

Kent Academic Repository

Full text document (pdf)

Citation for published version

Carruthers, Peter and Williams, David M. (2019) Comparative Metacognition. *Animal Behavior and Cognition*, 6 (4). pp. 278-288. ISSN 2372-5052.

DOI

<https://doi.org/10.26451/abc.06.04.08.2019>

Link to record in KAR

<https://kar.kent.ac.uk/78263/>

Document Version

Publisher pdf

Copyright & reuse

Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

Versions of research

The version in the Kent Academic Repository may differ from the final published version.

Users are advised to check <http://kar.kent.ac.uk> for the status of the paper. **Users should always cite the published version of record.**

Enquiries

For any further enquiries regarding the licence status of this document, please contact:

researchsupport@kent.ac.uk

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at <http://kar.kent.ac.uk/contact.html>



Comparative Metacognition

Peter Carruthers^{1*} and David M. Williams²

¹ Department of Philosophy, University of Maryland, USA

² Department of Psychology, University of Kent, UK

*Corresponding author (Email: pcarruth@umd.edu)

Citation – Carruthers, P., & Williams, D. M. (2019). Comparative metacognition. *Animal Behavior and Cognition*, 6(4), 278-288. <https://doi.org/10.26451/abc.06.04.08.2019>

Abstract – We argue that comparative psychologists have been too quick to jump to metacognitive interpretations of their data. We examine two such cases in some detail. One concerns so-called “uncertainty monitoring” behavior, which we show to be better explained in terms of first-order estimates of risk. The other concerns informational search, which we argue is better explained in terms of a first-order curiosity-like motivation that directs questions at the environment.

Keywords – Curiosity, Interest, Metacognition, Question, Risk evaluation, Uncertainty

Extensive evidence has been collected of flexible executively-controlled behavior in monkeys, apes, and other creatures (including birds) (Hanus, Mendes, Tennie, & Call, 2011; Smith, Couchman, & Beran, 2014; Taylor, Elliffe, Hunt, & Gray, 2010; von Bayern, Danel, Auersperg, Mioduszevska, & Kacelnik, 2018). Some of this evidence has been interpreted as demonstrating the presence of metacognitive processes in these animals. We will mostly focus on two such bodies of data here: evidence of “uncertainty monitoring” in monkeys, which is said to show their awareness of their own uncertainty (Smith, Beran, Couchman, & Coutinho, 2008; Smith, Shields, & Washburn, 2003); and evidence of ignorance-based searching and curiosity-like behavior in monkeys, which is said to show their awareness of their need for information (Beran & Smith, 2011; Hampton, Zivin, & Murray, 2004; Rosati & Santos, 2016). We will argue that the data are better explained in terms of first-order appraisals of risk, in the first case, and in terms of a set of first-order affective questioning attitudes, in the second.

Sometimes in this literature it is suggested that the mere fact that executive processes are “above” others in a hierarchy of control is enough to make those processes metacognitive (Smith et al., 2014). But if we abide by definitions of metacognition that are standard in the literature on human psychology, then this is a mistake. For in that literature a metacognitive process is said to be not just above, but *about*, other processes (Dunlosky & Metcalfe, 2009; Flavell, 1979; Nelson & Narens, 1990). Metacognition is defined in the human literature as “thought about thought,” or “thinking about [one’s own] thinking.” It is cognition that is *about* other aspects of cognition, not just controlling the latter but controlling it in a manner that in some way, or to some degree, involves representing it.

Many comparative psychologists have weakened this traditional understanding somewhat, in such a way that metacognition need not involve explicit, conceptual, representation of one’s own cognitive processes as such. There are only two possible interpretations of this move, however—unless the comparative literature is to part ways with its human counterpart altogether. One is that metacognition might involve some sort of nonconceptual perception-like monitoring or modeling of the first-order

cognitive process in question. The other is that it might involve thoughts about one's own thoughts without representing them as such, in the same sort of way that one can think "*That* might be dangerous" about an unrecognized shadowy shape in the forest. Here one thinks about what is, in fact, a bear (say), but without representing it as such. Note, however, that on each of these alternative interpretations, metacognition nevertheless involves representations of one's own cognitive processes. A nonconceptual perception of color or texture is still a representation of color or texture; and a thought about an unrecognized shadowy shape is still a representation of it.

The language of "first-order" versus "second-order" or "higher-order" cognitive processes may have contributed to the confusion surrounding these issues. For metacognitive processes are often described as higher-order ones. Yet there are perfectly good senses in which executive processes are "higher-order," in that they asymmetrically depend on the existence of lower-order processes that they control or modulate, as well as being more abstract and "advanced" than those processes. We should stress that in our own writing, however, "higher-order" is intended to characterize the *content* of the state or process in question. A higher-order state is here understood to be a state that *represents* or is *about* another. So in the way we employ the terminology, the thought, *the left-hand line is longer*, is a first-order one, whereas, *I am uncertain whether the left-hand line is longer*, is higher-order.

