear understanding of either the cognitive benefits conferred by a large brain, or the limits imposed by a small brain (Logan et al., 2018). Neuronal tissue is metabolically expensive to both build and operate, and there are many examples of the evolution of reduced neural tissue as a result of selective pressures to minimise the associated costs. From this perspective it is presumed a large brain must have evolved for something. But animals with small brains (particularly insects and birds) appear capable of a wide range of cognitive abilities that rival those of large-brained mammals (Perry, Barron and Cheng, 2013). Further, computational modelling has emphasised how even clever cognitive abilities can be demonstrated by neural networks with extremely small numbers of elements. These call into question many of the presumed benefits of large brains with large numbers of neurons. Here I draw on computational modelling to propose specific advantages of large brains with high neuron numbers for animals, and reconsider the comparative cognition literature to discuss the kinds of cognitive limitations imposed by very small brains.

Mapping Out The Landscape: A Multi-dimensional Approach To Behavioural Innovation
Rachael Brown

In the context of cognitive evolution, transformations in the "behavioural innovativeness" of species—broadly, the capacity to generate new or novel behaviours (Reader and Laland, 2003)—are often associated with significant evolutionary shifts in cognition (Arbilly and Laland, 2017). Whilst this assumption is intuitively and theoretically appealing, more empirical support is needed. Such empirical work is currently hampered by the lack of a good measure of behavioural innovation in the literature. Where, for example, does any particular instance of a novel behaviour lie within the broader scheme of behavioural innovativeness? Does this match up with what we know about the cognitive and neural complexity of species? This paper offers a solution to this measurement problem.

Existing approaches to behavioural innovation (e.g. Ramsey, Bastian and van Schaik, 2007) are largely binary (i.e., a behaviour is an innovation, or no; a cognitive mechanism is innovative, or not) with more nuanced approaches tending towards uni-dimensionality (Arbilly and Laland, 2017). I present a multi-dimensional framework for characterising and comparing putative cases of behavioural innovativeness. Using paradigm cases of novelty, I identify three key dimensions upon which a putative case of behavioural innovation varies: the novelty of the motor action, the novelty of the context in which the behaviour is deployed, and the novelty of the task or problem the behaviour responds to. I then demonstrate how we can use qualitative measures of each of these three dimensions for any given behavioural innovation to offer a fine-grained account of what constitutes that innovation. We can then compare putative cases of novelty along these dimensions.

This framework has a number of virtues: it allows us to capture similarities and differences between particular cases of behavioural innovation, it allows us to identify commonalities within clusters of behavioural innovation in the tree of life, and it ultimately allows us to identify the impact of major transitions in cognition on the innovativeness of species. Ultimately, it will allow us to test the idea that transitions in cognition generate transformations in behavioural innovativeness.

Associative Learning as a Null Model
Marta Halina

Morgan's Canon is a methodological principle widely adopted in research on animal cognition. The principle holds that when there are two or more possible explanations for an animal's behavior, psychologists should favor the explanation that posits the simpler or less sophisticated process. Philosophers have argued that this principle is deeply problematic. What counts as more or less sophisticated is ambiguous (Fitzpatrick, 2009; Sober, 2005). Simplicity has been used to distinguish between sensory and conceptual reasoning, stimulus-response mechanisms and conscious thought, associative and non-associative learning. In all of these cases, the contrasts are problematic (Andrews and Huss, 2014). Morgan's Canon also lacks clear justification as a general methodological principle. One commonly proposed justification is to interpret the Canon as a version of the epistemic virtue of parsimony, but parsimony takes many forms (ontological, explanatory, and evolutionary) and often these forms pull in different directions and rarely do they favor the conclusion urged by Morgan's Canon (Dacey, 2016; Sober, 2005).

The major transitions approach avoids framing cognitive diversity in terms of complex versus simple, or clever versus 'killjoy' explanations of cognition. On the other hand, the transitions framework supplies, I will argue, alternative "null models" for cognitive research. Building on Ginsburg and Jablonka's (2019) account of limited and unlimited associative learning (AL) as a major evolutionary transition, I show how AL is a compelling candidate for a null model in comparative psychology. Organisms with unlimited AL have a vast and open-ended capacity to learn and adjust to stimuli, increasing the possibility that they have solved a given cognitive task through AL. Associative learning is also phylogenetically deep: limited AL is found in at least six phyla and unlimited AL likely emerged in arthropods and vertebrates soon after they evolved (Bronfman, Ginsburg and Jablonka, 2016). These features of AL make it a promising candidate as a "null model" in the sense recently defended in the philosophical literature (Bausman, 2018). Although understanding AL as a null model should be distinguished from traditional accounts of Morgan's Canon, it helps explain to some extent the persistent appeal of the Canon in comparative research.

Explaining Neural Transitions through Resource Constraints
Colin Klein

One challenge in explaining neural evolution is the formal equivalence of a variety of different computational architectures. Well-known results show that various architectures, including neural networks with a single hidden layer and a nonlinear activation function, can be universal function approximators. Why change? The answer must involve the intense competition for resources—including time, space, and energy—that brains operate under (Sterling and Laughlin, 2015). I argue that such explanations are ultimately an abstract species of resource explanation (Klein, 2018), which are distinct from but complementary to explanations in terms of mechanical parts. Resource explanations play an important role in computer science, one that is often under-appreciated by philosophers of neuroscience (Aaronson, 2013). As a case study, I explore the internalization of external loops for sensorimotor feedback. This process is one plausible way by which major transitions in neural organisation may occur. I demonstrate that this process is driven by the interplay between competing demands on limited resources, and that transition by internalization trades off increased complexity against more efficient use of existing resources. The resulting framework offers an explanation of why major transitions can occur, and shows why organisms on either side of a transition boundary may have very similar cognitive capacities but very different potential for evolving new capacities.

