Mending wall*

Charles Rathkopf†
Daniel Dennett‡

Abstract

Heyes suggests that selective social learning comes in two varieties. One is common, domain general, and associative. The other is rare, domain specific, and metacognitive. We argue that this binary distinction cannot quite do the work she assigns it and sketch a framework in which additional strategies for selective social learning might be accommodated.

*This is a preprint of a commentary on Cecelia Heyes' book "Cognitive Gadgets: The Cultural Evolution of Thinking." The commentary appears in *Behavioral and Brain Sciences* in November of 2019.

†Institute for Brain and Behavior, Jülich Research Center
‡Center for Cognitive Science, Tufts University

Robert Frost's (1915) poem uncovers the costs and benefits of crisp boundaries. It is true both that sometimes "good fences make good neighbors" and that fences can stand in the way of understanding. In her insightful account of how genetic evolution and cultural evolution working together could have produced our modern human minds, Cecilia Heyes (2018) introduces and vividly names some valuable distinctions – most importantly between cognitive instincts and cognitive gadgets, between "Big Special" and "Small Ordinary" cognitive gifts, between mills and grist, and between metacognitive rules and other useful dispositions. But "something there is that doesn't like a wall," as Frost said, and that something is nature: the gradual effects of variation, differential reproduction, decay, and inauspicious birth. Heyes' distinctions are fine contributions to the task of explaining the evolutionary trajectory from animal to human cognition, but they are not all as crisp as she suggests. Acknowledging this can save her account from a variant of the well-worn chicken-and-egg problem. Which came first: cultural evolution or metacognition? To see why this looks like a problematic question for Heyes, note that her account provides support for each of the following four claims.

1. Cultural evolution requires transmission fidelity. (p. 112)
2. Transmission fidelity requires focused selectivity in social learning. (p. 111)
3. Focused selectivity in social learning requires metacognition. (p. 111)
4. Metacognition is a product of cultural evolution. (p. 107)

The apparent circularity implied by these claims stems from the fact that metacognition is described both as a product of cultural evolution and as one of its drivers. Of course, Darwin showed that chicken-and-egg problems like this are not nearly as perplexing as they first appear. The hint of paradox disappears as soon as we consider the role of intermediate forms. Metacognition must have emerged gradually – perhaps by means of a cultural evolutionary process that was noisier than the higher-fidelity process it subsequently made possible. And if metacognition did evolve gradually, there must be (or must have been) some intermediate cognitive form(s).

> I suggest that the crucial, culture-relevant difference between selective social learning in humans and other animals is that some

> human social learning is made selective by explicit metacognition (Shea et al., 2014): by conscious, reportable, domain-specific rules... (Heyes, 2018, pp. 105–106).

You can't follow an explicit rule that you don't understand, so comprehension is crucial in some human social learning, but what about the rest of it? Couldn't there be "rules" – don't there have to be "rules" – that are inexplicit, semi-understood free-floating rationales (Dennett, 1983, 2017) that modulate and control many of the competent behaviors that provided the evolutionary stepping stones to our current cognitive powers? Competence without comprehension must precede competence with comprehension.

Heyes does recognize one form of competence without comprehension in the domain of selective social learning. She discusses social learning strategies found in non-human apes (in addition to humans), and refers to them as planetary, cleverly reminding us of the fact that, just as planets follow Isaac Newton's rules without comprehending them, nonhuman apes can follow learning rules without any ability to reflect on them. For example, monkeys can learn to arrange a series of photos so that it matches the order provided by a human experimenter (Subiaul et al., 2004). Nothing in this behavior demands a metacognitive explanation. Domain-general associative learning would suffice for the monkey to learn that copying that particular human leads to a food reward. For Heyes, this sort of planetary social learning has little in common with the more sophisticated metacognitive variety that supports human cultural transmission. Heyes refers to uniquely human social learning strategies as cook-like, evoking the domain-specific and self-aware kind of social learning familiar to anyone who has tried to cook by following a written recipe.

If Heyes' goal is to remind us of the fact that not all learning requires consciousness and episodic memory, the distinction between planetary and cook-like learning strategies is helpful. However, if her goal is to understand the evolution of metacognition, then, regardless of whether that evolution is genetic or cultural, this binary distinction threatens to blind us to the messy middle ground between planetary and cook-like learning. As is often the case in thinking about evolutionary change, it may help to imagine a multidimensional space of possible learning strategies (Dennett, 2017; Godfrey-Smith, 2009). The planetary and cook-like varieties represent only two extremities in that space. Moreover, it is unlikely that evolution has managed to avoid visiting large subspaces in the interior. In Heyes' own view, the adaptations that mark our trajectory

through that space have been "Small and Ordinary" (p. 53), which rules out large saltation-like leaps.

