



## ARTICLE

<https://doi.org/10.1057/s41599-018-0200-y>

OPEN

# Cumulative culture and explicit metacognition: a review of theories, evidence and key predictions

Juliet Dunstone <sup>1</sup> & Christine A. Caldwell<sup>1</sup>

**ABSTRACT** A variety of different proposals have attempted to explain the apparent uniqueness of human cumulative culture as a consequence of underlying mechanisms that are also assumed to be uniquely well-developed in humans. Recently, Heyes and colleagues have proposed explicit (or Type 2) metacognition as a key feature of human cognition that might enable cumulative culture. In the current review we examine these arguments, and consider their plausibility. Firstly we consider whether distinctions between cognitive processes described as explicit/implicit, and Type 1/2 (or Systems 1/2), do indeed capture features that distinguish processes specific to human cognition, versus those that are shared with other species. In particular we consider whether this applies to distinctions relating to metacognitive processes. We also consider the ways in which explicit metacognitive processing might plausibly facilitate cumulative culture. We categorise the potential benefits as either optimising receiver behaviour, or optimising sender behaviour. Within both of these categories benefits could arise as a consequence of more effective representation of either one's own knowledge state, or that of others. We evaluate the current state of evidence supporting each of these potential benefits. We conclude by proposing methodological approaches that could be used to directly test the theory, and also identify which (if any) of the possible causal mechanisms may be implicated.

<sup>1</sup>University of Stirling, Stirling, UK. Correspondence and requests for materials should be addressed to J.D. (email: [juliet.dunstone1@stir.ac.uk](mailto:juliet.dunstone1@stir.ac.uk))

## Introduction

In the current article, we examine the view, recently proposed by Heyes (2016) and Shea et al. (2014), that distinctively human cultural evolution is attributable to capacities for explicit (or Type/System 2) metacognition. Essentially, these accounts argue that an ability to explicitly reflect upon states of knowledge, ignorance, and uncertainty, can fundamentally change the ways we use and share social information, and that these particular processes account for the characteristic forward progress of human culture. In the current review, we aim to evaluate these accounts, considering the evidence for their underlying assumptions, as well as the plausibility of mechanistic routes which could potentially link individual-level cognitive processes of explicit metacognition, with population-level outcomes resembling cumulative culture.

## Explanations of the distinctiveness of human cumulative culture

Cumulative cultural evolution is the process by which cultural traits (including behaviours, artefacts and tools) change over multiple episodes of social transmission to become more effective and beneficial to their users (Mesoudi and Thornton, 2018; Caldwell, 2018). In humans, this can lead to cultural traits evolving over many generations which could not have been invented by a single individual. Examples of human cumulative culture span a wide range of domains including abstract conceptual skills such as the cognitive tools provided by mathematical notation and operations (e.g. Bender and Beller, 2014), or survival skills such as lengthy food processing techniques that remove invisible toxins from raw ingredients (e.g. preparation of cycas seeds, Beck, 1992).

Many species have been shown to exhibit cultural traditions in the form of behavioural variation between populations which appears to be maintained by social learning, or evidence of the social diffusion of particular behavioural variants (Whiten et al., 2016). To highlight just a few examples, such evidence has been identified in chimpanzees (Hobaiter et al., 2014), humpback whales (Allen et al., 2013), and great tits (Aplin et al., 2015). Therefore, neither the capacity for culture, nor even the process of cultural evolution across generations (i.e. in the broader sense of any kind of incremental change arising as a consequence of social transmission, in the absence of improving functionality), is restricted to humans. In contrast, the cumulative improvement of traits across generations (sometimes referred to as the cultural “ratchet effect”, e.g. (Tomasello, 1990), is widely regarded to be unique to humans (Dean et al., 2014; Tennie et al., 2009).

Several previous theories have been proposed for why humans have an apparently unique capacity for cultural ratcheting. The most prominent view concerns particular learning mechanisms also proposed to be human unique. Humans are extremely adept at copying others’ behaviour (e.g. dubbed “*Homo imitans*” by Meltzoff, 1988) and seem to have a particular talent (or even a compulsion in some situations, Heyes, 2011) to do so. Conversely, non-human animals (henceforth animals) do not seem to exhibit the same proclivities to anything like the same degree. While there is evidence for action imitation in some species (e.g. marmosets: Voelkl and Huber, 2000), this occurs at much lower levels of accuracy, and in far more restricted contexts than in humans who are highly accurate copiers in domain general contexts. This has prompted theorists to propose that capacities for imitation and social learning may have represented critical cognitive developments in human evolution, allowing for cumulative culture (Lewis and Laland, 2012; Tomasello, 1999). Converging lines of evidence from computer models, and tournaments in which alternative strategies compete in simulated mixed populations

(Rendell et al., 2010) have found that the most successful strategies involve high rates of copying, and that model populations predominantly comprising social learners (simulated agents that “acquire a behavior performed by another individual, whether by observation of or interaction with that individual” (Rendell et al., 2010)) are more successful and develop more complex technologies.

There are however a number of reasons to question the notion that particular social learning mechanisms may account for human cumulative culture. Firstly, there is now mounting evidence of imitative abilities in other species (e.g. apes: Whiten et al., 2004, although see Call et al., 2005 for evidence counter to these claims) Secondly, in humans, experimental evidence has shown that cumulative culture can arise even when participants are restricted from observing others’ actions, such that these learners are forced to rely on non-imitative processes such as emulation of end products (Caldwell and Millen 2009). In addition, although cumulative culture necessarily involves social transmission as a mechanism for trait heritability, it is important to note that social learning alone cannot account for the increases in trait functionality that exemplify the process. The development of new technologies and behaviours also depends on innovation (Enquist et al., 2008; Lehmann et al., 2010). While copying error or accidental discovery, as well as intentional invention, can be sources of innovations (Caldwell et al., 2016; Henrich et al., 2008), it is clear that increased abilities for high-fidelity copying alone cannot explain cumulative increases in cultural behaviours and artefacts.

These considerations imply that cumulative culture may be explained not only by the mechanisms available to learners, but also the contexts in which these are employed, since it is the selective and strategic use of copying that accounts for its adaptiveness; high fidelity copying may be necessary for cumulative culture to emerge, but it is not sufficient. Indeed, models show that populations of flexible learners that can switch between social and individual learning at critical points outperform populations composed of only social learners, or only individual learners (Ehn and Laland, 2012; Enquist et al., 2007; Rendell et al., 2010).

Laland (2004) described a number of potential “Social Learning Strategies” (SLS) which could reflect adaptive rules regarding when to engage in social learning, and who to learn from. These included strategies such as: Copy when Uncertain, Copy the Majority, and Copy Successful Individuals. However, in spite of the fact that such strategies clearly have some potential to explain the selective retention of beneficial traits in learner populations, they nonetheless fail to provide an adequate explanation for the fact that cumulative culture appears to be restricted to humans. This is because a wide range of animals have also been shown to exhibit SLS. This includes social insects (Smolla et al., 2016), fish (Pike and Laland 2010) and bats (Jones et al., 2013).

However, Heyes (2018b) has drawn a distinction between social learning strategies that are based on “planetary” decision rules with those based on “cook-like” decision rules. Planetary SLS, like laws of planetary motion, capture regularities within the observable behaviour of the entities of interest, but the rules are only in the minds of those doing the describing. In contrast, “cook-like” SLS are more akin to the decision rules used by a cook following a recipe, i.e. they are explicitly represented within the mind of the agent. Heyes (2016, 2018b) has thus argued that it is these *explicitly metacognitive* SLS that account for the elaborate outcomes of human cumulative culture. Although there is no dissent regarding the fact that animal social learning strategies demonstrate adaptive flexibility, these “cook-like” SLS are assumed to permit a much higher degree of—insightful—flexibility, potentially optimising the effectiveness of social

transmission in a number of different ways (see Section 'How Might Explicit Metacognition Facilitate Cumulative Culture?').