References

Aaronson, S. (2013). Why philosophers should care about computational complexity. In B. J. Copeland, J. Posy Carl & O. Shagrir (Eds.), *Computability: Turing, gödel, church, and beyond* (pp. 261–327). MIT Press Cambridge, MA.

Andrews, K. & Huss, B. (2014). Anthropomorphism, anthropectomy, and the null hypothesis. *Philosophy and Biology*, *29*, 711–729.

Arbilly, M. & Laland, K. N. (2017). The magnitude of innovation and its evolution in social animals. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1848), 1–9.

Bausman, W. C. (2018). Modeling: Neutral, null, and baseline. *Philosophy of Science*, *85*(4), 594–616.

Bronfman, Z. Z., Ginsburg, S. & Jablonka, E. (2016). The transition to minimal consciousness through the evolution of associative learning. *Frontiers in psychology*, *7*, 1954.

Dacey, M. (2016). The varieties of parsimony in psychology. *Mind & Language*, *31*(4), 414–437.

Fitzpatrick, S. (2009). The primate mindreading controversy: A case study in simplicity and methodology in animal psychology. In R. W. Lurz (Ed.), *The philosophy of animal minds* (pp. 258–277). Oxford: Cambridge University Press.

Ginsburg, S. & Jablonka, E. (2019). *The evolution of the sensitive soul: Learning and the origins of consciousness*. MIT Press.

Godfrey-Smith, P. (2006). Local interaction, multilevel selection, and evolutionary transitions. *Biological Theory*, *1*(4), 372–380.

Klein, C. (2018). Mechanisms, resources, and background conditions. *Biology & Philosophy*, *33*(5-6), 36.

Logan, C. J., Avin, S., Boogert, N., Buskell, A., Cross, F. R., Currie, A., . . . Navarrete, A. F. et al. (2018). Beyond brain size: Uncovering the neural correlates of behavioral and cognitive specialization. *Comparative Cognition Society*, *30*, 55–90.

Perry, C. J., Barron, A. B. & Cheng, K. (2013). Invertebrate learning and cognition: Relating phenomena to neural substrate. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(5), 561–582.

Ramsey, G., Bastian, M. L. & van Schaik, C. (2007). Animal innovation defined and operationalized. *Behavioral and Brain Sciences*, *30*(4), 393–407.

Reader, S. M. & Laland, K. N. (2003). *Animal innovation*. Oxford University Press Oxford.

Sober, E. (2005). Comparative psychology meets evolutionary biology. In L. Daston & G. Mitman (Eds.), *Thinking with animals: New perspectives on anthropomorphism* (pp. 85–99). Columbia University Press.

Sterling, P. & Laughlin, S. (2015). *Principles of neural design*. MIT Press.

Szathmáry, E. & Smith, J. M. (1995). *The major transitions in evolution*. WH Freeman Spektrum Oxford, UK.

Trestman, M. (2013). The cambrian explosion and the origins of embodied cognition. *Biological Theory*, *8*(1), 80–92.

Biographical Descriptions

Andrew Barron is an Australian Research Council Future Fellow, and Deputy Head of the Department of Biological Sciences at Macquarie University. He is a neuroethologist, which is a discipline of neuroscience studying the neural mechanisms of natural animal behaviour. Most of his research focuses on insects, especially honey bees. Using advanced techniques to visualise, manipulate, map and record from the insect brain Barron's team has made important contributions to the understanding of fundamental behavioural systems such as cognition, navigation, social behaviour and learning and memory. Barron also conducts research to improve honey bee health and welfare. He is studying how bees and bee colonies are impacted by pesticide and disease stressors, and how to best intervene to help bee colonies under stress.

Rachael Brown is the Director of the Centre for Philosophy of the Sciences and Lecturer at the School of Philosophy in the Research School of Social Sciences at the Australian National University. She is also a member of Tempo & Mode: Centre for Macroevolution and Macroecology (ANU), Sydney-ANU Philosophy of Biology Group and an affilliate member of the Centre for Agency, Values and Ethics at Macquarie University. Brown works primarily at the intersection of the philosophy of biology, philosophy of cognitive science, and philosophy of science. She is particularly interested in the evolution of cognition and behaviour; the relationship between evo-devo and the Neo-Darwinian Synthesis; model-based reasoning in biology and philosophy; and methodological issues in the study of animal behaviour and cognition.

Marta Halina is University Lecturer in the Department of History and Philosophy of Science at the University of Cambridge. Halina founded and directs the Kinds of Intelligence program at the Leverhulme Centre for the Future of Intelligence, which draws on current work in psychology, neurobiology, computer science and philosophy to develop and critically assess notions of intelligence. Halina also co-organises the Animal-AI Olympics, which benchmarks current AI against animal species using a range of established animal cognition tasks. In addition to her philosophical writings on animal minds, artificial intelligence and scientific methods, Halina has designed and implemented experiments for testing the social cognitive abilities of nonhuman primates.

Colin Klein is an Associate Professor in the School of Philosophy at the Australian National University (ANU). He is an Australian Research Council Future Fellow, ANU Futures Awardee, and Foundations Lead of ANU's Humanising Machine Intelligence project. He is also a founder of the Australasian Society for Philosophy and Psychology. Klein works on philosophy of science and philosophy of mind, especially in methodological issues around philosophy of neuroscience. He has published widely, including in *Philosophy of Science*, *The British Journal for the Philosophy of Science*, Proceedings of the National Academy of Sciences, and Neuroimage. His 2015 book *What the Body Commands: The Imperative Theory of Pain*, with MIT Press, recently won the David Harold Tribe Prize.