An executive process isn't necessarily a metacognitive one, therefore. The mere fact that one cognitive process holds a controlling or modulating relation with respect to another doesn't make the first process *about* the second, and so doesn't make it metacognitive in nature, either. Attentional processes, for example, control and modulate perceptual ones, but no one thinks (nor should they think) that attention is a form of metacognition. Moreover, one process can monitor the output of another without in any sense representing it. For instance, the saliency attentional system monitors the contents of currently unattended perceptual and locally-activated memory representations, appraising the relevance of the contents thus represented, and competing for the resources of top-down attention when they are deemed relevant enough (Corbetta, Patel, & Shulman, 2008). But no one thinks that relevance-monitoring is metacognitive in nature. This is because what is monitored are the represented *contents* of the percepts and memories in question. There is nothing in the monitoring process that makes it *about* the mental states being monitored, as opposed to the worldly, represented, contents of those states. So monitoring, here, while executive, is entirely first-order-cognitive in nature.

We should emphasize that the root of our objections to the primate-metacognition literature is not that we think the data are better explained in mere associative terms. On the contrary, we fully accept that these (and many other) animals are capable of flexible executively-controlled decision making. Rather, our objections stem from what we take to be the best accounts of *human* decision making in tasks of the sort in question. For instance, in offering an account of so-called "uncertainty monitoring" in terms of first-order appraisals of risk, we rely on the now-extensive human literature on prospective reasoning and decision making (Gilbert & Wilson, 2005, 2007; Seligman, Railton, Baumeister, & Sripada, 2013). Risk appraisal is not thought to require metacognition.

We have, however, also been motivated by the following consideration. If metacognitive capacities were of ancient evolutionary provenance, reaching deep into the primate line, then one might expect humans to be a great deal better at metacognitive tasks than they actually are. Human metacognitive processing is generally only weakly accurate, at best (Dunlosky & Metcalfe, 2009; Fletcher & Carruthers, 2012). Moreover, it is late to develop in childhood, varies by culture, and seems to depend on idiosyncratically invented strategies as well as explicit instruction (Carr, Kurtz, Schneider, Turner, & Borkowski, 1989; Güss & Wiley, 2007; Shrager & Siegler, 1998). In contrast, human "mindreading" or "theory of mind" abilities develop in early infancy (Buttelmann, Carpenter, & Tomasello, 2009; Scott & Baillargeon, 2017; Southgate, Chevallier, & Csibra, 2010), in ways that are seemingly universal (Barrett et al., 2013), and form an essential foundation for much of human social life. This encourages us to think that metacognition actually results from self-directed mindreading, as one of us has argued at length elsewhere (Carruthers, 2011).

An adaptation can be present in an ancestral creature (monkeys) without undergoing further development or improvement in subsequent species (humans), of course. But note the importance of

learning and reasoning in human life-ways. Humans are hypersocial creatures. Almost everything we know is learned from others, and much of we do has to be negotiated collectively with others. Against this background one might think that if our ancestors really did have nascent innate capacities to improve learning and decision making through metacognitive monitoring and intervention, then the capacities in question would have come under intense selection pressure to become swifter, more reliable, and more sophisticated. That this seems *not* to have happened provides one initial consideration against the reality of monkey metacognition.

One might object against this argument that humans are unique in possessing explicit, exact, mathematical abilities, too; which are likewise slow to develop and dependent on explicit instruction. Yet no one thinks this consideration counts against the numerical abilities of an extensive range of nonhuman animals. Indeed, it is now widely accepted that animals can represent and reason about the approximate numerosity of sets (Feigenson, Dehaene, & Spelke, 2004). The difference in this case, however, is that exact-number cognition was a late cultural invention, probably only emerging a few thousand years ago along with market economies and fixed human settlements (Dehaene, Izard, Spelke, & Pica, 2008; Frank, Everett, Fedorenko, & Gibson, 2008). There has thus been little time for the pre-existing numerosity abilities shared with other animals to come under direct evolutionary selection. In contrast, humans (and probably other hominin species, too) have *always* been heavily dependent on learning from and negotiation with others.

Uncertainty Monitoring

This experimental paradigm has been extensively employed, especially with macaque monkeys, using a variety of smart variations designed to rule out associative explanations of the data (Couchman, Beran, Coutinho, Boomer, & Smith, 2012; Smith et al., 2008, 2014). The basic setup, however, is that the animals are faced with a primary discrimination-task of some sort (such as the density of pixels in a stimulus or the comparative length of two lines). Correct discrimination issues in a reward, whereas an incorrect choice issues in the mild penalty of a “time out” before the next trial can begin. All the while, however, the animals have the option of selecting an “opt out” symbol, which immediately advances the participant to the next trial, thus avoiding a time out. The key finding is that most macaques make adaptive use of the opt-out option, selecting it with greater frequency as the psychophysical difficulty of the primary discrimination-task increases. It is said that they do this because they are aware of their own uncertainty.