The proposal that explicit metacognitive processes may set human social learning apart from that of other animals is compelling, and persuasive theoretical arguments in its favour can be found in Shea et al., (2014), Heyes (2016), and Heyes (2018b). In the current review we consider the evidence in support of this explanation (which we refer to here as the Explicitly Metacognitive Cumulative Culture hypothesis, the EMCC), over and above competing alternatives. Recent literature (Heyes, 2018a) elaborates further on this argument to claim that these metacognitive strategies are themselves products of cultural evolution, as well as processes supporting it. Evaluating this extension of the theory is beyond the scope of this current review however, and as such we do not discuss this argument further.

Metacognition and the EMCC as discussed here also encompass capacities for theory of mind, or mentalising. As argued by Carruthers (2009) metacognition and theory of mind are not wholly distinct capabilities, and metacognition about one's own mind may in fact rely upon prerequisite capacities for mind-reading about others' minds. Indeed, some arguments proposing the utility of explicit metacognitive social learning strategies encompass mentalising as much as introspection (e.g., Heyes, 2016, when asking "who knows?"), but see section 'Optimisation of receiver behaviour, due to understanding of others' knowledge states' for a more detailed discussion of this not uncontroversial distinction.

We begin below by examining a key assumption of the EMCC hypothesis, which is that explicit metacognitive processes are restricted to humans.

### Explicit metacognition as a uniquely human feature

The literature contains numerous claims of metacognitive ability in animals, across a broad range of species (e.g. monkeys: Smith et al., 2009; chimpanzees: Beran et al., 2015; dolphins: Smith et al., 1995; pigeons: Sole et al., 2003; rats: Foote and Crystal, 2007; Templer et al., 2017; and even bees: Perry and Barron, 2013). However, the EMCC rests on the assumption that the experimental paradigms used in these studies are assessing qualitatively different phenomena from the type of metacognition required for cumulative culture, which is assumed to be unique to humans. In this section we examine theories and evidence underlying the assumption that alternative methodologies in metacognition research may be evaluating fundamentally different cognitive processes, and that certain types of metacognition may indeed be manifested only in humans.

**Dual processing theories of cognition.** As noted above, a critical point in the EMCC is that only humans have conscious access to their social learning decision-making rules, whereas social learning in other animals is driven by automatic processes of which the agents themselves are unaware. It is this difference that is proposed to account for the unusual prevalence of cumulative culture in humans, compared with other species. It is important to note that the EMCC does not imply that *all* human cognition involves conscious access, or indeed that all examples of social learning in humans are based on explicit processes. Rather, the hypothesis draws on theories of human cognition which propose the existence of two systems, or two processing types. We present an overview of this body of literature below.

Theories of dual processes for various aspects of cognition have been relatively widespread since the 1970s (e.g. Wason and Evans, 1974). These theories state that there are two different modes of higher cognitive processing; one which is generally automatic, fast acting, non-conscious and based on associative mechanisms, and

one which is conscious, slower to act and rule-based (see Evans and Stanovich (2013), for a summary of attributes typically associated with each of the processing types). These two alternatives are generally referred to as either *Systems* (Systems 1 and 2), or *Process Types* (Type 1 and 2), to capture the automatic (1) and rule-based (2) cases, respectively. Although the idea of different *Process Types* offers a less theoretically loaded framework, which is potentially more consistent with a wider range of empirical evidence (e.g. reflecting a continuum, rather than a dichotomy, of alternative cognitive mechanisms), it is the *Systems* label which has been associated with the idea that there may be distinctively human modes of cognition ((Epstein, 1994); (Stanovich, 1999, 2004)). Accordingly, it is the *Systems* label that has been used in the literature relating dual process theories to human cultural evolution (Shea et al., 2014; Heyes, 2016; Heyes, 2018b). In the current review we use both terms.

In relation to the issue of human distinctiveness, some accounts hold that System 1 is phylogenetically ancient, and therefore shared with other animals, whereas System 2 is more recently evolved and likely to be unique to humans. Dual processing theories have been used as a framework for the interpretation of a diverse range of psychological phenomena, from decision making (Evans, 2007), learning (Dienes and Perner, 1998) and social cognition (Smith and DeCoster, 2000).

Evidence for dual processing comes from dual-task studies, and tasks which apply strict time pressures. This is because System 2 processes are argued to be taxing on executive functions and working memory capacity, as well as generally taking longer. Dual tasks are designed to put an additional load on finite cognitive capacities. If two tasks require the same cognitive mechanisms this creates a bottleneck in processing, resulting in delayed response or impaired performance in one or both tasks (Pashler, 1994). This means a dual task can detect what level of processing is being used to complete a task; if task performance is unimpeded by a concurrent working memory or executive function load it is likely to be an automatic (or System 1) process, whereas if working memory load significantly reduces speed or accuracy of responding it is likely to be System 2. Participants under a working memory load have been found to make more incorrect responses based on salient information rather than logical reasoning when completing conjunction fallacy problems, or logic puzzles such as the Wason Selection Task (De Neys, 2006). The application of a strict time pressure may also prevent the use of System 2, as it would not allow the longer processing time needed. This effect has also been found using the Wason Selection Task (Roberts and Newton, 2001).

There is also some neurological evidence for dual processing systems: distinct brain activations for using logic based (System 2) and belief based (System 1) solutions to problems have been found using fMRI (Goel and Dolan, 2003). Additionally, NIRS analysis found that areas implicated in incongruent reasoning trials were not activated when the same tasks were performed under additional cognitive load (Tsujii and Watanabe, 2009). McClure et al., (2004) found activation in the prefrontal cortex during reasoning about future monetary rewards but not during immediate decision making. This activation was found in a similar neurological region as areas associated with metacognition and executive functions.

Dual processing theories are by no means universally accepted; see Keren and Schul (2009) or Osman (2004) for some objections. However, critical accounts have generally focused on lack of precise definitions, or evidence of overlap between the proposed dichotomy of characteristics between System 1 and System 2, concluding that the two systems cannot be considered distinct and isolable. However, Evans and Stanovich (2013) have argued that this is merely a poor interpretation of the literature, and that

most of the features commonly described as differentiating Type 1 and 2 are just correlates typical of the processing types, and that these should not be expected to operate in a categorical, mutually-exclusive fashion.

Evans and Stanovich (2013) describe their theory of dual processing as a “default-interventionist” view (e.g. p227), meaning that the majority of cognition relies on System 1 processes unless the available response is incorrect or does not meet with the task goals, in which case System 2 will intervene.

**Metacognition and dual process distinctions.** It is central to the EMCC that it is the agents’ conscious access to their decision-making rules that allows human metacognitive processes to generate such highly adaptive outcomes from social learning (e.g. by allowing learners to seek out the most appropriate models dependent on their specific goal, Heyes (2018b), or by allowing those in possession of knowledge to broadcast their degree of confidence or uncertainty, as well as their choices, Shea et al., (2014)). It is perhaps unsurprising then that these theories have emphasised the importance of the System 1/System 2 distinction. However, in order to fully understand and evaluate these accounts, it is important to also consider metacognitive processes more generally, and then to turn to the question of how these relate to the dual process framework detailed above.

Metacognition, as originally defined by Flavell (1979), can be thought of as knowledge about one’s own cognitive processes, or cognition about cognition. Flavell split this into four separate components: knowledge (of your own cognitive abilities and of learning processes), experiences (current feelings of certainty or doubt), goals (objectives you have in order to achieve your current cognitive task) and actions (behaviours employed to achieve these objectives). Subsequent research has instead divided metacognition into declarative metacognitive knowledge (corresponding to Flavell’s knowledge) and procedural metacognition. Procedural metacognition has then been commonly divided into monitoring and control processes (Flavell’s experiences and actions, respectively) (Roebbers, 2017).

The term metacognition therefore encompasses a wide variety of phenomena. Accordingly, it has been used to describe a broad spectrum of findings identified in a wide range of contexts, from detecting tiny changes in perceptual stimuli (Deroy et al., 2016) to deliberately allotting revision time for exams or correcting errors in a piece of written text (Sannomiya and Ohtani, 2015). In addition, some authors consider metacognition to encompass understanding of others’ cognition, as well as one’s own, as part of a wider cognitive capacity for metarepresentation relating to mental states (Kuhn, 2000; Misailidi, 2010). In a large proportion of experimental research, metacognition has typically been operationalised as judgements of confidence in performance of an activity just completed (Judgements of Confidence; JOC), or ratings of prospective performance in an activity about to be completed (Feeling of Knowing; FOK) (Nelson and Narens, 1990)

Given the specific emphasis on explicit/System 2 metacognition within the EMCC, it is worth considering here examples of metacognitive phenomena which could be classified as implicit or System 1 processes, and how these differ from those assumed to implicate explicit or System 2 processes. Any paradigm involving direct report of degree of confidence, doubt or uncertainty, necessarily requires awareness of these states, and therefore would be classified as implicating explicit metacognition. However other, more indirect, methods have also been used as means of evaluating metacognition, particularly within animal studies.

