Levelling the playing field in studying cumulative cultural evolution: conceptual and methodological advances in nonhuman animal research

Rawlings, B¹,², Legare, C.H¹. Brosnan, S.F.²,³ and Vale, G.L²,³

1. Department of Psychology, The University of Texas at Austin, Texas, USA.

2. National Center for Chimpanzee Care, Michale E. Keeling Center for Comparative Medicine and Research, The University of Texas MD Anderson Cancer Center, Bastrop, TX, USA

3. Department of Psychology, Language Research Center, Neuroscience Institute and Center for Behavioral Neuroscience, Georgia State University, Atlanta, GA, USA


Abstract
Cumulative cultural evolution (CCE), the improvement of cultural traits over generations via social transmission, is widely believed to be unique to humans. The capacity to build upon others’ knowledge, technologies, and skills has produced the most diverse and sophisticated technological repertoire in the animal kingdom. Yet, inconsistency in both the definitions and criteria used to determine CCE and the methodology used to examine it across studies may be hindering our ability to determine which aspects are unique to humans. Issues regarding how improvement is defined and measured, and whether some criteria are empirically testable are of increasing concern to the field. In this article, we critically assess the progress made in the field and current points of debate, from conceptual and methodological perspectives. We discuss how inconsistency in definitions is detrimental to our ability to document potential evidence of CCE to nonhuman animals. We build on Mesoudi and Thornton’s (2018) recently described core and extended CCE criteria to make specific recommendations about, from a comparative lens, which criteria should be used as evidence of CCE. We evaluate existing data from both wild and captive studies of nonhuman animals using these suggestions. We finish by discussing issues currently faced by researchers studying CCE in nonhuman animals, particularly nonhuman primates, and provide suggestions that may overcome these concerns and move the field forward.

Keywords: cognition; comparative psychology; cumulative cultural evolution; innovation; nonhuman animals; culture; social learning.
Over recent decades, the topic of cumulative cultural evolution (CCE) – in which cycles of innovation and social learning lead to adaptive modifications accumulating over historical time - has become a focal topic in biological and social sciences (Caldwell & Millen, 2008; Dean et al., 2014; Legare, 2017; Mesoudi & Thornton, 2018; Tennie et al., 2009). CCE is key to the sophisticated technologies, customs and knowledge which pervade our lives, and is the reason humans have such developments as complex social institutions, rituals, biomedicine, agriculture, and literature (Boyd & Richerson, 1985; Henrich, 2015; Legare & Nielsen, 2020; Muthukrishna & Henrich, 2016). Computers, for example, are not the invention of any one individual, but are the product of centuries of cumulative cultural improvement; they have evolved from steam-based analytical engines in the early 1800s, to Alan Turing’s seminal machine capable of algorithmic logic, to large, single-circuit digital desktops computers, all of which are products of building upon others’ output (with hundreds of iterations in between), paving the way for the lightweight and computationally powerful laptops we use today. Likewise, over generations, horticultural societies have developed sophisticated processing techniques to process otherwise toxic foods, such as cycads and cassava to provide new sources of food (Beck, 1992; Henrich & Henrich, 2010; Wilson & Dufour, 2002). Researchers have extensively examined what underpins CCE, its evolutionary origins and whether it exists outside of humans, driven by the aim of understanding the success of our species (reviewed in Caldwell et al., 2020; Dean et al., 2014; Tennie et al., 2009; Vale, Carr, et al., 2017).