One consideration often said to support a metacognitive interpretation of the data, is that when humans undertake the same tasks they not only display very similar response profiles to those of macaques, but they report that they opt out when they do *because they are aware of their uncertainty* (Couchman, Coutinho, Beran, & Smith, 2010; Smith, 2005). But this is weak evidence, at best. One reason is that humans are consummate and obligatory mindreaders. So when asked to provide a retrospective explanation of their behavior, they will correctly interpret it as having been guided by uncertainty. That they do so provides no reason to think that they were aware of their own uncertainty at the time they decided to opt out, however. And even if they *were* aware of it at the time of deciding, it is quite another matter to claim that they chose as they did *because of* metacognitive awareness of uncertainty, rather than because of first-order uncertainty itself. Indeed, we have known since Nisbett and Wilson (1977) that people are really quite poor at identifying the causal relationships among mental states—even among states they are aware of—such as knowing which among their thoughts and feelings caused them to act. So while we accept that people choose as they do because they are uncertain, awareness of uncertainty might not be necessary, or might not play any causal role in performance.

We think a better explanation of the data (both macaque and human) is one framed in terms of risk-based decision making. When humans make decisions non-discursively (without engaging in verbally-based reasoning) they do so through *prospection*. They represent the actions available to them, anticipate their likely outcomes, and respond affectively to the represented actions and outcomes (Gilbert & Wilson, 2005, 2007; Seligman et al., 2013). Faced with a difficult discrimination, one will foresee that

neither of the primary response options is likely to issue in success—anticipating, instead, a boring (and reward-free) timeout. In contrast, the opt-out option is known to move one straight to the next trial, and so will be appraised as preferable by comparison. We suggest that macaques, too, are appraising the riskiness of a difficult choice, and then act adaptively to avoid unwelcome consequences. Nothing metacognitive is required.

One initial piece of evidence consistent with this suggestion comes from recent findings with capuchin monkeys. It had initially been found that capuchins, unlike macaques, do *not* make adaptive use of the opt-out option (Beran, Smith, Coutinho, Couchman, & Boomer, 2009), and this was interpreted as marking a divide between primates capable of metacognition and those that are not. But Carruthers and Ritchie (2012) offered a different interpretation, drawn from the perspective of their risk-based decision-making account. Reflecting on the differences between capuchins and macaques in foraging behavior (leaf-and-fruit eating versus generalist) and other factors, they suggested that capuchins may just be less sensitive to risk. This has now been confirmed. Beran, Perdue, Church, and Smith (2016) find that when capuchins are confronted with a much riskier version of the task (requiring them to make a discrimination among six items rather than just two) they *do* start to make adaptive use of the opt-out option. (Note that in the original paradigm, animals who adopt a guessing strategy have a 50% chance of receiving a reward anyway; in this new paradigm, the chances of success with such a strategy drop to below 20%.)

Informed by considerations like those sketched above, Nicholson, Williams, Grainger, Lind, and Carruthers (2019) set out to test whether uncertainty-monitoring tasks of the sort conducted with monkeys are genuinely metacognitive in nature, while at the same time testing (against a pair of alternatives) their view that metacognition results from self-directed mindreading. They tested matched groups of neurotypical people and adults with autism spectrum disorder (ASD), using both an implicit uncertainty-monitoring task modeled closely on the ones used with animals, as well as an explicit judgment-of-confidence task using the same sorts of materials. Judgment-of-confidence tasks require participants to give verbal reports of certainty about their knowledge and are de facto measures of metacognitive ability used widely in the human literature on metacognition. Nicholson and colleagues reasoned that if implicit uncertainty monitoring is genuinely metacognitive in nature, then these two tasks should share some resources, and people's performance on them should thus be correlated. This should be true even if the implicit task is nonconceptual or does not represent uncertainty as such, provided that the kind of metacognitive monitoring involved provides an input for people's judgments in the explicit judgment-of-confidence task. But this is not what Nicholson and colleagues found. On the contrary, the correlation between performance on the two tasks was nearly zero in both groups of participants.

In addition, and as they had predicted, Nicholson et al. (2019) found that ASD performance on the implicit task was no different from that of neurotypical controls, whereas their performance on the explicit judgment-of-confidence task was impaired. This is just what would be predicted if judgments of confidence employ some of the same resources as third-person mindreading, given that people with ASD are known to be impaired in the latter (and as Nicholson and colleagues further confirmed in their sample). Moreover, performance on the implicit task was uncorrelated with measures of mindreading ability, whereas mindreading accounted for a significant amount of variance on the judgment-of-confidence task.

Of course, these data fail to establish conclusively that uncertainty-monitoring tasks do not require metacognitive resources of any kind. For it might be that implicit and explicit forms of metacognition are two wholly separate capacities, one of which involves self-directed mindreading and the other of which does not. But there now seems little to motivate such a view. Moreover, taking this position would require giving up on the claim that the abilities manifested in monkeys are a first step on the road to explicit human metacognition and self-awareness. And in addition, the probative value of people's explicit metacognitive judgments about what they are doing when they undertake an implicit uncertainty-monitoring task appears to have been decisively undercut.

Overall, Nicholson et al.'s (2019) findings support the view that explicit forms of metacognition employ some of the same cognitive resources as mindreading, and that implicit tasks of the sort conducted with monkeys are not genuinely metacognitive in nature. Rather, the best account of the latter

tasks is one framed in terms of first-order forms of risk-based decision making, rather than metacognitive monitoring.