The typical methodological paradigm used to solicit “metacognitive” behavioural responses in the absence of verbal report, involves offering an “opt-out” option within a decision making

task, which can be used adaptively by the participant to avoid the risks associated with particularly difficult trials. Adaptive use of the option to opt out is assumed to reflect the subject’s appreciation of their own uncertainty, and therefore metacognition. Such designs have been used to support claims of metacognitive ability in nonhuman primates (e.g. macaques: Smith et al., 2008). Alternative methodologies involve ‘information-seeking’ paradigms, where a participant can seek additional information before making a decision, which have also been used to support claims of metacognitive awareness in primates (e.g. Call and Carpenter, 2001).

However, these experiments remain contentious as demonstrations of true metacognitive ability, as adaptive performance could be explained by responses being driven by first-order states of anxiety elicited by the uncertainty of the situation, rather than second-order reflection on the state of uncertainty itself (Carruthers and Ritchie, 2012). This is explained most thoroughly by Carruthers (2008), who has argued that first-order beliefs, along with other basic mechanisms such as signal detection theory, are just as capable of explaining the findings. The account can be summarised thus: the participant is presented with two choices that carry equal valence (the animal is equally motivated to both choose and not choose either option). As soon as a third option (the opt-out or uncertain option) is presented this option automatically becomes the most attractive, especially given its reinforcement history of being associated with a small reward. As this explanation is simpler, in terms of the cognition required by the animal participants, Carruthers argues convincingly that this is a more parsimonious explanation than those ascribing metacognitive capacities to animals. A similar account from Hampton (2009) described how a range of studies claiming to show animal metacognition could be explained by environmental or behavioural cues, or direct competition between a choice to act, or make a “metacognitive” action. These accounts make it clear that although metacognitive introspection could in principle explain the results of the studies in question, the plausibility of such interpretations is seriously challenged by the availability of simpler explanations. Under the dual process framework, the animals’ performance in these studies would therefore likely be classified as System 1, or implicit, metacognition (e.g. Shea et al., 2014).

There are unavoidable challenges involved in establishing whether implicit and explicit metacognitive responses depend on different cognitive processes, especially if our ultimate motivation is to determine whether one is a distinctive feature of human cognition. Adult humans are necessarily capable of both, and we can only use non-verbal measures with animals due to the language requirements of direct assessments of explicit metacognition. However, patterns of emergence during human development potentially provide another source of evidence that could shed light on the relationship between implicit and explicit metacognitive behaviour, and whether implicit adaptive responding can occur in the absence of explicit competence.

Behavioural tests analogous to the “opt-out” paradigms used with animals (described above) have been used to demonstrate implicit metacognitive ability in very young children, from 20 months to around 5 years old. Behavioural tests of metacognitive competence have also included assessment of spontaneous information-seeking prior to committing to a response. These paradigms also potentially provide an insight into implicit reactions to the state of ignorance, without necessarily implicating metacognitive awareness of that state. These studies will not be described in detail here, as reports of implicit metacognitive measures are not directly relevant to the current review. However, please see Bernard et al., (2015), Goupil

et al., (2016) or Call and Carpenter (2001) amongst others for examples of the paradigms in question.

In spite of the early development of such behavioural responses to uncertainty (analogous to the evidence from animals), evidence of explicit metacognitive understanding only appears to emerge later. Although verbal reports can be readily obtained from preschool aged children, studies requiring them to verbalise their own state of knowledge nonetheless indicate that they have difficulty doing this and that when they do they show a pervasive bias towards overestimation of their own knowledge and performance (e.g. Rohwer et al., 2012).

The earliest examples of accurate performance based on explicit measures of metacognition come from children of around four to five years old. For example, Rohwer et al., (2012) found that only children older than five could provide reports about what they did not know in a partial exposure task. Cultice et al., (1983) also found accurate explicit metacognitive responding in children aged four and five years old, asked to name familiar individuals from their photograph. When children were unable to spontaneously recall the name themselves, they could respond with reasonable predictive accuracy to the question: “If I told you a lot of names, do you think you would know or remember which one was her name?”

The adaptive responses to uncertainty identified in children younger than four years old (including those aged one and two, e.g. Goupil et al., 2016, and Call and Carpenter, 2001) appears strikingly similar to the behaviour of animals in opt-out and information seeking studies. However, this kind of competence appears to precede the ability to provide explicit, accurate evaluations of states of knowledge, which apparently only develops some years later. This would therefore seem to corroborate accounts which propose that successful performance on the alternative task types is underpinned by different processes, and that the animal studies therefore do not provide evidence of explicit metacognition. This is consistent with the EMCC’s assumption that System 2 metacognitive capacities are specific to humans.

**How might explicit metacognition facilitate cumulative culture?** As an explanation for distinctively human cumulative culture, the EMCC rests on two fundamental assumptions. The first of these is the corresponding distinctiveness of explicit metacognition (as examined in the preceding section). The second of these is that the resulting reflective awareness of states of knowledge, ignorance and uncertainty (identified as the defining feature of explicit metacognition) offers significant benefits with regard to the optimisation of social information use, in ways that could explain the ratchet-like advances which distinguish human culture from the traditions of other species. Having considered the first of these premises in the preceding sections, we now turn to the second. What basis is there, either evidential or logical, for believing that explicit metacognition might enable cumulative culture? What are the potential routes by which this might occur? We hope that by clarifying the potential links between explicit metacognition and cumulative culture we can identify areas where evidence is lacking, with a view to informing future research efforts investigating the EMCC.

Explicit metacognition could potentially enable cumulative culture in a number of different ways. Below, we categorise the potential benefits as arising from receiver behaviour, or sender behaviour. Within both of these categories benefits could arise as a consequence of more effective representation of one’s own knowledge state, or that of others. It should be noted at this point that the existing accounts of the EMCC focus, respectively, on optimisation of sender behaviour due to understanding of own

knowledge state (Shea et al., 2014) and optimisation of receiver behaviour due to understanding of others’ knowledge states (Heyes, 2016), both detailed below.

*Optimisation of receiver behaviour, due to understanding of own knowledge state.* In much the same way that metacognitive awareness is assumed to facilitate academic performance (Dunlosky and Metcalfe, 2009) it is possible that it could similarly enable cumulative culture for reasons that are not inherently linked to how an agent understands or interacts with others. Awareness of one’s own knowledge state would allow learners to seek out new information when necessary, and recognise when updating their knowledge might be beneficial. This awareness might also be crucial when acquiring a new skill or knowledge is likely to require a protracted period of effortful practice before mastery is achieved. Essentially, we would predict that such awareness would result in social information being used in a much more optimal fashion than would otherwise be possible, encouraging highly strategic social information seeking, as well as direction of effort towards innovation when social information sources are judged to be inadequate. To our knowledge, the role of this kind of reflective awareness in directing one’s own learning has not been investigated within the social learning and cultural evolution literature. However, some authors have alluded more tangentially to the importance of self-focussed strategic effort in social learning. For example, Galef (2013) has stated: “in the case of skiing, there is no learning to do an act from seeing it done. Rather, there is learning by observation that an act is possible.... [A] novice can ... select from within her available repertoire of movements.... Then, over time, she can bring that first approximation into greater accord with the demonstrated act.” (p. 125). Galef (2013) also suggests that such learning may be particularly important for cumulative culture.

