Accepted manuscript (in press, copyedited)

Behavioral and Brain Sciences - Commentary on Heyes

Tinkering with cognitive gadgets: Cultural evolutionary psychology meets active inference

Paul Benjamin Badcock,[a,b,c]

Axel Constant,[d,e,f]

Maxwell James Désormeau Ramstead[e,f,g,h]

[a]*Centre for Youth Mental Health, The University of Melbourne, Parkville, VIC 3052, Australia;* [b]*Melbourne School of Psychological Sciences, The University of Melbourne, VIC 3010, Australia;* [c]*Orygen, the National Centre of Excellence in Youth Mental Health, Parkville, VIC 3052, Australia;* [d]*Charles Perkins Centre, Johns Hopkins Drive, The University of Sydney, NSW 2006, Australia;* [e]*Wellcome Trust Centre for Neuroimaging, University College London WC1N3BG, United Kingdom;* [f]*Culture, Mind, and Brain Program, McGill University, 1033 Pine Avenue, Montreal, QC, Canada;* [g]*Department of Philosophy, McGill University, Montreal, QC, H3A 2T7, Canada;* [h]*Division of Social and Transcultural Psychiatry, Department of Psychiatry, McGill University, Montreal, QC, H3A 1A1, Canada.*

pbadcock@unimelb.edu.au

axel.constant.pruvost@gmail.com

maxwell.d.ramstead@gmail.com

Abstract

Cognitive Gadgets offers a new, convincing perspective on the origins of our distinctive cognitive faculties, coupled with a clear, innovative research program. Although we broadly endorse Heyes' ideas, we raise some concerns about her characterisation of evolutionary psychology and the relationship between biology and culture, before discussing the potential fruits of examining cognitive gadgets through the lens of active inference.

Heyes' (2018) *Cognitive Gadgets* presents a compelling and erudite case for the influence of cultural evolution on the emergence of the distinctive cognitive gadgets unique to humans. Her argument is a convincing one – she offers a lucid treatment of four empirically supported examples to substantiate her claims, and highlights a number of viable research avenues to open her theory to scientific scrutiny. *Cognitive Gadgets* no doubt stands to make a meaningful contribution to current thinking about the evolutionary roots of human cognition. Admittedly, our opinion is also biased – *cultural evolutionary psychology* resonates with our own work, and we hope to see its widespread adoption as a research program.

There are, however, a number of points worth raising about Heyes' misleading characterisation of evolutionary psychology. First, the decidedly gene-centric view she attributes to members of the "High Church" of evolutionary psychology is somewhat specious – to understand and model gene-environment interactions and the ways in which social and physical ecologies impact the development of human cognitive specialisations is a core commitment of the field (Geary & Bjorklund 2000; Krebs 2003; Ploeger et al. 2008). Heyes also overlooks the heterogeneity of evolutionary psychology, broadly construed. For example, consistent with her own

teleosemantic perspective, proponents of an "evolutionary systems" approach adopt an expanded view of inheritance, which extends beyond the gene to incorporate other forms of intergenerational information transmission, including epigenetic effects and the inheritance of cultural artefacts and patterned cultural practices (Badcock 2012; Caporael 2001; Hendriks-Jansen 1996; Laland 2017; Lickliter & Honeycutt 2003; Ramstead et al. 2016; Roepstorff et al. 2010).

By emphasising Chomsky's (1965) "poverty of stimulus" criterion and our innate cognitive "starter kit," Heyes also creates the impression that evolved, biological traits that help us adapt to the social milieu should largely be relegated to early biases, which are then refined by nurture and culture. However, there are clearly many adaptive traits that show a strong biological basis and profoundly affect social cognition and behaviour throughout the lifespan. Obvious examples include the gross morphology and physiological properties of the brain (Friston 2010); hormonal and neuromodulatory systems (Katz & Harris-Warrick 1999; McGlothlin & Ketterson 2008); reward, mood, and affective systems (Gray 1994; Nettle & Bateson 2012); personality traits (Bouchard & Loehlin 2001; Nettle 2006); and sensitive periods of development (e.g., puberty) that fine-tune our adaptation to different socio-environmental contexts across the life course (Frankenhuis & Fraley 2017; Geary & Bjorklund 2000).