CCE has impacted all human cultures across a diverse set of domains, including knowledge of foraging subsistence-oriented societies (Reyes-García et al., 2016; Salali et al., 2016), tool complexity in Tasmania (Henrich, 2004), hunting techniques of the Central Inuit in the Canadian Arctic (Boyd et al., 2013), structure and transmissibility of language (Kirby et al., 2008) and folktales (Acerbi et al., 2017; Tehrani, 2013) as well as religious practices (Norenzayan & Gervais, 2011). Key to CCE is the social dissemination of improved skills, knowledge, behaviors and artifacts within groups that lead to adaptive changes. Many nonhuman animal species are capable of social learning (learning by directly observing other organisms or the byproducts of their actions, Heyes, 1994) and some even show cultural traditions – group typical behavior that is socially transmitted (Laland & Janik, 2006). Population-specific foraging, tool use, social and communicative behaviors across taxa such as apes, monkeys, birds, and cetaceans are a result of the social transmission of these behaviors within groups (Allen et al., 2013; Aplin et al., 2015; Gruber et al., 2015; van Schaik, 2003a; Whitehead & Rendell, 2014; Whiten, 2019; Whiten et al., 1999). Yet, concrete evidence for cultural change in a way that cumulatively improves existing skillsets over time outside of humans remains elusive (Caldwell, 2020; Mesoudi & Thornton, 2018).

In particular, it may be challenging to show adaptive change over historical time in nonhuman species, much less in experimental contexts, because their cultural traditions either lack artifacts (i.e., are behavioral) or show less apparent change. However, while no other species have anything close to our sophisticated technologies, symbolic languages and institutions, as we present here, more basic forms of cultural improvement may be present across a range of taxa (Wilks et al., 2021). Understanding these will help us understand the evolutionary history of humans’ impressive capacity for cumulative culture.

Indeed, conclusions about the uniqueness of cumulative culture in humans may be precluded by a lack of consistency in both the definitions and criteria used to determine CCE and the methodology used to examine it across different animal species. Concerns regarding how improvement is defined, whether it is measured the same way across studies, and whether some criteria are empirically testable are of increasing interest to many researchers (Caldwell, 2020; Mesoudi & Thornton, 2018).
et al., 2020; Mesoudi & Thornton, 2018; Miton & Charbonneau, 2018; Reindl et al., 2020; Schofield et al., 2018). After all, inconsistent use of definitions means studies are not comparable, making it difficult to assess how CCE may differ across contexts, much less across different species. The latter of these is essential to understand the evolutionary history of CCE and how it came to be such a dominant force in humans.

The primary goal of this paper is to provide a critical assessment of conceptual and methodological advances within the field of CCE with respect to non-human animals. We begin with a review of how definitions to assess CCE have evolved, particularly concerning current debates surrounding scientific parsimony. After considering how these debates and developments have impacted our ability to attribute CCE to nonhuman animals, we make specific recommendations about which criteria should be used. We then use our recommended criteria to assess the existing evidence from both wild and captive studies of nonhuman animals, before discussing what research is needed going forward to advance the field. Throughout the manuscript, we focus largely on nonhuman primates as, based on their phylogenetic and cultural proximity to humans, they have been the focus of the majority of nonhuman animal CCE research - particularly chimpanzees (Whiten, 2019). We also argue that the field must continue to expand beyond primates however, and thus, where relevant, we include literature from non-primates.

The evolution of definitions and assessment of cumulative cultural evolution

Early characterizations of cultural improvement were relatively broad, focusing on the ‘ratchet effect’ - a process whereby the complexity of cultural traits (such as skills, knowledge, technology, and customs that are transmitted socially within communities) are improved through generations of innovations and their subsequent social transmission (Tomasello et al., 1993). By this definition, cycles of modifications to, and ensuing social dissemination of, cultural traits allow the ratcheting of knowledge with little/no ‘backward slippage’ (loss of knowledge or reversion to previous, less efficient behaviour). These cycles were suggested to culminate in products that no single individual could invent from scratch within their lifetime (Boyd & Richerson, 1996; Dean et al., 2012, 2014). Attribution of human uniqueness of CCE was based on the assumption that only humans possess the required capacity for high fidelity social learning to faithfully maintain improvements over time and prevent backward slippage. This view continues to be contentious, however, as (i) similarly high-fidelity social learning may also be present in some nonhuman animals (Boesch et al., 2020; Horner & Whiten, 2004; Loretto et al., 2020; Musgrave et al., 2020; Whiten et al., 1996, 2009) and (ii) experimental research with humans has shown that cumulative technological improvement can occur without high fidelity social learning (Caldwell & Millen, 2009; Zwirner & Thornton, 2015, although see Wasielewski, 2014).