As noted above, this would be a route by which explicit metacognition might be critical to generating cumulative culture without the effects being restricted to *social* learning specifically. In the accounts of both Shea et al., (2014) and Heyes (2016), explicit metacognition is assumed to facilitate cumulative culture because it helps agents make inferences about others’ knowledge (Heyes, 2016) or provide information to others (Shea et al., 2014). It is perhaps not surprising that, in attempting to explain a phenomenon which itself certainly does depend on social learning, authors have focused on explanations which would specifically facilitate that type of learning over and above others. However, we would suggest that explicit metacognition might potentially facilitate efficient use of any kind of vicariously acquired information, as well as helping in any situation where habitual or automatic responses may need to be overridden due to the availability of up-to-date, or situation-specific, information which indicates that these are not appropriate. This account would be consistent with the assumptions of the “default interventionist” views of dual-process cognition described earlier, which posit that System 2 intervenes only when automatic System 1 processes are inadequate for the task in question. Thus, explicit decision rules, reasoning processes and learning strategies are likely to be intrinsically associated with situations where default responses (based on personal reinforcement history and/or genetically inherited behavioural biases) will be ineffective.

Although such situations will be by no means restricted to contexts involving social information use, the need to override default automatic and habitual responses may be a prevalent feature of these contexts. Consider a situation in which a new possibility becomes apparent to an agent through vicarious exposure to another’s behaviour; for example, the agent might observe that plentiful food resources, such as tubers, could be found underground. Taking full advantage of this new

information might necessitate an immediate switch in foraging strategy, overriding habitual responses which have been directly reinforced on multiple occasions. Although similar exposure to new information might occur outside of social contexts (e.g., if tubers were to be revealed as an incidental outcome of disturbance of the ground surface), the behaviour of others is perhaps particularly likely to provide information of immediate utility. Furthermore, once transgenerational accumulation of knowledge was in evidence, social sources would then effectively become repositories of particularly valuable information that might be otherwise hard to acquire. Therefore the benefits of this type of learning might be most apparent in social contexts, even though the learning mechanisms themselves would be general-purpose ones, not specifically adapted for use in social contexts.

Essentially, the suggestion here is that System 2 metacognition may be critical due to the high “executive function” demands of the type of social learning likely to be involved in cumulative culture. Overlap between the concepts of executive function and metacognition have been acknowledged in the existing literature (e.g. Roebers, 2017). Indeed, some research effort has already been targeted at the question of whether executive function limitations (specifically difficulties with inhibition) might explain the absence of cumulative culture in chimpanzees (e.g. Davis et al., 2016). We would see such an explanation as falling under the umbrella of the broader EMCC, within this particular category of optimisation of receiver behaviour due to understanding of own knowledge state.

*Optimisation of sender behaviour, due to understanding of own knowledge state.* Understanding of one’s own state of knowledge can also potentially facilitate cumulative culture by influencing sender behaviour, increasing the likely benefit to others of oneself as a source of social information. Access to one’s own level of confidence or uncertainty means that this information can be conveyed to others, alongside actual behavioural decisions. This would then allow others to make more strategic use of that social information, weighting information more heavily when a source reports confidence, or disregarding conflicting information when a source reports high levels of uncertainty. It is this aspect of the EMCC that forms the focus of Shea et al.’s, (2014) argument. There is some experimental evidence suggesting that this kind of metacognitive communication does indeed improve the efficacy of social information use. For example Bahrami et al., (2010) studied pairs of participants completing a low-level perceptual decision-making task. When members of a pair had similar visual acuity, they performed better as a pair than they did individually, as long as they were given the opportunity to communicate freely. The authors concluded that this benefit was attributable to the participants providing accurate estimates of their own confidence level within their communication.

*Optimisation of receiver behaviour, due to understanding of others’ knowledge states.* Although there is still considerable debate over whether metacognition relating to one’s own mind involves the same processes as metacognition regarding the mind of others (e.g. see Carruthers, 2009), when it comes to *explicit* metacognition, it certainly seems likely that understanding one’s own mind, and understanding those of others, are likely to be linked, given the degree of reflective awareness involved (even if the specifics of which understanding comes first may be unclear; see the various models outlined in Carruthers, 2009). Nonetheless, it is worth noting that arguments in support of the EMCC that place the emphasis on explicit understanding of *other’s* minds (e.g. Heyes, 2016), are using the term metacognition in a context that, in other areas of the literature, would be regarded as non-standard, and possibly even controversial (e.g. Nichols and Stich, 2003).

Furthermore, the literature previously discussed in this review, which relates to the question of whether particular types of metacognition may be unique to humans, may not be strictly relevant in addressing this particular interpretation of the EMCC, since an ability to evaluate one’s own confidence or uncertainty may or may not predict one’s understanding of others as mental agents. However, if anything, it is probably much easier to make an argument that an explicit understanding of others’ minds is restricted to humans. “Theory of mind” (e.g. Premack and Woodruff, 1978), has been a focus of much empirical enquiry and many theoretical analyses in both comparative and developmental psychology, and therefore we do not intend to reiterate findings or conclusions in depth here. But a number of accounts have proposed separate systems for mindreading as a means to reconcile behavioural findings suggesting some tracking of other’s mental states in toddlers and animals (e.g. Krupenye et al., 2016; Southgate et al., 2007), with consistent evidence that explicit understanding of others’ beliefs does not develop until around the age of four in children (e.g. Wellman et al., 2001), as well as the failure of nonhuman apes in an equivalent nonverbal analogue task (Call and Tomasello, 1999). Apperly and Butterfill’s (e.g. Apperly and Butterfill, 2009) two-systems account is perhaps the most high profile of the theories that have been proposed to reconcile these findings (although others exist, e.g. Perner and Roessler, 2012). However, here it suffices to note that it is a relatively widespread view that implicit and explicit tests of understanding of others’ mental states may be measuring different processes.

It follows fairly logically, then, to conclude that a System 2, or explicit, understanding of others’ mental states might give an agent a significant advantage in their use of social information, allowing them to use this more flexibly and in accordance with the most up-to-date information about who is likely to be an effective model (in line with Heyes’s distinction between cook-like and planetary-like decision rules). However, in spite of the convincing rationale for this potential advantage, to our knowledge no empirical studies to date have tested whether an explicit understanding of social sources as mental agents confers benefits over and above implicitly represented strategies. For example, it might be expected that with advancing age, children become capable of using social information in increasingly sophisticated ways, perhaps overriding general purpose biases and heuristics when new information about others’ actual knowledge comes to light. For example, recognising that the actions of a single knowledgeable individual are likely to be more valuable than the same number of actions from multiple uninformed individuals. In the absence of such evidence, this particular assumption of how the EMCC might operate may well be plausible, but it nonetheless remains highly speculative.

*Optimisation of sender behaviour, due to understanding of others’ knowledge states.* Explicit understanding of others’ mental states might also bring about changes in sender behaviour, as well as that of the receiver. Even in animals, social learning is not necessarily restricted to the use of inadvertent cues acquired from others as a consequence of incidental observation of behaviours performed only in the interests of the actor themselves. Therefore senders can play an active role in social transmission, and the finer details of how they do so may be significant. In animals, behaviour that functions to teach others has been documented in a number of different species (e.g. meerkats: Thornton and McAuliffe, 2006; ants: Franks and Richardson, 2006). However, as with animals’ social learning “strategies”, this is a further example where the adaptive function of the behaviour is assumed not to be driven by the agent’s *understanding* of that function. This is therefore very different from teaching as it would normally be

interpreted in humans, which would generally be expected to implicate some degree of recognition of the part of the teacher, regarding the potential effect of the behaviour on another's knowledge or skill level. The question pertinent to the EMCC then, is whether sender behaviour can facilitate the learning of others much more effectively when senders have an understanding of others' states of knowledge or ignorance. As with the mechanism described in the previous section, a logical argument for this can be constructed with very little difficulty. At the very least, such an understanding would open out the potential contexts within which teaching could occur, whereas (to extend the analogy) "planetary" teaching behaviour would be expected to be restricted to contexts involving an extended selection history (including species-typical behaviours, such as particular predation skills as in meerkats, or well-defined categories of episodic knowledge, such as routes to food sources as in ants). Caldwell et al., (2017) have previously argued that intentional teaching may be particularly valuable for supporting cumulative culture, since almost by definition cumulative culture is likely to involve novel behavioural variants that are not part of the species-typical repertoire.