More critically, *Cognitive Gadgets* tends to neglect the fundamental role of biology in shaping our cultural worlds. Dynamical simulation studies have provided proof of principle that individual differences in adaptive decision rules (e.g., mating or social learning strategies) create marked changes in the self-organisation of social norms and cultural dynamics (Kenrick et al. 2003; Molleman et al. 2014). More substantively, evolutionary psychologists have accumulated a wealth of evidence to

suggest that cognitive traits favoured by natural selection exert a powerful influence on the sorts of cultural expectations, norms, and practices that are likely to evolve – including those surrounding communal sharing and morality (Kameda et al. 2005; Krebs 2008), in-group versus out-group behaviours (Brewer 2007), and behaviours involving social exchange (Wischniewski et al. 2009). Such work reminds us that culture shouldn't be individuated from the evolved biobehavioural dynamics of the individuals that comprise it (Kenrick et al. 2003; Lehman et al. 2004). The idea that we gain remarkable cognitive capacities via cultural learning almost goes without saying – but this should not detract from a dialectical view that sees biological and cultural inheritance as mutually constitutive.

As such, our main reservation with *Cognitive Gadgets* is that it promotes a sharp distinction between nature, nurture, and culture. As evidenced by our own work in this area (Constant et al. 2018; Kirmayer & Ramstead 2017; Ramstead et al. 2016; 2019), we certainly agree that cultural influences play an essential role in both the inheritance and development of our adaptive cognitive specialisations. However, we do not think that these three sources can be pried apart so easily – our interest lies more in the ways they *interact* in order to produce human phenotypes. Human biology is also a cultural biology; and human culture is realised by interacting biological systems within a shared material niche (Kirmayer & Ramstead 2017). Indeed, for over 200,000 years, the main selection pressure on human survival has been the capacity to access and leverage accumulated sociocultural information (Henrich 2015; Hrdy 2009; Tomasello 2014). Heyes wouldn't deny this – she proposes that humans begin with a genetically specified "starter kit" that is geared towards navigating the sociocultural world, which allows more sophisticated forms of social cognition, like literacy and mindreading, to develop. Arguably, such "innate"

propensities only ever emerged because of the increasing importance of sociocultural information for human survival.

Accordingly, we think that one of the most interesting questions raised by *Cognitive Gadgets* relates to the mechanisms that explain *how* biological, social, and cultural dynamics interact. This appeals to an *evolutionary systems* perspective that is able to capture both the ways in which biological and cultural evolution shape individual minds, and the ways in which individual minds shape culture and biology (Badcock 2012; Caporael 2001; Kenrick 2001). Of course, such multilevel dynamics are challenging to understand, let alone study.

Fortunately, a promising approach has emerged from neuroscience and theoretical biology that has the potential to provide such a multiscale modelling strategy, called *active inference*. A descendent of *predictive coding* schemes of the brain (see Clark 2013), active inference is a mathematical formulation that describes how living systems are able to maintain themselves within a limited range of phenotypic states; that is, within the set of states in which they expect to find themselves, on average and over time. It explains how biological systems appear to resist the natural tendency to dissipate into their environment by fulfilling biologically instantiated (Bayesian) *prior beliefs*, or expectations about the way the world unfolds (Friston 2010; Friston et al. 2009). In short, organisms are driven by the biological imperative to maintain homeostasis via action-perception loops that actively minimise "surprise." This framework has recently been extended to explain the evolution, development, and multiscale dynamics of living systems in general (Friston 2013; Kirchhoff et al. 2018; Ramstead et al. 2018). Here, we would like to concentrate on two of our own complementary approaches derived from active inference, which we believe are particularly relevant to Heyes' proposal.

On the one hand, Heyes' appeal to our unique cognitive specialisations connects with a new theory of the human brain, cognition, and behaviour called the *hierarchically mechanistic mind* (HMM); see Badcock et al. (2019a; 2019b). This model rests on two fundamental claims. The first follows active inference by suggesting that the brain is a complex adaptive system comprising hierarchically organised neurocognitive mechanisms that function to reduce the dispersion or decay of our sensory and physical states by producing action-perception cycles that seek to minimise surprise (Badcock et al. 2019a). The second claim follows Tinbergen's (1963) four questions in ethology (i.e., adaptation, phylogeny, ontogeny, and mechanism) by suggesting that neural form and function can only be understood in terms of the broader evolutionary, intergenerational, developmental, and real-time processes that act on human phenotypes, which are differentially illuminated by major paradigms in psychology (i.e., evolutionary psychology, evo-devo, developmental psychology, and psychology's subdisciplines respectively; Badcock et al. 2019b). Thus, in order to understand a phenotypic trait, researchers need to develop multiscale hypotheses that synthesise findings from diverse fields of psychological inquiry in order to explain both *why* that trait is adaptive; along with *how* it emerges from evolutionary, intergenerational, developmental, and real-time processes (for an application to depression, see Badcock et al. 2017). The HMM situates Heyes' work within a broader meta-theory of psychological inquiry that sees cultural evolutionary psychology as but one viable approach to understanding the evolution of human traits – a paradigm, like evo-devo, that concentrates on the group-level, *intergenerational* dynamics that bridge human evolution and development, thereby driving phylogenetic change. This *complements* insights from other evolutionary paradigms; it certainly doesn't *contradict* them.