A large body of subsequent research has since highlighted theoretical and experimental challenges to these early definitions (Caldwell et al., 2020; Dean et al., 2014; Haidle & Schlaudt, 2020; Mesoudi & Thornton, 2018; Miton & Charbonneau, 2018; Reindl et al., 2020; Schofield et al., 2018). For instance, while almost all researchers agree that CCE requires cultural traits to show improvement over time by a process of building upon previous generations (Caldwell, 2020; Davis et al., 2016; Dean et al., 2014; Legare, 2017; Mesoudi & Thornton, 2018; Sasaki & Biro, 2017; Tennie et al., 2009; Weston & Jackson, 2018; Whiten, 2017a), improvement is not always defined or measured consistently. From an evolutionary perspective, improvement should represent adaptive modifications over historical time (Boyd et al., 2011; Derex & Mesoudi, 2020; Fay et al., 2019; Henrich, 2004,
2015; Winters, 2019), yet experiments only measure short term improvement rather than long term adaptive changes. There is thus a disconnect between what, globally, we aim to know about cumulative improvement, and what we can measure in laboratory settings.

Experimentally, improvement is typically described as an increase in complexity or efficiency of a behavior (Dean et al., 2014), but these constructs themselves are inconsistently measured (Schofield et al., 2018). Complexity, for example, can refer to the number of behavioral steps or techno-units required to achieve a goal (Boesch et al., 2019; Davis et al., 2019), differences in the hierarchical structure of behaviors (Boesch et al., 2019; Byrne & Byrne, 1993) or overall level of difficulty (Caldwell et al., 2017). Efficiency can be measured in terms of speed, learnability, a reduction of behavioral steps, calorie intake, convenience, or security (Davis et al., 2016; Schofield et al., 2018; Stewart et al., 2007; Yamamoto et al., 2013). Given that studies differ substantially in how they measure these constructs, we suggest clearly stating units of measurements will provide clarity, help to refine cross-study evaluations and will allow a clearer assessment of the impact of improvement on other outcomes such as adaptive value, payoffs, or productivity (Schofield et al., 2018).

A broader conceptual issue is that definitions of CCE may refer to both cultural processes (cumulative improvement) or cultural products (behavioral outputs) (Reindl et al., 2020). Process-based definitions, as defined by Reindl and colleagues, require improvements in complexity or efficiency over generations of social transmissions but are not concerned with the actual quantity of complexity or efficiency of the final product (McGuigan et al., 2017; Mesoudi & Thornton, 2018). Product-based definitions, conversely, define the improved complexity/efficiency as that which goes beyond what a naive individual could invent within their lifetime (Boyd & Richerson, 1996; Reindl et al., 2017, 2020; Tennie et al., 2016). Researchers generally elect for either process or product-based definitions and this has significant consequences for the underlying cognitive mechanisms involved (Rawlings & Legare, 2021), the extent to which CCE is uniquely human or not, and even whether different cohorts of the same species share CCE (e.g., children, adults, or both: Reindl et al., 2020).

Using different criteria also moves the goalposts concerning what is classified as CCE and hinders our capacity to make fair cross-species comparisons, and thus to fully understand the evolutionary origins of CCE. These issues also present important philosophical questions regarding comparisons of CCE behavior between humans and nonhuman animals. To what degree should researchers define phenomena such as CCE based on their experimental tractability (i.e., process-based criteria)? Should we limit definitions to what is testable? Study designs hinge on such questions; without tethering our definitions to empiricism, we cannot scientifically evaluate them. Yet, capturing a phenomenon based on what is possible to empirically examine risks overlooking very important aspects of the behavior (i.e., product-based criteria).