Search Behavior

Monkeys, when ignorant, will search for the information they need to achieve their goals, and this is said to display metacognitive awareness of their own lack of knowledge (Beran & Smith, 2011; Hampton et al., 2004; Kornell, Son, & Terrace, 2007; Rosati & Santos, 2016). For instance, if they know that food is hidden in one of two locations, but do not know which, they will search those locations before committing to a choice. Humans completing such a task would explain that they looked first because they did not know where the target was but wanted or needed to know. Note that the desire cited here is a metacognitive one (to obtain knowledge), grounded in a metacognitive belief (that one lacks knowledge). Parsimony might suggest that the same explanation should be given for the monkeys' behavior, too.

Nevertheless, Carruthers (2018, 2020) makes the case that these explanations (both human and animal) are in error—albeit an error that is deeply embedded in common-sense psychology. As noted earlier, mindreading abilities are early to develop in infancy, and seem likely to rely on a small number of conceptual primitives, including THINK, WANT, SEE, HEAR, and TELL (Baillargeon, Scott, & Bian, 2016; Carruthers, 2013). But it seems that neither CURIOSITY nor a concept for the motivation underlying instrumental search (which lacks a lexicalized name in English) are among them. Those concepts therefore need to be constructed in development out of primitive components, perhaps as WANT TO THINK. This would explain why (until recently) all philosophers and cognitive scientists who have written on the topic of curiosity have addressed it in metacognitive terms—as involving a desire for knowledge or true belief, or as an intrinsic motivation to learn, or something of the sort (see Foley, 1987; Goldman, 1999; Williamson, 2000, among philosophers; and see Blanchard, Hayden, & Bromberg-Martin, 2015; Gruber, Gelman, & Ranganath, 2014; Kidd & Hayden, 2015; Litman, 2005, among psychologists.) Even Loewenstein's (1994) well-known “information gap” theory of curiosity, which *sounds* as if it might not require metacognition, is actually framed in metacognitive terms. Curiosity is said to arise from “a discrepancy between what one knows and what one *wishes to know*” (p. 93, emphasis added).

Recently, however, a small group of philosophers has argued that curiosity is best analyzed as one of a number of first-order *questioning attitudes* (Carruthers, 2018; Friedman, 2013; Whitcomb, 2010). As expounded by Carruthers (2018), curiosity is an affective emotion-like motivating state that takes a question for its content rather than a proposition. Someone who is curious about a new couple who have moved into the neighborhood is in a state with a content such as, *who they are*, or *what they are like*. As with other affective states, curiosity directly activates motor plans that have been shaped by evolution and/or learning to issue in satisfaction of the attitude, such as approaching to look, asking verbal questions, and more. These motor-activations provide an impetus to act that will need to be executively suppressed if the action is not to be completed. (Compare the impetus to run when afraid.) But curiosity can also motivate forms of executively-controlled planning, leading one to seek ways of achieving an answer to the embedded question. (Again, compare how someone who is afraid may evaluate a number of potential means of avoiding the object of fear.)

Both instrumental search and curiosity-driven behavior are said to be motivated by affective states with contents like, *where the food is*, *what that is*, *what will happen next*, and so on. So long as these states remain active, they normally continue to motivate behavior until a matching—question-answering—belief is acquired (*the food is in the box*, *that is a predator*, *a fight happened*, and so on). Notice that in some cases no overt behavior may be required beyond the maintenance of attention. This is the emotion of *interest*, which is a questioning attitude with a content such as, *what will happen*. Someone who is interested in something has their attention held by events happening in full view (requiring no overt investigative behavior) until the underlying question gets answered, or until some other motivation wins out for the control of behavior.

Importantly for our purposes, questioning attitudes, like other affective states, are sensitive to background *appraisals*, which provide the conditions under which they become activated. But in this case

the background condition is an *absence* or a *lack*. Curiosity, for example, can be triggered by a failure of recognition, motivating one to approach closer, look more closely, and so on. Likewise, the emotion underlying instrumental search is triggered by a failure to access information needed to achieve some goal (but not under that—metacognitive—description). For example, a hungry food-caching bird may search its memory for the location of a stash. If it fails to come up with anything, that will prompt a question, *where food is*, which will motivate the animal to begin a foraging search. And a monkey that has seen food hidden in one of two potential locations but without seeing which, will be prompted into a questioning state with the content, *where the food is*, or perhaps the content, *whether the food is in A or in B*. That state then motivates behavior designed to issue in an answer to the question-content, such as acquiring the belief, *the food is in B*.

Note that *interest* (the onset of which often manifests as *surprise*) is frequently listed among the basic emotions, common to many other creatures besides ourselves (Ekman, 1992; Izard, 1977, 2007). Moreover, Kidd and Hayden (2015) argue that interest and attentional search should both be thought of as forms of curiosity, showing that they are widespread across the animal kingdom. They also review evidence that the same valuational circuits underlie each.

In addition, in the course of his extensive neuropsychological work with animals (mostly rats and cats), Panksepp (1998) identifies one of the basic mammalian motivational states as one of SEEKING. (Panksepp uses capital letters when referring to the emotion-kinds he investigates in order to emphasize that they might not correspond neatly to any folk-psychological kinds.) This is the state that motivates an animal to explore its environment but can also be harnessed in the service of other motivations, including mate-search and foraging-search. (Compare the way in which fear can be harnessed not just for avoidance of things that are intrinsically threatening, but also instrumentally, motivating avoidance of things that stand in the way of important goals.) Especially striking, Panksepp reports that when humans have received direct electrical stimulation of putative components of the SEEKING brain-network, the most common introspective report they offer is one of intense *interest*—including feelings of alertness, excitement, and the sense that something very interesting is happening or about to happen.