To borrow Heyes' own term, active inference also supplies a plausible "force theory" of the nested socio-environmental dynamics responsible for the evolution and development of cognitive gadgets at the level of the individual. This brings us to our second complementary approach, called the *variational approach to niche construction* (VANC) (Constant et al. 2018)).

The VANC considers niche construction – i.e., implicit and explicit modifications of the environment – as a corollary of active inference, whereby embrained expectations guiding adaptive action-perception loops come to be encoded in the material layout of human niches. Take, for instance, desire paths. As they cut through a grassy field on their daily commute, people implicitly leave traces that inform other pedestrians of the possible intentions of those who crossed the path before; for example, "I want to reach the eastern exit of the park." By engaging the well-worn path, a novice agent can zero in on the optimal route without having any knowledge of the park's design. Our approach here is consistent with Heyes' – all this agent would require is the propensity to let herself be guided by the path. This disposition is made possible, presumably, by a minimal *starter kit* that includes basic *cooperative sensory and motor dispositions*, as well as some expectations regarding what she herself desires to do; for example, "I expect that my action will lead me to the eastern exit." By engaging the desire path, the agent will further wear down the trail, thereby increasing its reliability for others. Of course, another example of niche construction, which relates directly to Heyes' treatment of language, is written text.

The point here is that, through niche construction, humans produce culturally specific behavioural patterns encoded in the constructed artefacts that populate their niche, which they can then recruit to support the performance of various tasks. In so doing, they often implicitly and automatically converge on statistically recurrent

behaviours, which, following active inference, are the least surprising ones (i.e., those that characterise the local cultural phenotype) (Constant et al., 2019). Over evolutionary time, cultural evolution scaffolds and finesses progressively complex "nurtural" networks of externally realised expectations (in patterned cultural practices and constructed niches), thereby guiding and transmitting increasingly sophisticated cultural behaviour (Constant et al. 2018; Ramstead et al. 2016; Veissière et al. 2019).

Under both the HMM and VANC, cognitive gadgets can be described as heritable *adaptive priors* that underlie (neuronally encoded) expectations about the dynamics of the social world and guide our action-perception cycles towards unsurprising states (see Badcock et al. 2019b). According to this perspective, such priors have emerged from the reciprocal interplay of biological, sociocultural, and ecological dynamics over evolutionary time because they have afforded a reliable means to reduce socio-environmental uncertainty. In other words, cognitive gadgets can be thought to entail hierarchical architectures of adaptive prior expectations encoded at multiple levels and sites, spanning neural systems, human phenotypes, social interactions, culturally specified motor patterns, and ecological structures.

Arguably, this notion adds to Heyes' proposal in two important ways. First, it avoids the questionable claim that our distinctive cognitive faculties are chiefly cultural products by suggesting that nature, nurture, and culture operate synergistically to optimise our phenotypes and eco-niches over evolutionary, intergenerational, and developmental timescales. Shedding light on such faculties requires recourse to research that spans the full breadth of evolutionary psychology (cultural or otherwise), not to mention allied disciplines like anthropology, biology, and ecology. Second, by subsuming active inference, the HMM and VANC afford a single, common language that allows us to describe both biological and cultural

influences on human phenotypes mathematically; to model their interactions computationally; to test these models via simulation studies; and then compare the outcomes of such *in silico* research with real-world experiments and observations (Badcock et al. 2019a; Ramstead et al. 2018).

In sum, our approach to human cognition builds on Heyes', but blurs the lines between nature, nurture, and culture – proposing instead a single, generic information theoretic mechanism (i.e., active inference) that expresses itself in different, complementary ways across all three. With this in mind, we suggest that active inference would make a powerful addition to the explanatory toolbox of cultural evolutionary psychology.