Further, the use of the criterion that CCE leads to traits (such as adaptive behaviours, skills or artifacts) beyond which an individual can invent within their lifetime is increasingly being questioned because it is impossible to experimentally examine in long lived species such as humans and apes (Caldwell et al., 2016; Mesoudi & Thornton, 2018; Miton & Charbonneau, 2018; Schofield et al., 2018). This criterion emphasizes the power of collective knowledge and provides a broader framework for CCE (in addition to measures of improvements in complexity and efficiency). Experiments typically reduce generations or lifetimes to a matter of minutes or hours, and thus cannot adequately or ethically assess whether a solitary...
individual can invent a product of CCE in their lifetime (Mesoudi & Thornton, 2018; Miton & Charbonneau, 2018).

In an attempt to overcome some of the definitional discrepancies outlined above, Mesoudi and Thornton (2018) recently described both core and extended criteria for attributing CCE to a population (which can be applied to assess CCE for any behavior found in a given population, of any animal species). The core criteria, representing the minimum requirements for a population to display CCE, include (i) a change in behavior (or product of behavior, such as an artifact), typically due to asocial learning, followed by (ii) the transfer via social learning of that novel or modified behavior to other individuals or groups, where (iii) the learned behavior causes a performance improvement, which is a proxy of genetic and/or cultural fitness, with (iv) the previous three steps repeated in a manner that generates sequential improvement over time. According to this account, criteria i and ii equate to cultural change (i.e., the transmission of innovations), the addition of criterion iii would be evidence of cultural evolution (where the innovation causes improvement), and to distinguish cultural evolution from cumulative cultural evolution, criterion iv is required (where this improvement repeats over time). The extended criteria, which may or may not be present and may be scaffolded by different socio-cognitive mechanisms than the core criteria, include i) multiple functionally dependent cultural traits (for example, when new innovations are contingent upon previous ones), ii) diversification into multiple lineages (such as different types of projectile tools for hunting), iii) recombination across lineages (combining cultural traits to develop new ones), and iv) cultural exaptation (i.e., a change in a cultural trait’s function), and v) cultural niche construction (i.e., where cultural traits impact the selective environments of other biological or cultural traits). Criterion such as these provide a valuable metric in which to classify candidate behaviours as evidence of CCE or not.

Here we build upon these recent criteria proposed by Mesoudi & Thornton (2018) to define CCE from a comparative perspective. From a practical standpoint, we believe a criterion needs to be empirically tractable and so we focus on criteria that can be assessed and experimentally tested in nonhuman species. To this end, the CCE criteria we think are most suitable for nonhuman animal testing (with particular consideration of nonhuman primates) would be core criteria i-iii (the transmission (ii) of a new behaviour (i) which causes performance improvements (iii)). We further suggest that improvements should be those that surpass what an individual can accomplish when exposed to the same conditions (iv), although we recognize that determining what is needed for conditions to be the same can be difficult. This distinguishes cases of cultural change and cultural accumulation (where behaviors are added to repertoires, but they are not more complex; Dean et al., 2014) from CCE, which requires measurable improvement. It is important to note that although prior cultural knowledge can influence the inventive capacities of individuals, testing whether group behaviors exceed individual achievements provides a useful benchmark to distinguish CCE from cultural evolution. This means, to our definition, CCE could be attributed to novel behaviors which are improvements upon previous versions (beyond an individual’s abilities), and which are transmitted to others.

We propose that an extended criterion should include that these steps are repeated over generations (which is core criterion iv from Mesoudi and Thornton, 2018), because for some animal species, including nonhuman primates, assessing generational improvements is highly difficult (Caldwell et al., 2020). In captivity, generational transmission chain experiments (where information is passed along chains of individuals and improvement is measured) are logistically problematic, requiring access to sufficient sample sizes and facilities that permit
this type of testing (issues discussed in detail later). Obtaining generational data from wild populations requires longitudinal observations, which is equally, or more, difficult. Thus, while we agree that evidence of invention and social spread leading to improvement (criteria i-iii) being repeated over time would be optimal, it is not always feasible to measure. Therefore, while we find the use of criterion extremely valuable for classifying behaviours as evidence of CCE, we suggest that a more inclusive definition, applicable to diverse taxa may also help further our understanding of the evolution of CCE. Using these criteria, we now present and assess evidence of CCE in non-human primates and, where applicable, other species, based on studies of wild and captive populations (Table 1).