In addition to broadening the contexts across which teaching can occur, an understanding of others' minds may also render teaching behaviour far more effective, due to the ability to gauge one's own behaviour in response to the apparent needs of the learner. Teachers can selectively show or perform particular features of what is to be transmitted, with a view to making this maximally informative, based on their own understanding of what might benefit a learner. Furthermore, an understanding of the mind of the learner also allows for adjustments to be made online during teaching, in direct response to the learner's level of success. Mistakes can be corrected, or misunderstandings clarified, and redundancy can be avoided by skipping elements already mastered. A similarly high level of responsiveness might be unlikely in the absence of sensitivity to the meaning of potential cues to knowledge and competence.

There is some literature documenting developmental changes in teaching behaviour in young children which appears consistent with this. Ronfard and Corriveau (2016) studied how children aged between three and five years old taught a game to puppet characters that had demonstrated differing levels of competence. They found that children's ability to monitor the relative accuracy of the puppets improved with age, and that older children tailored their instruction more precisely to the apparent needs of the learner, more often directly addressing the specific errors of individual puppets. This finding is consistent with the assumption that increasing awareness of others' mental states can facilitate transmission by altering sender behaviour.

*Optimisation of the sender-receiver interaction due to understanding of minds of self and other.* It should be noted that the above categorisations are not intended to be regarded as mutually exclusive; indeed it would be surprising in some cases if they operated in complete isolation from one another. In addition, for each of these two categorisations, it might be expected that benefits arising from the interaction between the two alternative mechanisms could be more than the sum of their individual parts (i.e. the combination of smart behaviour on the part of both sender and receiver, or the combination of understanding of one's own knowledge in relation to others', might be particularly effective in generating cumulative culture). For example, an interaction between an experienced individual (who is motivated to impart their knowledge) and a naïve partner (who is motivated to learn), will likely be most effective when each recognises the other's motivation. In the categorisations detailed above we have only discussed communication in the context of sender

behaviour, but communication on the part of the receiver may also have a powerful role to play once there is a mutual appreciation of a shared motivation. This allows the receiver to effectively communicate what the *sender* may need to know, in order to provide the most effective guidance. Clearly, such bidirectional cooperative interactions involve high levels of flexibility in the behaviour of both the sender and receiver, informed by their understanding of both their own, and their partner's, state of knowledge. For the receiver to effectively communicate their needs, this is likely to include not just a representation of the sender's mental state, but a representation of the sender's representation of the receiver's mental state (second order theory of mind, e.g., Perner and Wimmer, 1985), which they are in a position to correct, update, or augment.

In such contexts the breakdown of roles into "sender" and "receiver" becomes significantly less clear-cut. Consistent with this, it has been shown that *both* members of a pair can improve their performance on certain tasks through two-way information sharing (Bahrami et al., 2010). Bahrami et al., found that such benefits only occurred when participants were able to communicate freely, and thus share their confidence levels in addition to their own initial best guess, consistent with the idea that these benefits arise due to metacognitive competence relating to both communicating one's own level of knowledge, and the interpretation of others'. There may be particular value in being able to interpret another's knowledge state relative to one's own, in ways that make each interacting agent simultaneously both a provider of social information (through their influence on another's success level), and also a beneficiary (through their own improved performance).

It needs to be acknowledged that metacognition is not infallible; people are often under- or over-confident when rating their performance (for example see Metcalfe and Dunlosky, 2008; Miller and Geraci, 2011). However, this shortcoming in self-regulation may be overcome by the shared nature of *explicit* metacognition: Bang et al., (2017) found a collective benefit in making collective perceptual judgements when 'poorly calibrated' groups of participants (groups where the more confident members were not the more accurate or skilled members) matched their confidence levels. This may suggest that explicitly sharing metacognitive information about confidence, such as in the scenarios described by Shea et al., (2014), would help to counteract negative effects of poor metacognitive accuracy on personal decision making.

*How can the EMCC be tested?.* Currently, evidence for the EMCC remains very limited. The accounts proposed by Shea et al., (2014) and Heyes (2016) are built on indirect inference, drawing links between apparent differences in metacognitive awareness in humans versus animals, and the plausibility of metacognition facilitating cumulative culture (sometimes supported by evidence suggesting that outcomes of social learning may be influenced by the availability of metacognitive information). However, in order to effectively evaluate these proposals, more direct evidence is now required. Firstly, there is a need for further empirical evidence that experimentally manipulates the availability of metacognitive resources and/or information, in order to look for direct impact on outcomes of social learning. Secondly, there is a need for studies that fully operationalise cumulative culture, as opposed to studying single transmission events, or looking at interactions only at the level of the dyad. We would expect the combination of these methods to elicit results that showed a reduction in ratchet-like behaviour over generations. That is, methods that are shown to produce accumulation of improvement over generations under normal conditions would no longer

show this accumulation when those tasks are carried out in conditions that prevent access to or use of system-2.

Identifying empirical evidence of a causal link between explicit metacognition and cumulative culture is critical in order to establish that the EMCC has some explanatory power over and above other speculative explanations of cumulative culture in terms of other apparently uniquely human features. There are a multitude of features that differentiate humans from other animals, and it is often not difficult to make an argument for the involvement of a particular cognitive or behavioural trait in cumulative culture. What will distinguish such proposals is the availability of empirical evidence that convincingly demonstrates a causal link between the feature or trait in question and outcomes of social learning.

Brain imaging techniques may be used to identify if there are correlations between brain regions activated when using adaptive social learning strategies and capacities for cumulative culture, and those activated when making explicit metacognitive judgements; the EMCC would predict strong correlations in these areas. However, this would not provide direct evidence of a causal link between explicit metacognition and cumulative culture.

Studies which experimentally manipulate the availability of metacognitive resources or information are therefore required to test the EMCC. This is likely to be considerably easier to do for studies investigating the effects of sender behaviour, compared with those focusing on the abilities of the receiver. Accordingly, studies already exist (e.g., Bahrami et al., 2010, described previously) which have experimentally manipulated opportunities for communication, and therefore the potential for sharing metacognitive information, which demonstrate positive impacts on the effectiveness of social information. However, it is much harder to manipulate the extent to which a receiver can employ explicit metacognition in their interpretation of others' behaviour, since it is not possible to simply remove human capacities for metacognition. Nonetheless, we can envisage at least two potentially fruitful avenues of investigation which would allow some insight into the effects of the availability of metacognitive resources on the part of the receiver. The first of these would involve the use of dual task methods, described previously. The EMCC specifically implicates System 2 involvement, and it should be possible to block or impede the involvement of System 2 using dual tasks that also place demands on executive function. This could therefore act as a proxy for restricting explicit metacognition directly, with the expectation that reduced access to explicit processing would restrict participants' ability to interpret (and also share) social information in ways that could be critical for generating cumulative culture.

The expected outcomes of such tasks would be a reduced capacity to make the requisite social learning decisions required for cumulative culture to emerge, and therefore a reduction or absence of ratcheting in tasks that would ordinarily have been shown to produce a ratchet effect in the laboratory.

A further promising approach would be to investigate the effects of developmental changes in metacognitive competence on performance in social learning paradigms, by studying both in young children of a range of ages. The EMCC predicts strong correlations between the emergence of explicit metacognitive competence, executive function capacities and proficiency in strategic social learning tasks.

It should be noted however, that none of the approaches discussed would allow direct manipulation of the involvement of explicit metacognition. Whilst dual task methods offer potential for experimental manipulation, this would be premised on an assumption that these functioned to block explicit metacognition. Interpretation of such results would therefore be strengthened considerably by the existence of additional evidence validating

this assumption, which to our knowledge has yet to be tested. We know of no studies to date which have investigated the involvement of System 2 processing (i.e., as assessed through evidence of interference under dual task conditions) in explicit reports of metacognition such as judgements of confidence and feelings of knowing (JOC and FOK). Such tests would provide key evidence in evaluating the EMCC, which would inform both theory and method.

Neuroimaging approaches are also somewhat limited in their scope as although they may demonstrate the involvement of brain areas associated with explicit metacognition, they are not necessarily able to show whether participants are unable to produce ratcheting effects in cultural evolution tasks without the involvement of these areas.