Few comparative researchers have been tempted to claim that either rats or cats are capable of metacognition (but see Kirk, McMillan, & Roberts, 2014; Templer, Lee, & Preston, 2017). So, a first-order explanation in terms of questioning attitudes seems preferable. Moreover, we know that pigeons, as well as monkeys, will opt to give up a significant portion of their eventual reward in order to learn whether a reward is or is not coming shortly thereafter (Bromberg-Martin & Hikosaka, 2009; Gipson et al., 2009). These animals will select an option that reliably signals whether a food-item is coming a few seconds later, even though this choice has no impact on the likelihood of the reward itself, and even though that choice always results in a smaller reward, if there is one. (Compare how one might pay a premium in advance in order to learn via express mail rather than regular mail whether one has won a lottery.) Moreover, we know that the reward systems in animals respond positively to the prospect of the informative option independently of their responses to the prospect of the food itself (Blanchard et al., 2015).

It is also worth noting that for more than thirty years now, researchers have employed *expectancy violation* paradigms to investigate human infants' "core knowledge" of the world around them (Baillargeon, Spelke, & Wasserman, 1985; Spelke & Kinzler, 2007). When an infant has had her expectations violated she will look longer at the object or event than when the latter was expected. Since no one has suggested that infants of this age are capable of metacognition, the behavior is best explained in terms of questioning attitudes directed at the environment, such as the question, *how that happened*. And indeed, infants don't just passively attend to expectancy-violating objects or events, but they preferentially learn from them, and if given the opportunity they will explore them in ways related to the nature of the initial expectation (Stahl & Feigenson, 2017). A 10-month-old infant shown an object apparently passing through a solid wall, for example, will try banging it on a surface when later given an opportunity to handle the object (seemingly asking whether it is solid); whereas the same infant shown an object that seems to remain in the air without support will thereafter repeatedly drop it (as if inquiring whether it can float). Likewise, Kidd, Piantadosi, and Aslin (2012, 2014) show that infants as young as

seven months will preferentially attend to stimuli that are neither already familiar nor too complex to comprehend, seemingly maximizing the extent to which their questioning attitudes are being satisfied.

Questioning attitudes make a good deal of sense in the light of evolution, of course. For almost all animals need to acquire information about the structure and layout of their environments. In the simplest cases, animals might just move through the environment at random, hoovering up information as they go. But most will need to direct targeted searches of the environment—not revisiting regions that have recently been explored, for example, and focusing most closely on those that have yielded rewards in the past. For instance, while a foraging bear might take a semi-random walk through its environment, its behavior at particular junctures seems likely to be motivated by questions that are informed by the bear's background knowledge, such as, *whether there are berries on that bush*, or, *what is under that log*. And the experiments cited earlier are sufficient to show that we need to postulate more than just hunger to explain the bear's behavior, since in the right circumstances curiosity and hunger can compete and get traded off against each other.

But if all mammals and birds (at least) have questioning, curiosity-like, motivational states, then why is it that capuchins fail at look-first-when-ignorant tasks, whereas macaques and apes succeed (Basile, Hampton, Suomi, & Murray, 2009; Beran, Smith, & Perdue, 2013)? And why is it that pigeons fail to seek the stimuli required for success in a match-to-sample task, whereas macaques succeed (Beran & Smith, 2011; Roberts et al., 2009)? These divergent findings may reflect nothing more than differences in cognitive style among species, not unlike the differences between individual humans in what psychologists call “need for cognition.” Capuchins have a problem-solving style of a try-everything-really-quickly sort, whereas macaques are more cautious and more reflective. (We have already noted the divergence between these two species in attitudes to risk.) This does not show that they lack attitudes that would motivate information seeking altogether. The same may be true of pigeons, who, as we noted earlier, *do* have curiosity-like motivations that can compete with foraging ones in the right circumstances.

We have been building a case that information-seeking experiments that have been claimed to show metacognition in monkeys do not really do so. Rather, what we see are the effects of first-order questioning attitudes. These are sensitive to salient forms of ignorance, but without representing ignorance. How might this be tested directly, however? One possibility would be to conduct dual-task studies with humans. If information seeking is genuinely metacognitive, then it should be interfered with by a concurrent metacognitive task. One might imagine experiments in which people need to seek information (or not) to succeed in the task, where this could be combined either with a concurrent explicit judgment-of-confidence task (in one group of participants) or with some other control-task matched for first-order difficulty (in another). Our prediction would be that there would be no difference between the two in terms of the extent to which they interfere with information seeking. Alternatively, one could just look for correlations between people's performance in information-seeking tasks and explicitly metacognitive ones that are equally demanding of working-memory resources. For if the former are genuinely metacognitive in nature, then they should share domain-specific resources with the latter.

Other Tasks

While our main focus here has been on uncertainty-monitoring tasks and information-seeking tasks, we should say something briefly before closing about other tasks that have been said to provide evidence of metacognitive abilities in animals.