Cumulative culture in non-human animals: Evidence from wild and captive populations

As evidence for culture in nonhuman animals grows (Aplin et al., 2015; Gruber et al., 2019; Laland & Galef, 2009; Whiten, 2017b, 2021; Whiten et al., 1999), researchers are examining whether nonhumans are capable of CCE. Various approaches have been developed, from documenting complex traits in wild populations, sifting through historical datasets, to experimentally manipulating behavioral complexity and efficiency in laboratory-controlled settings.

CCE in non-human primates: Evidence from the wild

Reports of group-level traditions in wild nonhuman primate populations, similar to human culture, are now well-documented (see Boesch et al., 2019; McGrew, 1992; Whiten et al., 1999 for chimpanzees [Pan troglodytes]; Hohmann & Fruth, 2003 for bonobos [Pan paniscus]; Perry, 2011 for capuchins [genus Cebus]; Robbins et al., 2016 for gorillas [Gorilla gorilla]; van Schaik, 2003 for orangutans [Pongo]). Some of these traditions, such as chimpanzee nut-cracking, vary in their complexity across populations, leading some to suggest they have undergone successive refinements and that they, therefore, represent examples of CCE (Boesch, 2003).

Perhaps the most famous traditions include chimpanzees’ community-specific styles of constructing and using tools during extractive foraging that are not easily attributed to ecological or genetic dissimilarities. This strengthens the conclusion that these inventions (criterion (i)) spread by social learning (criterion (ii)). Taking nut-cracking as one example, chimpanzees in Bossou, Guinea, and Taï Forest, Côte d’Ivoire (Whiten et al., 2001) display sophisticated procedures of using several tools to open the hard-shelled nuts found in their habitats (see also Coelho et al., 2015 and Eshchar et al., 2016 for similar cases of nut-cracking in capuchin monkeys). These two populations use wooden and stone hammers in conjunction with an anvil, and more rarely, with a third component of a stone to stabilize the anvil (Boesch et al., 2019; Luncz & Boesch, 2014; Matsuzawa, 1994; Sugiyama, 1997; Sugiyama & Koman, 1979). Researchers have suggested that composite (two or more tools used together to achieve a goal, here hammers and anvils) and compound tool use (combining elements to make a single unit, here the wedge added to the anvil) represent elaborations of the simpler percussive act of hitting (soft) shells with one’s arm, or a single tool, perhaps indicative of some form of CCE (Boesch, 2003). Neighboring groups’ specificity in their material selections also have consequences for the efficiency of these cultural traits, as measured by foraging speed and the number of actions required to crack open nuts (Luncz et
al., 2018). However, while community differences in complexity and efficiency may indicate trait modification, there is a lack of evidence of directionality to confirm that simple trait variants preceded more sophisticated ones. Without such evidence, it is difficult to assert that the behavior has undergone beneficial modifications (uncertainty of criterion (iii)).

A further concern here is whether the learned behaviors are beyond what may be invented by a solitary chimpanzee in comparable conditions (criterion (iv)). This has also proven difficult to test experimentally. Chimpanzees in the wild begin to successfully crack nuts by approximately 3.5 years of age (Inoue-Nakamura & Matsuzawa, 1997) and master the behavior when they are 5-6 years old (Estienne et al., 2019). Moreover, there appears to be a sensitive period of between 3.5 and 10 years of age during which they can develop the skill (Biro et al., 2003; Inoue-Nakamura & Matsuzawa, 1997; Matsuzawa, 1994). Both the extended learning period and the limited age during which learning appears to occur make it difficult to study in laboratory settings. A recent study found no evidence that 13 naive captive chimpanzees, aged 10-52, would spontaneously crack nuts with tools after 92 hours of practice (Neadle et al., 2020), which may suggest it is not easily invented. However, the authors also reported no evidence that chimpanzees could socially learn this behavior, perhaps suggesting that the chimpanzees either received too little exposure or were past the sensitive period (Neadle et al., 2020). Indeed, in a group of 2–6-year-old chimpanzees, individuals 3 years and over learned to crack open nuts on a stone anvil after observing conspecific demonstrations within days (Marshall-Pescini & Whiten, 2008a). This contrasted an age-matched control group that showed no such learning until they received social demonstrations (Marshall-Pescini & Whiten, 2008a), highlighting the importance of social learning in the development of nut-cracking technologies during early chimpanzee ontogeny.