Evidence from developmental approaches, whilst offering insights into the potential for cumulative culture both before and after the development of explicit metacognition, would be necessarily pseudo-experimental, involving no attempt to experimentally manipulate the variable of interest, making it much more difficult to identify a causal association. Nonetheless, if relationships were to be found between individual-level measures of metacognitive ability, and individual-level measures of social learning proficiency, especially if these persisted when controlling for age, this would provide fairly convincing support for the EMCC. Certainly, in spite of their respective limitations, both dual task and developmental approaches, and to some extent neuroimaging approaches, have the potential to provide much stronger support than circumstantial evidence of common exclusivity to humans.

In addition, we would also suggest that a truly robust test of the EMCC would involve laboratory simulation of cumulative culture (e.g. Caldwell and Millen, 2008, 2009), rather than the study of single transmission events, or dyadic interactions. If ultimately the EMCC aims to explain the (group-level) phenomenon of human cumulative culture, it is critical to show that the involvement or otherwise of explicit metacognition does actually impact on the degree to which learning benefits accrue over multiple generations (e.g. Caldwell, 2018). Thus, experimental designs using transmission chain or micro-society paradigms (e.g. Mesoudi and Whiten, 2008) would provide a key source of evidence in evaluating the EMCC.

Finally, we also propose that research should be targeted at identifying which of the routes described in Section 'How Might Explicit Metacognition Facilitate Cumulative Culture?' account for any link found between explicit metacognition and cumulative culture. We have argued that in principle all are plausible. However, a full account would specify which of these (whether in isolation or combination) appeared to be critical to supporting ratchet effects in cultural evolution.

## Conclusion

To date, there is as yet no generally accepted theory explaining the apparent uniqueness of human cumulative culture. The theories recently proposed by Heyes (2016, 2018b) and Shea et al., (2014), which implicate the use of explicit metacognition and System 2 cognition (or Type 2 processes) have the potential to provide a convincing account of distinctively human culture. Here we have used the term the Explicitly Metacognitive Cumulative Culture hypothesis (EMCC), to refer to any view proposing that System 2 processes allow human learners to use metacognition in ways that facilitate social learning. We have also proposed a number of different routes by which System 2 metacognition might have potential to enable cumulative culture, through optimising the behaviour of either the sender or receiver

behaviour, based on an explicit understanding of the mental states of either oneself or others.

We have established that, to date, there has been little or no empirical work directly testing these proposals. Indirect evidence is available which provides some support for the view that the implicit metacognitive competence identified in animals depends on processes distinct from explicit metacognition. There is also some support for the view that information transmission may become more effective with increasing metacognitive competence (at least on the part of the sender), and that having the opportunity to communicate metacognitive confidence levels, in addition to task responses themselves, can also increase benefits of social learning. However, there are significant gaps in the literature, particularly from the point of view of establishing the mechanistic links (see section 'How Might Explicit Metacognition Facilitate Cumulative Culture?') between apparently distinctively human explicit metacognition, and the evolutionary anomaly of cumulative culture. In particular, we see a need for studies involving laboratory tasks which operationalise the group-level phenomenon of cumulative culture, rather than focussing on single transmission events. We have also highlighted that dual task methods, understood to restrict the use of System 2, have not as yet been exploited within the literature on social learning and cultural evolution, and that these offer a potentially powerful tool for experimentally manipulating the availability of cognitive resources needed for explicit metacognitive processing. We have further suggested that developmental research in human children could shed valuable light on this topic, as children's advancing metacognitive competence offers a natural experiment permitting investigation of the resulting effects on the efficacy of social learning, through increasingly flexible and sophisticated behaviour (whether in the role of sender or receiver).

In conclusion therefore, we consider that the EMCC has considerable promise as a potential explanation for the elaborateness of human culture in relation to the behavioural traditions of other animals. Further research is now warranted in order to test key assumptions and flesh out the details of the links.

### Data availability

Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

Received: 10 July 2018 Accepted: 19 November 2018

Published online: 04 December 2018

### References

- Allen J, Weinrich M, Hoppitt W, Rendell LE (2013) Network-based diffusion analysis reveals cultural transmission of lobeail feeding in Humpback Whales. *Science* 340(6131):485–488. <https://doi.org/10.1126/science.1231976>
- Aplin LM, Farine DR, Morand-Ferron J, Cockburn A, Thornton A, Sheldon BC (2015) Experimentally induced innovations lead to persistent culture via conformity in wild birds. *Nature* 518(7540):538–541. <https://doi.org/10.1038/nature13998>
- Apperly IA, Butterfill SA (2009) Do humans have two systems to track beliefs and belief-like states? *Psychol Rev* 116(4):953–970. <https://doi.org/10.1037/a0016923>
- Bahrami B, Olsen K, Latham PE, Roepstorff A, Rees G, Frith CD (2010) Optimally interacting minds. *Science* 329:1081–1086
- Bang D, Aitchison L, Moran R, Herculano-Houzel S, Rafiee B, Mahmoodi A, Summerfield C (2017) Confidence matching in group decision-making. *Nat Human Behav* 1(6):117. <https://doi.org/10.1038/s41562-017-0117>
- Beck W (1992) Aboriginal preparation of Cycas seeds in Australia. *Econ Bot* 46(2):133–147. <https://doi.org/10.1007/BF02930628>
- Bender A, Beller S (2014) Mangarevan invention of binary steps for easier calculation. *Proc Natl Acad Sci* 111(4), 1322–1327. <https://doi.org/10.1073/pnas.1309160110>
- Beran MJ, Perdue BM, Futch SE, Smith JD, Evans TA, Parrish AE (2015) Go when you know: Chimpanzees' confidence movements reflect their responses in a computerized memory task. *Cognition* 142:236–46. <https://doi.org/10.1016/j.cognition.2015.05.023>
- Bernard S, Proust J, Clément F (2015) Procedural metacognition and false belief understanding in 3- to 5-year-old children. *PLoS ONE* 10(10). <https://doi.org/10.1371/journal.pone.0141321>
- Caldwell CA (2018) Using experimental research designs to explore the scope of cumulative culture in humans and other animals. *Topics Cogn Sci*
- Caldwell CA, Cornish H, Kandler A (2016) Identifying innovation in laboratory studies of cultural evolution: rates of retention and measures of adaptation. *Philos Trans R Soc B* 371(1690):20150193. <https://doi.org/10.1098/rstb.2015.0193>
- Caldwell CA, Millen AE (2008) Experimental models for testing hypotheses about cumulative cultural evolution. *Evol Human Behav* 29(3):165–171. <https://doi.org/10.1016/j.evolhumbehav.2007.12.001>
- Caldwell CA, Millen AE (2009) Social learning mechanisms and cumulative cultural evolution. *Psychol Sci* 20(12):1478–1483. <https://doi.org/10.1111/j.1467-9280.2009.02469.x>
- Caldwell CA, Renner E, Atkinson M (2017) Human teaching and cumulative cultural evolution. *Rev Phil Psychol* <https://doi.org/10.1007/s13164-017-0346-3>
- Call J, Carpenter M (2001) Do apes and children know what they have seen? *Anim Cogn* 3(4):207–220. <https://doi.org/10.1007/s100710100078>
- Call J, Carpenter M, Tomasello M (2005) Copying results and copying actions in the process of social learning: Chimpanzees (Pan troglodytes) and human children (Homo sapiens). *Anim Cogn* 8(3):151–163. <https://doi.org/10.1007/s10071-004-0237-8>
- Call J, Tomasello M (1999) A nonverbal false belief task: the performance of children and great apes. *Child Dev* 70(2):381–395. <https://doi.org/10.1111/1467-8624.00028>
- Carruthers P (2008) Meta-cognition in animals: a skeptical look. *Mind Lang* 23(1):58–89. <https://doi.org/10.1111/j.1468-0017.2007.00329.x>
- Carruthers P (2009) How we know our own minds: the relationship between mindreading and metacognition. *Behav Brain Sci* 32(2):121–138–182. <https://doi.org/10.1017/S0140525X09000545>
- Carruthers P, Ritchie JB (2012) The emergence of metacognition: affect and uncertainty in animals. In: Beran MJ, Brandl J, Perner J, Proust J (eds) *Foundations of Metacognition*. Oxford University Press: Oxford
- Cultice JC, Somerville SC, Wellman HM (1983) Preschoolers' memory monitoring: feeling-of-knowing judgments. *Child Dev* 54(6):1480–1486
- Davis SJ, Vale GL, Schapiro SJ, Lambeth SP, Whiten A (2016) Foundations of cumulative culture in apes: Improved foraging efficiency through relinquishing and combining witnessed behaviours in chimpanzees (Pan troglodytes). *Sci Rep* 6:1–12. <https://doi.org/10.1038/srep35953>
- Dean LG, Vale GL, Laland KN, Flynn E, Kendal RL (2014) Human cumulative culture: a comparative perspective. *Biol Rev* 89(2):284–301. <https://doi.org/10.1111/brv.12053>
- De Neys W (2006) Automatic-heuristic and executive-analytic processing during reasoning: chronometric and dual-task considerations. *Q J Exp Psychol* 59(6):1070–1100. <https://doi.org/10.1080/02724980543000123>
- Derooy O, Spence C, Noppeney U (2016) Metacognition in multisensory perception. *Trends Cogn Sci* 20(10):736–747. <https://doi.org/10.1016/j.tics.2016.08.006>
- Dienes Z, Perner J (1998) A theory of implicit and explicit knowledge (1999), 735–808
- Dunlosky J, Metcalfe J (2009) *Metacognition*. SAGE Publications, Thousand Oaks
- Ehn M, Laland KN (2012) Adaptive strategies for cumulative cultural learning. *J Theor Biol* 301:103–111. <https://doi.org/10.1016/j.jtbi.2012.02.004>
- Enquist M, Eriksson K, Ghirlanda S (2007) Critical social learning: a solution to rogers' s paradox of nonadaptive culture. *Am Anthropol* 109(4):727–734. <https://doi.org/10.1525/AA.2007.109.4.727.728>
- Enquist M, Ghirlanda S, Jarrick A, Wachtmeister C-A (2008) Why does human culture increase exponentially? *Theor Popul Biol* 74(1):46–55. <https://doi.org/10.1016/j.tpb.2008.04.007>
- Epstein S (1994) Integration of the cognitive and the psychodynamic unconscious. *Am Psychol* 49(8):709–724. <https://doi.org/10.1037/0003-066X.49.8.709>
- Evans TA (2007) Hypothetical thinking: dual processes in reasoning and judgement. Psychology Press, New York
- Evans TA, Stanovich KE (2013) Dual-process theories of higher cognition: advancing the debate. *Perspect Psychol Sci* 8(3):223–241. <https://doi.org/10.1177/1745691612460685>
- Flavell JH (1979) Metacognition and cognition monitoring: a new area of cognitive developmental inquiry. *Am Psychol* 34(10):906–911. 34
- Footo AL, Crystal JD (2007) Metacognition in the rat. *Curr Biol* 17(6):551–555. <https://doi.org/10.1016/j.cub.2007.01.061>
- Franks NR, Richardson T (2006) Teaching in tandem-running ants. *Nature* 439(7073):153–153. <https://doi.org/10.1038/439153a>