Consider, first, the sorts of tests that are alleged to show that monkeys have metacognitive access to their own memories. In a match-to-sample task where the difficulty can be varied by altering the wait-time, monkeys will choose to opt out of taking the memory-test more often in cases where they have to wait longer, where remembering is more difficult (Hampton, 2001, 2005, 2009). It is said that they are *aware* that they no longer remember the original stimulus, which is why they opt out. Of course, it is true that the monkeys, just like humans, *might* have such metacognitive awareness. Like us, they might be aware of themselves as remembering or failing to remember something. But the data fail to support this interpretation.

Presumably the monkeys, like humans, solve the task by retaining in working memory a representation of the target stimulus. But for them (just as for humans) such a task can be hard, since distraction will result in that representation being lost. So the more time elapses between presentation and test, the more chance there is that the information in working memory has become fragmentary and incomplete, or lost altogether. The monkey knows that its task is to touch the item on the test screen that matches the one presented earlier. If it retains in working memory a representation of that item, then it should appraise the chances of success in the test to be good, and therefore opt to take it. In contrast, if the monkey no longer has a working-memory representation at all, or only has one that is significantly degraded, then it should experience the opt-out option as the most positive and so select that, rather than opt to take the test. Success in this task requires the monkey to *have* a memory, not to *monitor* its memory. And opting out adaptively requires the monkey to respond appropriately if it *lacks* any specific memory of the stimulus. This does not require it to be aware of itself *as* lacking a memory.

Might these ways of using memory amount to a form of *implicit* meta-memory? By acting in one way in the presence of a memory, and in another in its absence, does the animal display some sort of procedural meta-memory? It is hard to discern any grounds for saying so. For that one is disposed to act in one way if one has a memory and in another if one does not is just what it *is* to have or lack a memory. This just describes the normal first-order causal role of memory in the cognitive and decision-making processes of creatures that possess memory-states. There need be nothing *meta* about it.

Finally, consider the “go when you know” tasks administered to chimpanzees by Beran et al. (2015). When the animals had completed a delayed match-to-sample task correctly, they moved to a hopper to collect a briefly-available reward, whereas when they had answered incorrectly they did not. This is said to show that they were aware of their own levels of confidence in their responses. But this, too, seems unwarranted. The animals know that in order to be eligible receive a reward, they need to touch the picture on the screen that matches the one that had been there a few seconds previously. So all that is required is a first-order judgment of the form: *that one* [the one that has just been touched] *was on the screen previously*. Although grounded in a memory of the previous sample, no doubt, this judgment only requires the animals to *have* a memory, not to metacognize it as such.

We suggest that human adults and children will solve the task in the same way: if one still has in working memory an image of one of the pictures now presented, one will touch that and move to retrieve the food; otherwise not. A human might well report that she went to the hopper because she knew she had got it right. And this explanation is correct, if “getting it right” means “touching the picture that was the same as the sample.” But if the human says that she acted as she did because she was *aware of* having got it right, then she goes beyond what she is in a position to know, we submit. Even if she was so aware, her action might have had a simpler first-order cause; one that is shared with other animals.

Conclusion

In our view, many comparative psychologists have been too quick to interpret their data as evidence of metacognition in animals. In part this may result from a bias in favor of neo-Cartesian—first-person-first—accounts of meta-representational capacities generally (Carruthers, Fletcher, & Ritchie, 2012). But it may also result from too great a readiness to accept what humans are inclined to say about the causes of their own behavior when participating in similar tasks. In our view, metacognition plays a far smaller role in adult human cognition than people are intuitively inclined to think. We are misled in part because, as consummate mindreaders, we can readily construct explanations for our own behavior that seem metacognitive in nature when combined with the assumption that we have transparent access to our own minds. But we are also misled because the mindreading faculty itself mis-conceptualizes some of the mental-state-kinds in its domain (curiosity and interest in particular). Thus, in our view, comparative psychology needs to pay closer attention to the best theories of *human* psychology.