Another case for CCE in chimpanzees has been made for their termite fishing behaviors. In a large-scale study, Boesch et al. (2020) collated data from multiple chimpanzee communities, documenting a detailed analysis of the technical elements they use to access termites. Overall, there was impressive diversity in the techniques employed, with 38 different elements involved, 30 of which could not easily be explained by community ecological conditions. Some newly observed community-specific elements (circumstantial evidence of inventions, criterion (i)), such as orientation, for instance by leaning on their elbow or lying on their side, were suggested to have spread by social learning (criterion (ii)). Groups further differed in the specific combinations of elements they used, some of which include tool modifications (e.g., making of brush-tipped probes) that can lead to improvements in behavioral efficiency (Sanz et al., 2009). Such community-specificity in the complexity of the behavioral sequences used to extract termites, as with the nut-cracking example, resemble cumulative additions to behaviors (see Boesch et al., 2020; Dean et al., 2014). Without observation of its inception and development over time, however, we cannot be sure that these behaviors evolved in complexity or efficiency (Boesch et al., 2020; although see Sanz et al., 2009; criterion (iii)), nor that they exceed individuals’ ability to invent (criterion (iv)). Though individual discovery of the exact community-specific styles of termite fishing (involving numerous behavioral elements) seems improbable, we cannot rule this out with current data.
To circumvent speculation over the developmental course of recently observed traits, researchers have examined long-term datasets spanning multiple generations (thus meeting our extended criteria). Schofield et al. (2018) reviewed over 60 years of observations of Japanese macaques’ (Macaca fuscata) food-washing behaviors. This dataset included the first instance of sweet potato washing followed by wheat washing (criterion (i)), allowing researchers to track the spread (criterion (ii)) and progressive enhancements to these behaviors since their inception (criterion (iii)). Various modifications were observed which may represent cumulative increases in complexity and efficiency (Schofield et al., 2018). Similar longitudinal observations have been conducted with white-faced capuchins (Perry, 2011; Perry et al., 2003, 2017), tracing the development and diffusion of a number of social conventions, but there is question over whether they have improved over time, and their reinvention suggests they do not surpass what individuals can invent for themselves (Dean et al., 2014).

Tracing innovations and behavioral change across generations is extremely time-intensive, but promises to shed light on the cumulative cultural abilities of other species. Although there remains a question over the role that social learning (criterion (ii)) plays in the transmission of behaviors such as these, as it is not directly observed (Caldwell et al., 2020; Fiore et al., 2020), we hope that continued observations of wild communities’ inventions and their spread will shed more light on whether CCE is shared with other animals.

CCE in non-human primates: Experimental evidence

To our knowledge, it was not until 2008 that the first experiment was conducted to explicitly test whether chimpanzees display CCE. Marshall-Pescini & Whiten (2008b) presented 11 young chimpanzees with a honey-dipping task which could be solved using the relatively simple act of dipping a tool into a contraption to gain honey or a more complex action of first removing a bolt before leveraging open a lid to reveal a large compartment of honey and peanuts. Overall, five chimpanzees demonstrated the ability to dip for honey, either personally discovering the action during baseline trials (criteria (i)), or acquiring it following social demonstrations (criterion (ii)). The more complex action, by contrast, was not acquired, despite participants receiving a substantial number of demonstrations (not meeting (iii)). This was not because the more optimal behavior was too challenging for them - task naïve chimpanzees easily discovered the complex act through personal exploration (not meeting criterion (iv)). The authors thus raised the proposition that the chimpanzees displayed behavioral conservatism, such that their known behavior (‘dipping’) interfered with the adoption of the more productive alternative that was otherwise in their ability to acquire (Marshall-Pescini & Whiten, 2008b) - a hypothesis that inspired further studies (which we discuss below).