- Galef BG (2013) Imitation and local enhancement: detrimental effects of consensus definitions on analyses of social learning in animals. *Behav Process* 100:123–130. <https://doi.org/10.1016/j.beproc.2013.07.026>
- Goel V, Dolan RJ (2003) Explaining modulation of reasoning by belief. *Cognition* 87:B11–B22. [https://doi.org/10.1016/S0010-0277\(02\)00185-3](https://doi.org/10.1016/S0010-0277(02)00185-3)
- Goupil L, Romand-monniere M, Kouider S (2016) Infants ask for help when they know they don't know. *Proc Natl Acad Sci* 113(13), 3492–3496. <https://doi.org/10.1073/pnas.1515129113>
- Hampton RR (2009) Multiple demonstrations of metacognition in nonhumans: converging evidence or multiple mechanisms? *Comp Cogn Behav Rev* 4:17–28. <https://doi.org/10.3819/ccbr.2009.40002>
- Henrich J, Boyd R, Richerson PJ (2008) Five misunderstandings about cultural evolution. *Human Nat* 19(2):119–137. <https://doi.org/10.1007/s12110-008-9037-1>
- Heyes C (2011) Automatic imitation. *Psychol Bull* 137(3):463–483. <https://doi.org/10.1037/a0022288>
- Heyes C (2016) Who knows? Metacognitive social learning strategies. *Trends Cogn Sci* 20(3):204–213. <https://doi.org/10.1016/j.tics.2015.12.007>
- Heyes C (2018a) *Cognitive gadgets*. Harvard University Press, Cambridge
- Heyes C (2018b) Enquire within: cultural evolution and cognitive science. *Phil Trans R Soc London* 373(1743). <https://doi.org/10.1098/rstb.2017.0051>
- Hobaiter C, Poisot T, Zuberbühler K, Hoppitt W, Gruber T (2014) Social network analysis shows direct evidence for social transmission of tool use in wild chimpanzees. *PLoS Biol* 12(9):e1001960. <https://doi.org/10.1371/journal.pbio.1001960>
- Jones PL, Ryan MJ, Flores V, Page RA (2013) When to approach novel prey cues? Social learning strategies in frog-eating bats. *Proc R Soc B* 280(1772). <https://doi.org/10.1098/rspb.2013.2330>
- Keren G, Schul Y (2009) Two is not always better than one. *Perspect Psychol Sci* 4(6):533–550. <https://doi.org/10.1111/j.1745-6924.2009.01164.x>
- Krupenye C, Kano F, Hirata S, Call J, Tomasello M (2016) Great apes anticipate that other individuals will act according to false beliefs. *Science* 354(6308):110–114. <https://doi.org/10.1126/science.aaf8110>
- Kuhn D (2000) Metacognitive development. *Curr Dir Psychol Sci* 9(5):178–181. <https://doi.org/10.1111/1467-8721.00088>
- Laland KN (2004) Social learning strategies. *Anim Learn Behav* 32(1):4–14. <https://doi.org/10.3758/BF03196002>
- Lehmann L, Feldman MW, Kaeuffer R (2010) Cumulative cultural dynamics and the coevolution of cultural innovation and transmission: An ESS model for panmictic and structured populations. *J Evolut Biol* 23(11):2356–2369. <https://doi.org/10.1111/j.1420-9101.2010.02096.x>
- Lewis HM, Laland KN (2012) Transmission fidelity is the key to the build-up of cumulative culture. *Philos Trans R Soc B: Biol Sci* 367(1599):2171–2180. <https://doi.org/10.1098/rstb.2012.0119>
- McClure SM, Laibson DI, Loewenstein G, Cohen JD (2004) Separate neural systems value immediate and delayed monetary rewards. *Science* 306(August 2016):503–507. <https://doi.org/10.1126/science.1100907>
- Meltzoff AN (1988) The human infant as Homo imitans. In: Zenall T, Galef G (eds) *Social learning: psychological and biological perspectives*. Psychology Press: New York, pp. 318–340
- Mesoudi A, Thornton A (2018) What is cumulative cultural evolution? *Proc R Soc B* <https://doi.org/10.1098/rspb.2018.0712>
- Mesoudi A, Whiten A (2008) The multiple roles of cultural transmission experiments in understanding human cultural evolution. *Philos Trans R Soc B* 363(1509):3489–3501. <https://doi.org/10.1098/rstb.2008.0129>
- Metcalfe J, Dunlosky J (2008) *Metamemory*. *Cognitive Psychology of Memory*. pp. 349–362
- Miller TM, Geraci L (2011) Unskilled but aware: reinterpreting overconfidence in low-performing students. *J Exp Psychol: Learn Mem Cogn* 37(2):502–506. <https://doi.org/10.1037/a0021802>
- Misailidi P (2010) Children's metacognition and theory of mind: bridging the gap. In: Efklides A, Misailidi P (eds) *Trends and prospects in metacognition research*. Springer US, New York, p. 279–292
- Nelson TO, Narens L (1990) Metamemory: a theoretical framework and new findings. *Psychol Learn Motiv* 26:125–173. [https://doi.org/10.1016/S0079-7421\(08\)60053-5](https://doi.org/10.1016/S0079-7421(08)60053-5)
- Nichols S, Stich SP (2003) *Mindreading: an integrated account of pretence, self-awareness, and understanding other minds*. Oxford University Press, New York
- Osman M (2004) An evaluation of dual-process theories of reasoning. *Psychon Bull Rev* 11(6):988–1010. <https://doi.org/10.3758/BF03196730>
- Pashler H (1994) Dual-task interference in simple tasks: data and theory. *Psychol Bull* 116(2):220–244. <https://doi.org/10.1037/0033-2909.116.2.220>
- Perner J, Roessler J (2012) From infants' to children's appreciation of belief. *Trends Cogn Sci* 16(10):519–525. <https://doi.org/10.1016/j.tics.2012.08.004>
- Perner J, Wimmer H (1985) John thinks that Mary thinks that: attribution of second order beliefs by 5-year-old to 10-year-old children. *J Exp Child Psychol* 39(3):437–471. [https://doi.org/10.1016/0022-0965\(85\)90051-7](https://doi.org/10.1016/0022-0965(85)90051-7)
- Perry CJ, Barron AB (2013) Honey bees selectively avoid difficult choices. *Proc Natl Acad Sci USA* 110(47), 19155–9. <https://doi.org/10.1073/pnas.1314571110>
- Pike TW, Laland KN (2010) Conformist learning in nine-spined sticklebacks' foraging decisions. *Biol Lett* 6(4):466–468. <https://doi.org/10.1098/rsbl.2009.1014>