References

Badcock, P.B. (2012). Evolutionary systems theory: a unifying meta-theory of psychological science. *Review of General Psychology, 16*, 10-23. doi:10.1037/a0026381

Badcock, P.B., Davey, C., Whittle, S., Allen, N.B., & Friston, K.J. (2017). The depressed brain: an evolutionary systems theory. *Trends in Cognitive Sciences, 21(3),* 182–194. doi:10.1016/j.tics.2017.01.005

Badcock, P.B., Friston, K.J., & Ramstead, M.J.D. (2019a). The hierarchically mechanistic mind: A free-energy formulation of the human psyche. *Physics of Life Reviews.* Advance online publication. doi:10.1016/j.plrev.2018.10.002

Badcock, P. B., Friston, K. J., Ramstead, M. J., Ploeger, A., & Hohwy, J. (2019b). The hierarchically mechanistic mind: An evolutionary systems theory of the human brain, cognition, and behavior. *Cognitive, Affective, & Behavioral Neuroscience*. Advance online publication. doi:10.3758/s13415-019-00721-3

Bouchard, T.J., & Loehlin, J.C. (2001). Genes, evolution, and personality. *Behavior Genetics, 31(3),* 243-273. doi:10.1023/A:1012294324713

Brewer, M. B. (2007). The importance of being we: Human nature and intergroup relations. *American Psychologist, 62*(8), 728-738. doi:10.1037/0003-066x.62.8.728

Caporael, L.R. (2001). Evolutionary psychology: Toward a unifying theory and a hybrid science. *Annual Review of Psychology, 52,* 607-628. doi:10.1146/annurev.psych.52.1.607

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181-204. doi:10.1017/S0140525X12000477

Constant, A., Ramstead, M. J., Veissiere, S. P., Campbell, J. O., & Friston, K. J. (2018). A variational approach to niche construction. *Journal of The Royal Society Interface, 15*(141), 20170685. doi:10.1098/rsif.2017.0685

Constant, A., Ramstead, M., Veissière, S., & Friston, K. J. (2019). Regimes of expectations: An active inference model of social conformity and decision making. *Frontiers in Psychology, 10*, 679. doi:10.3389/fpsyg.2019.00679

Frankenhuis, W.E., & Fraley, R.C. (2017). What do evolutionary models teach us about sensitive periods in psychological development? *European Psychologist, 22,* 141-150. doi:10.1027/1016-9040/a000265

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews: Neuroscience, 11*, 127-138. doi:10.1038/nrn2787

Friston, K. (2013). Life as we know it. *Journal of the Royal Society Interface, 10(86),* 20130475. doi:10.1098/rsif.2013.0475

Friston, K.J., Daunizeau, J., & Kiebel, S.J. (2009). Reinforcement learning or active inference? *PLOS ONE, 4*(7), e6421. doi:10.1371/journal.pone.0006421

Geary, D.C., & Bjorklund, D.F. (2000). Evolutionary developmental psychology. *Child Development, 71(1),* 57-65. doi:10.1111/1467-8624.00118

Gray, J.A. (1994). Personality dimensions and emotion systems. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 329–331). New York: Oxford University Press.

Hendriks-Jansen, H. (1996). *Catching ourselves in the act*. Cambridge, MA: MIT Press.

Henrich, J. (2015). *The secret of our success: how culture is driving human evolution, domesticating our species, and making us smarter*. Princeton, NY: Princeton University Press.

Heyes, C. (2018). *Cognitive gadgets: The cultural evolution of thinking*. Cambridge, MA: Harvard University Press.

Hrdy, S.B. (2009). *Mothers and others*. Cambridge MA: Harvard University Press.

Kameda, T., Takezawa, M., & Hastie, R. (2005). Where do social norms come from? The example of communal sharing. *Current Directions in Psychological Science, 14*(6), 331-334. doi:10.1111/j.0963-7214.2005.00392.x

Katz, P. S., & Harris-Warrick, R. M. (1999). The evolution of neuronal circuits underlying species-specific behavior. *Current Opinion in Neurobiology, 9(5),* 628-633. doi:10.1016/s0959-4388(99)00012-4

Kenrick, D. T. (2001). Evolutionary psychology, cognitive science, and dynamical systems: Building an integrative paradigm. *Current Directions in Psychological Science, 10*, 13-17. doi:10.1111/1467-8721.00104

Kenrick, D. T., Li, N.P., & Butner, J. (2003). Dynamical evolutionary psychology: Individual decision rules and emergent social norms. *Psychological Review, 110*(1), 3-28. doi:10.1037/0033-295X.110.1.3

Kirchhoff, M., Parr, T., Palacios, E., Friston, K., & Kiverstein, J. (2018). The Markov blankets of life: autonomy, active inference and the free energy principle. *Journal of The Royal Society Interface, 15*(138), 20170792. doi:10.1098/rsif.2017.0792

Kirmayer, J.L., & Ramstead, M.J.D. (2016). Embodiment and enactment in cultural psychiatry. In C. Durt, T. Fuchs, & C. Tewes (Eds.), *Embodiment, enaction,*

and culture: Investigating the constitution of the shared world. Cambridge, MA: MIT Press.