The next study of CCE in nonhumans was by Dean and colleagues (2012), who took a comparative approach on a larger scale, comparing the performance of group-housed chimpanzees and capuchins to nursery aged children. Experiments that make such cross-species comparisons are particularly important as they can identify trait similarities and differences across species and shed light on their evolutionary histories (Liebal & Haun, 2012). All three species were posed an extractive foraging task similar to Marshall-Pescini and Whiten (2008), but with three sequential and increasingly difficult steps that revealed progressively valuable rewards (Dean et al., 2012). Whereas chimpanzees and capuchins had difficulty in reaching complex task solutions, either by individually discovering them or by
learning socially from a proficient conspecific, many of the children quickly solved the task. Children also displayed several socio-cognitive processes related to their task success that were largely absent in the other primates, including sharing rewards (altruistic resource donation) and knowledge (via teaching), as well as copying the actions of group members in their group (imitation). High-fidelity social learning mechanisms, including imitation and teaching, support the transmission of particularly complex knowledge or cultural traits, for which other mechanisms, such as trying to recreate products through backward engineering (emulation), may be insufficient (Caldwell et al., 2017; Morgan et al., 2015) and are thus important contributors to the level of complexity a species can reach. This suite of psychological processes were hence suggested to be key in supporting children’s ability to transition to more sophisticated and rewarding techniques, enabling them to outperform the non-human participants (Legare, 2019).

A limitation of this study was the absence of an asocial control to ensure that individuals could not independently solve the task (criterion (iv)). To address this, a new population of children were tested on the same task previously presented to groups, to see if they could solve the puzzlebox through personal exploration (Reindl et al., 2020). Notably, 9 out of 35 children discovered all three task levels without social demonstrations. Children in this study failed to meet the authors’ product-orientated definition of cumulative cultural evolution, which requires the behavioral products of group tested participants to exceed those of individuals (Reindl et al., 2020; criterion (iv)). Instead, the authors proposed that Dean et al. (2012) had shown elementary cumulative cultural learning that was process orientated, involving change through invention and social learning but without requiring the end result (task level) to surpass what individuals could invent.

Transmission chains, replacement methods and open diffusions

Subsequent studies have experimentally examined CCE in other animals, including birds (Sasaki & Biro, 2017) and baboons (Claidière et al., 2014), using transmission chains to closed group open diffusion type tasks (described below) to reveal evidence for cultural improvement. Transmission chain designs involve one participant, or ‘model’, providing task-relevant social information to a naïve observer before they perform the same task, with the latter then acting as the model to a new individual in the chain, and so on (Whiten & Mesoudi, 2008). Using this method, baboons were given a pattern recognition task in which the output of the previous participant formed the stimuli for the following participant in the chain, and so on (Claidière et al., 2014, also see Saldana et al., 2019). Over time, pattern reproduction improved (increased performance, criterion (iii)) as there emerged structure in the lineages and across ‘generations’, similar to how human languages evolve (Kirby et al., 2008). Importantly, repeated exposure to one’s own learning output did not yield the same level of change, suggesting that chains comprising of different individuals were responsible for improving the learnability of the task (meeting criterion (iv)).