- Premack D, Woodruff G (1978) Does the chimpanzee have a theory of mind? *Behav Brain Sci* 4:515–526
- Rendell LE, Boyd R, Cownden D, Enquist M, Eriksson K, Feldman MW, Laland KN (2010) Why copy others? Insight from the social learning strategies tournament. *Science* 328(5975):208–213. <https://doi.org/10.1126/science.1184719.Why>
- Rendell LE, Fogarty L, Laland KN (2010) Rogers' paradox recast and resolved: population structure and the evolution of social learning strategies. *Evolution* 64(2):534–548. <https://doi.org/10.1111/j.1558-5646.2009.00817.x>
- Roberts MJ, Newton EJ (2001) Inspection times, the change task, and the rapid-response selection task. *Q J Exp Psychol A* 54(4):1031–1048. <https://doi.org/10.1080/713756016>
- Roebbers CM (2017) Executive function and metacognition: towards a unifying framework of cognitive self-regulation. *Dev Rev* 45:31–51. <https://doi.org/10.1016/j.dr.2017.04.001>
- Rohwer M, Kloof D, Perner J (2012) Escape from metaignorance: how children develop an understanding of their own lack of knowledge. *Child Dev* 83(6):1869–1883. <https://doi.org/10.1111/j.1467-8624.2012.01830.x>
- Ronfard S, Corriveau KH (2016) Teaching and preschoolers' ability to infer knowledge from mistakes. *J Exp Child Psychol* 150:87–98. <https://doi.org/10.1016/j.jecp.2016.05.006>
- Sannomiya M, Ohtani K (2015) Does a dual-task selectively inhibit the metacognitive activities in text revision? *Think Skills Creat* 17:25–32. <https://doi.org/10.1016/j.tsc.2015.04.002>
- Shea N, Boldt A, Bang D, Yeung N, Heyes C, Frith CD (2014) Supra-personal cognitive control and metacognition. *Trends Cogn Sci* <https://doi.org/10.1016/j.tics.2014.01.006>
- Smith ER, DeCoster J (2000) Dual-process models in social and cognitive psychology: conceptual integration and links to underlying memory systems. *Personal Social Psychol Rev* 4(2):108–131. [https://doi.org/10.1207/S15327957PSPR0402\\_01](https://doi.org/10.1207/S15327957PSPR0402_01)
- Smith JD, Beran MJ, Couchman JJ, Coutinho MVC (2008) The comparative study of metacognition: sharper paradigms, safer inferences. *Psychon Bull Rev* 15(4):679–691. <https://doi.org/10.3758/PBR.15.4.679>
- Smith JD, Redford JS, Beran MJ, Washburn DA (2009) Rhesus monkeys (*Macaca mulatta*) adaptively monitor uncertainty while multi-tasking. *Anim Cogn* 13(1):93–101. <https://doi.org/10.1007/s10071-009-0249-5>
- Smith JD, Schull J, Strote J, McGee K, Egnor R, Erb L (1995) The uncertain response in the bottlenosed dolphin (*Tursiops truncatus*). *J Exp Psychol General* 124(4):391–408. <https://doi.org/10.1037/0096-3445.124.4.391>
- Smolla M, Alem S, Chittka L, Shultz S (2016) Copy-when-uncertain: bumblebees rely on social information when rewards are highly variable. *Biol Lett* 12(6):208–213. <https://doi.org/10.1098/rsbl.2016.0188>
- Sole LM, Shettleworth SJ, Bennett PJ (2003) Uncertainty in pigeons. *Psychon Bull Rev* 10(3):738–745. <https://doi.org/10.3758/BF03196540>
- Southgate V, Senju A, Csibra G (2007) Action anticipation through attribution of false belief by 2-year-olds. *Psychol Sci* 18(7):587–592. <https://doi.org/10.1111/j.1467-9280.2007.01944.x>
- Stanovich KE (1999) *Who is rational? Studies of individual differences in reasoning*. Psychology Press, Mahwah
- Stanovich KE (2004) *The robot's rebellion: finding meaning the age of Darwin*. University of Chicago Press, Chicago
- Templer VL, Lee KA, Preston AJ (2017) Rats know when they remember: transfer of metacognitive responding across odor-based delayed match-to-sample tests. *Animal Cogn* 1–16. <https://doi.org/10.1007/s10071-017-1109-3>
- Tennie C, Call J, Tomasello M (2009) Ratcheting up the ratchet: on the evolution of cumulative culture. *Philos Trans R Soc Lond Ser B* 364(1528):2405–15. <https://doi.org/10.1098/rstb.2009.0052>
- Thornton A, McAuliffe K (2006) Teaching in wild meerkats. *Science* 313(5784):227–229. <https://doi.org/10.1126/science.1128727>
- Tomasello M (1990) Cultural transmission in tool use and communicatory signaling of chimpanzees? In: Parker S, Gibson K (eds) "Language" and intelligence in monkeys and apes: comparative developmental perspectives. Cambridge University Press, Cambridge, p 274–311
- Tomasello M (1999) The human adaptation for culture. *Annu Rev Anthropol* 28:509–529. <https://doi.org/10.1017/S0140525X0003123X>
- Tsujii T, Watanabe S (2009) Neural correlates of dual-task effect on belief-bias syllogistic reasoning: a near-infrared spectroscopy study. *Brain Res* 1287:118–125. <https://doi.org/10.1016/j.brainres.2009.06.080>
- Voelkl B, Huber L (2000) True imitation in marmosets. *Anim Behav* 60(2):195–202. <https://doi.org/10.1006/anbe.2000.1457>

- Wason PC, Evans TA (1974) Dual processes in reasoning? *Cognition* 3(2):141–154. [https://doi.org/10.1016/0010-0277\(74\)90017-1](https://doi.org/10.1016/0010-0277(74)90017-1)
- Wellman HM, Cross D, Watson J (2001) Meta-analysis of theory of mind development: the truth about false belief. *Child Dev* 72(3):655–684
- Whiten A, Caldwell CA, Mesoudi A (2016) Cultural diffusion in humans and other animals. *Curr Opin Psychol* 8:15–21. <https://doi.org/10.1016/j.copsyc.2015.09.002>
- Whiten A, Horner V, Litchfield CA, Marshall-Pescini S (2004) How do apes ape? *Learn Behav* 32(1):36–52. <https://doi.org/10.3758/BF03196005>

### Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No. 648841 RATCHETCOG ERC-2014-CoG.

### Additional information

**Competing interests:** The authors declare no competing interests.

**Reprints and permission** information is available online at <http://www.nature.com/reprints>

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018