Krebs, D.L. (2003). Fictions and facts about evolutionary approaches to human behaviour: Comment on Lickliter and Honeycutt (2003). *Psychological Bulletin, 129*, 842–847. doi:10.1037/0033-2909.129.6.842

Krebs, D.L. (2008). Morality: An evolutionary account. *Perspectives on Psychological Science, 3*(3), 149-172. doi:10.1111/j.1745-6924.2008.00072.x

Laland, K. N. (2017). *Darwin's unfinished symphony: How culture made the human mind*. Princeton, NJ: Princeton University Press.

Lehman, D.R., Chiu, C.Y., & Schaller, M. (2004). Psychology and culture. *Annual Review of Psychology, 55,* 689-714. doi:10.1146/annurev.psych.55.090902.141927

Lickliter, R., & Honeycutt, H. (2003). Developmental dynamics: toward a biologically plausible evolutionary psychology. *Psychological Bulletin, 129,* 819-835. doi:10.1037/0033-2909.129.6.819

McGlothlin, J.W., & Ketterson, E.D. (2008). Hormone-mediated suites as adaptations and evolutionary constraints. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 363*(1497), 1611-1620. doi:10.1098/rstb.2007.0002

Molleman, L., Van den Berg, P., & Weissing, F.J. (2014). Consistent individual differences in human social learning strategies. *Nature Communications, 5,* 3570. doi:10.1038/ncomms4570

Nettle, D. (2006). The evolution of personality variation in humans and other animals. *American Psychologist, 61*(6), 622-631. doi:10.1037/0003-066X.61.6.622

Nettle, D., & Bateson, M. (2012). The evolutionary origins of mood and its disorders. *Current Biology, 22*(17), R712-R721. doi:10.1016/j.cub.2012.06.020

Ploeger, A., van der Maas, H. L., & Raijmakers, M. E. (2008). Is evolutionary psychology a metatheory for psychology? A discussion of four major issues in psychology from an evolutionary developmental perspective. *Psychological Inquiry, 19*(1), 1-18. doi:10.1080/10478400701774006

Ramstead, M.J.D., Badcock, P., & Friston, K.J. (2018). Answering Schrödinger's question: A free-energy formulation. *Physics of Life Reviews, 24*, 1-16. doi:10.1016/j.plrev.2017.09.001

Ramstead, M.J., Veissière, S.P., & Kirmayer, L.J. (2016). Cultural affordances: scaffolding local worlds through shared intentionality and regimes of attention. *Frontiers in Psychology, 7*, 1090. doi:10.3389/fpsyg.2016.01090

Ramstead, M.J.D., Constant, A., Badcock, P.B., & Friston, K.J. (2019). Variational Ecology and the physics of sentient systems. *Physics of Life Reviews.* Advance online publication. doi:10.1016/j.plrev.2018.12.002

Roepstorff, A., Niewöhner, J., & Beck, S. (2010). Enculturing brains through patterned practices. *Neural Networks, 23*(8-9), 1051-1059. doi:10.1016/j.neunet.2010.08.002

Tinbergen, N. (1963). On aims and methods in ethology. Zeitschrift für Tierpsychologie, 20, 410-433. doi: 10.1111/j.1439-0310.1963.tb01161.x

Tomasello, M. (2014). *A natural history of human thinking*. Cambridge, MA: Harvard University Press.

Veissière, S. P. L., Constant, A., Ramstead, M. J. D., Friston, K. J., & Kirmayer, L. J. (2019). Thinking through other minds: A variational approach to cognition

and culture. *Behavioral and Brain Sciences*. Advance online publication. doi:10.1017/S0140525X19001213.

Wischniewski, J., Windmann, S., Juckel, G., & Brüne, M. (2009). Rules of social exchange: Game theory, individual differences and psychopathology. *Neuroscience & Biobehavioral Reviews, 33*(3), 305-313. doi:10.1016/j.neubiorev.2008.09.008