References

- Baillargeon, R., Scott, R., & Bian, L. (2016). Psychological reasoning in infancy. *Annual Review of Psychology*, *67*, 159–186.
- Baillargeon, R., Spelke, E., & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition*, *20*, 191–208.
- Barrett, H., Broesch, T., Scott, R., He, Z., Baillargeon, R., Wu, D.,...Laurence, S. (2013). Early false-belief understanding in traditional non-western societies. *Proceedings of the Royal Society B: Biological Sciences*, *280*, 20122654.
- Basile, B., Hampton, R., Suomi, S., & Murray, E. (2009). An assessment of memory awareness in tufted capuchin monkeys (*Cebus apella*). *Animal Cognition*, *12*, 169–180.
- Beran, M., Perdue, B., Church, B., & Smith, J. D. (2016). Capuchin monkeys (*Cebus apella*) modulate their use of an uncertainty response depending on risk. *Journal of Experimental Psychology: Animal Learning and Cognition*, *42*, 32–43.
- Beran, M., Perdue, B., Futch, S., Smith, J. D., Evans, T., & Parrish, A. (2015). Go when you know: Chimpanzees' confidence movements reflect their responses in a computerized memory task. *Cognition*, *142*, 236–246.
- Beran, M., & Smith, J. D. (2011). Information seeking by rhesus monkeys (*Macaca mulatta*) and capuchin monkeys (*Cebus apella*). *Cognition*, *120*, 90–105.
- Beran, M., Smith, J. D., Coutinho, M., Couchman, J., & Boomer, J. (2009). The psychological organization of "uncertainty" responses and "middle" responses: A dissociation in capuchin monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes*, *35*, 371–381.
- Beran, M., Smith, J. D., & Perdue, B. (2013). Language-trained chimpanzees (*Pan troglodytes*) name what they have seen but look first at what they have not seen. *Psychological Science*, *24*, 660–666.
- Blanchard, T., Hayden, B., & Bromberg-Martin, E. (2015). Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. *Neuron*, *85*, 602–614.
- Bromberg-Martin, E. & Hikosaka, O. (2009). Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron*, *63*, 119–126.
- Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, *112*, 337–342.
- Carr, M., Kurtz, B., Schneider, W., Turner, L., & Borkowski, J. (1989). Strategy acquisition and transfer among American and German children: Environmental influences on metacognitive development. *Developmental Psychology*, *25*, 765–771.
- Carruthers, P. (2011). *The opacity of mind*. Oxford: Oxford University Press.
- Carruthers, P. (2013). Mindreading in infancy. *Mind and Language*, *28*, 141–172.
- Carruthers, P. (2018). Basic questions. *Mind & Language*, *33*, 130–147.
- Carruthers, P., (in press). Questions in development. In L. Butler, S. Ronfard, & K. Corriveau (Eds.), *The questioning child: Insights from psychology and education*. Cambridge: Cambridge University Press.
- Carruthers, P., Fletcher, L., & Ritchie, J. B. (2012). The evolution of self-knowledge. *Philosophical Topics*, *40*, 13–37.
- Carruthers, P. & Ritchie, B. (2012). The emergence of metacognition: Affect and uncertainty in animals. In M. Beran, J. Brandl, J. Perner, & J. Proust (Eds.), *Foundations of metacognition* (pp. 76-93). Oxford: Oxford University Press.
- Corbetta, M., Patel, G., & Shulman, G. (2008). The reorienting system of the human brain: From environment to theory of mind. *Neuron*, *58*, 306–324.
- Couchman, J., Beran, M., Coutinho, M., Boomer, J., & Smith J. D. (2012). Evidence for animal metaminds. In M. Beran, J. Brandl, J. Perner, and J. Proust (Eds.), *Foundations of metacognition* (pp. 21-35). Oxford: Oxford University Press.
- Couchman, J., Coutinho, M., Beran, M., & Smith, J. D. (2010). Beyond stimulus cues and reinforcement signals: A new approach to animal metacognition. *Journal of Comparative Psychology*, *124*, 356–368.
- Dehaene, S., Izard, V., Spelke, E., & Pica, P. (2008). Log or linear? Distinct intuitions of the number scale in Western and Amazonian indigene cultures. *Science*, *320*, 1217–1220.
- Dunlosky, J. & Metcalfe, J. (2009). *Metacognition*. Washington, DC: Sage Publications.
- Ekman, P. (1992). Are there basic emotions? *Psychological Review*, *99*, 550–553.
- Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*, 307–

314.