In our view, Heyes' already excellent treatment of selective social learning could be enriched by acknowledging and then exploiting the inner regions of this space. The attraction of this expansion can be seen clearly when we compare Heyes' treatment of status-based selective learning with her treatment of age-based selective learning. Heyes describes a study by McGuigan (2013), in which 5-year-old children can get help solving a puzzle-box problem from different classes of adults. It turns out that 5-year-olds can rank the social status of adults, and then tune their social learning strategy to ensure that only high-status models get copied. Although this looks like a case of children regulating their social learning in just the way cultural evolutionary theory requires, Heyes offers this study as an example of domain general, associative, and non-metacognitive learning. Contrast this with her most prominent example of genuine metacognition, the rule that instructs us to copy digital natives. Metacognition is defined as thinking about thinking. So what makes this rule metacognitive? Can't we construe it as a rule about what to do, rather than a rule about what to think? One might say to oneself: "If, in the future, you happen to be thinking about which app to download, copy digital natives." If one were to subvocalize that sentence, it would be a clear case of metacognition in Heyes' sense. But we see no reason that a rule with this content must be acquired by such metacognitive means.

In these two cases, we have an exemplar of purportedly non-metacognitive learning that looks meta, and an exemplar of purportedly metacognitive learning that doesn't look quite so meta after all. One might interpret this as evidence that Heyes' distinction between genuine metacognition and merely planetary social learning is not as mutually exclusive as she makes it out to be. Instead, our suggestion is that these two styles of social learning are not jointly exhaustive. They simply leave out many of the more complex kinds of learning rules that don't fit either category neatly.

The benefits of embracing the messy middle are not exhausted by the opportunity to improve the conceptual framework we use to characterize human social learning. Consider the literature on so-called rational imitation in chimpanzees. Chimpanzees raised in captivity will imitate a human who turns on a light switch with her forehead more often when the human seems to choose that method freely, compared to a condition in which it appears there is no choice, because the experimenter's hands were full (Buttelmann et al., 2007). This

suggests a kind of social learning selectivity that isn't purely planetary, since it displays some sensitivity to the possibility of there being a rationale behind the forehead technique. Or, think of the second-order confidence "judgments" of monkeys (Middlebrooks and Sommer, 2012). Monkeys will place large bets on judgments they are highly likely to get right, and smaller bets on judgments they are less likely to get right. These decisions might accurately be described as only *sorta* (Dennett, 2013) metacognitive, and none the worse for that.

Our suggestion is not that these partial cases might, appearances to the contrary, suffice for cultural evolution. Rather, our suggestion is that phenomena like these provide clues about the kind of learning strategies that occupy the messy middle ground between planetary and cook-like learning. In fact, Heyes' wonderful term, "gadget," is ideally suited to play the role of a semi-understood, semi-appreciated found object that an agent might put to good use without fully understanding why.

References

Buttelmann, D., Carpenter, M., Call, J., and Tomasello, M. (2007). Enculturated chimpanzees imitate rationally. *Developmental science*, 10(4):F31–F38.

Dennett, D. C. (1983). Intentional systems in cognitive ethology: The "panglossian paradigm" defended. *Behavioral and Brain Sciences*, 6(3):343–355.

Dennett, D. C. (2013). *Intuition pumps and other tools for thinking*. WW Norton & Company.

Dennett, D. C. (2017). *From Bacteria to Bach and Back: The Evolution of Minds*. Penguin UK.

Frost, R. (1915). *North of Boston*. Henry Holt.

Godfrey-Smith, P. (2009). *Darwinian populations and natural selection*. OUP Oxford.

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

McGuigan, N. (2013). The influence of model status on the tendency of young children to over-imitate. *Journal of experimental child psychology*, 116(4):962–969.

Middlebrooks, P. G. and Sommer, M. A. (2012). Neuronal correlates of metacognition in primate frontal cortex. *Neuron*, 75(3):517–530.

Shea, N., Boldt, A., Bang, D., Yeung, N., Heyes, C., and Frith, C. D. (2014). Supra-personal cognitive control and metacognition. *Trends in cognitive sciences*, 18(4):186–193.

Subiaul, F., Cantlon, J. F., Holloway, R. L., and Terrace, H. S. (2004). Cognitive imitation in rhesus macaques. *Science*, 305(5682):407–410.