- Flavell, J. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, *34*, 906–911.
- Fletcher, L., & Carruthers, P. (2012). Metacognition and reasoning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*, 1366–1378.
- Foley, R. (1987). *The theory of epistemic rationality*. Cambridge, MA: Harvard University Press.
- Frank, M., Everett, D., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition*, *108*, 819–824.
- Friedman, J. (2013). Question-directed attitudes. *Philosophical Perspectives*, *27*, 145–174.
- Gilbert, D., & Wilson, T. (2005). Affective forecasting: Knowing what to want. *Current Directions in Psychological Science*, *14*, 131–134.
- Gilbert, D., & Wilson, T. (2007). Prospection: Experiencing the future. *Science*, *317*, 1351–1354.
- Gipson, C., Alessandri, J., Millher, H., & Zentall, T. (2009). Preference for 50% reinforcement over 75% reinforcement by pigeons. *Learning and Behavior*, *37*, 289–298.
- Goldman, A. (1999). *Knowledge in a social world*. Oxford: Oxford University Press.
- Gruber, M., Gelman, B., & Ranganath, C. (2014). States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron*, *84*, 486–496.
- Güss, C., & Wiley, B. (2007). Metacognition of problem-solving strategies in Brazil, India, and the United States. *Journal of Cognition and Culture*, *7*, 1–25.
- Hampton, R. (2001). Rhesus monkeys know when they remember. *Proceedings of the National Academy of Sciences*, *98*, 5359–5362.
- Hampton, R. (2005). Can rhesus monkeys discriminate between remembering and forgetting? In H. Terrace & J. Metcalfe (Eds.), *The missing link in cognition* (pp. 272–295). Oxford: Oxford University Press.
- Hampton, R. (2009). Multiple demonstrations of metacognition in nonhumans: Converging evidence or multiple mechanisms? *Comparative Cognition and Behavior Reviews*, *4*, 17–28.
- Hampton, R., Zivin, A., & Murray E. (2004). Rhesus monkeys (*Macaca mulatta*) discriminate between knowing and not knowing and collect information as needed before acting. *Animal Cognition*, *7*, 239–246.
- Hanus, D., Mendes, N., Tennie, C., & Call, J. (2011). Comparing the performances of apes (*Gorilla gorilla*, *pan troglodytes*, *Pongo pymaeus*) and human children (*Homo sapiens*) in the floating peanut task. *PLoS One*, *6*, e19555.
- Izard, C. 1977: *Human emotions*. New York, NY: Plenum Press.
- Izard, C. 2007: Basic emotions, natural kinds, emotion schemas, and a new paradigm. *Perspectives on Psychological Science*, *2*, 260–80.
- Kidd, C., & Hayden, B. (2015). The psychology and neuroscience of curiosity. *Neuron*, *88*, 449–460.
- Kidd, C., Piantadosi, S., & Aslin, R. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS ONE*, *7*, e36399.
- Kidd, C., Piantadosi, S., & Aslin, R. (2014). The Goldilocks effect in infant auditory attention. *Child Development*, *85*, 1795–1804.
- Kirk, C., McMillan, N., & Roberts, W. (2014). Rats respond for information: Metacognition in a rodent? *Journal of Experimental Psychology: Animal Learning and Cognition*, *40*, 249–259.
- Kornell, N., Son, L., & Terrace, H. (2007). Transfer of metacognitive skills and hint seeking in monkeys. *Psychological Science*, *18*, 64–71.
- Litman, J. (2005). Curiosity and the pleasures of learning: Wanting and liking new information. *Cognition and Emotion*, *19*, 793–814.
- Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, *116*, 75–98.
- Nelson, T., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *The psychology of learning and information* (Vol. 26). New York: Academic Press.
- Nicholson, T., Williams, D.M., Grainger, C., Lind, S., & Carruthers, P. (2019). Relationships between implicit and explicit uncertainty monitoring and mindreading: Evidence from autism spectrum disorder. *Consciousness and Cognition*, *70*, 11–24.
- Nisbett, R., & Wilson, T. (1977). Telling more than we can know. *Psychological Review*, *84*, 231–295.
- Panksepp, J. (1998). *Affective neuroscience*. Oxford: Oxford University Press.
- Roberts, W., Feeney, M., McMillan, N., MacPherson, K., Musolino, E., & Petter, M. (2009). Do pigeons (*Columba*

- livia*) study for a test? *Journal of Experimental Psychology: Animal Behavior Processes*, 35, 129–142.
- Rosati, A., & Santos, L. (2016). Spontaneous metacognition in rhesus monkeys. *Psychological Science*, 27, 1181–1191.
- Scott, R., & Baillargeon, R. (2017). Early false-belief understanding. *Trends in Cognitive Sciences*, 21, 237–249.
- Seligman, M., Railton, P., Baumeister, R., & Sripada, C. (2013). Navigating into the future or driven by the past. *Perspectives on Psychological Science*, 8, 119–141.
- Shrager, J., & Siegler, R. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science*, 9, 405–410.
- Smith, J. D. (2005). Studies of uncertainty monitoring and metacognition in animals and humans. In H. Terrace & J. Metcalfe (Eds.), *The missing link in cognition* (pp. 242–271). Oxford: Oxford University Press.
- Smith, J. D., Beran, M., Couchman, J., & Coutinho, M. (2008). The comparative study of metacognition: Sharper paradigms, safer inferences. *Psychonomic Bulletin & Review*, 15, 679–691.
- Smith, J. D., Couchman, J., & Beran, M. (2014). Animal metacognition: A tale of two comparative psychologies. *Journal of Comparative Psychology*, 128, 115–131.
- Smith, J. D., Shields, W., & Washburn, D. (2003). The comparative psychology of uncertainty monitoring and meta-cognition. *Behavioral and Brain Sciences*, 26, 317–373.
- Southgate, V., Chevallier, C., & Csibra, G. (2010). Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Developmental Science*, 13, 907–912.
- Spelke, E., & Kinzler, K. (2007). Core knowledge. *Developmental Science*, 10, 89–96.
- Stahl, A., & Feigenson, L. (2017). Expectancy violations promote learning in young children. *Cognition*, 163, 1–14.
- Taylor, A., Elliffe, D., Hunt, G., & Gray, R. (2010). Complex cognition and behavioral innovation in New Caledonian crows. *Proceedings of the Royal Society B: Biological Sciences*, 277, 2637–2643.
- Templer, V., Lee, K., & Preston, A. (2017). Rats know when they remember: Transfer of metacognitive responding across odor-based delayed match-to-sample tests. *Animal Cognition*, 20, 891–906.
- von Bayern, A., Danel, S., Auersperg, A., Mioduszewska, B., & Kacelnik, A. (2018). Compound tool construction by New Caledonian crows. *Nature Scientific Reports*, 8, 15676.
- Whitcomb, D. (2010). Curiosity was framed. *Philosophy and Phenomenological Research*, 81, 664–687.
- Williamson, T. (2000). *Knowledge and its limits*. Oxford: Oxford University Press.