Similar improvements across transmission chains have been reported in homing pigeons (Columba livia domestica), where the homing routes of birds later in chains were more efficient than earlier ‘generations’ and solo individuals (Sasaki & Biro, 2017). ‘Generational’ change was explored by first pairing a bird with an established homing route with a naïve flight partner. Once this naïve partner had flown with their partner, they became the
experienced participant paired with the next naïve subject in the chain, until five successive
generations were formed. Birds in the transmission chain showed measurable improvements
in their routes, such that the fifth generation’s routes were 1.2 km shorter than those of the
first-generation (meeting the extended criterion of generational improvements). This
improvement over time was also more pronounced than for solo birds repeating the same
number of flights, which produced only a 0.05 km reduction in their route - therefore meeting
criterion (iv).

Improvements in behavioural efficiency have also recently been documented in great tits
(Parus major) using the ‘replacement’ method. This method is similar to transmission chains
except each generation is a group of individuals instead of a dyad (Chimento et al., 2021).
Here, birds could gain a reward by either pushing open a door in a demonstrated but
inefficient direction, or they could innovate the efficient, quicker alternative by pushing the
door from the opposite side. For birds that were in the ‘static’ control condition, repeatedly
testing the same groups birds, the efficient behaviour was discovered (easily invented, thus
not meeting criterion (iv)) but was not widely adopted. For the birds in the ‘turnover’
condition, repeatedly replacing two of the six birds in each ‘generation’ with naïve ones, the
efficient alternative was not just invented (criterion (i)) but also widely adopted (social
transmission: criterion (ii) and improvement (iii)). Thus, naïve birds appeared key to
populations’ transition to the improved (quicker) solution by overcoming experienced
individuals’ tendency to stick with their practiced methods (Chimento et al., 2021).

As transmission chains and the replacement method are usually not feasible with
chimpanzees (discussed in detail below), recent studies have employed ‘closed group’ or
‘open diffusion’ task designs (Whiten & Mesoudi, 2008). Typically, these tasks involve
training a chimpanzee on a behavior that subsequently serves as a model to the rest of their
group. Davis et al. (2016) tested whether groups of chimpanzees would relinquish a
practiced, inefficient solution for a more optimal (quicker) one introduced by a trained
demonstrator. As with Marshall-Pescini & Whiten (2008b), chimpanzees generally
perseverated on their inefficient behaviors, and only switched to the more efficient behavior
when their practiced technique became almost impossible to perform (see also Harrison &
Whiten, 2018; Hopper et al., 2011; Hrubesch et al., 2009; Manrique & Call, 2011, but see
Jacobson & Hopper, 2019). This indicated a degree of conservatism in chimpanzees, similar
to that observed in great tits (Chimento et al., 2021), where known behaviors interfered with
the adoption of an alternative, more efficient one that was shown to be within their ability to
invent (Davis et al., 2016; not meeting criterion (iv)). Further investigation indicated that
behavioral flexibility was linked to the solution complexity and payoff incentives, as
chimpanzees switched from a simple, known behavior to a newly observed and better paying
solution of equal difficulty, but more rarely relinquished known behaviors when they were
complex or were of similar difficulty but not higher paying (Davis et al., 2019). Abandoning
an act that may have taken considerable effort or practice to learn may show reluctance to
forgo well-rehearsed techniques or habits that were known to achieve a goal and may explain
behavioral stasis once a certain level of complexity is reached.

Similar closed group designs have been used to examine whether chimpanzees can socially
learn more complex (and efficient) solutions from a conspecific demonstrator (Vale, Davis, et
al., 2017). Chimpanzees acquired behaviors that surpassed what individuals invented in the
absence of any social information (in comparable conditions). However, groups of naïve chimpanzees appeared to pool their skills to discover the same behaviors. One individual discovered part of the behavioral sequence of interest (unscrewing a stop valve to make a functional straw), before a second individual used their discarded tool to obtain a juice reward; a sequence that spread to others albeit at a slower rate and attained by fewer individuals than occurred in the model seeded groups. These data therefore only provide suggestive evidence for cumulative advancements in this technological (tool deconstruction) task as the role of social learning was not clear cut (criterion (ii)). The same population of chimpanzees was recently tested on a task that afforded participants to construct, rather than deconstruct, tools to gain progressively valuable rewards (Vale et al., 2020). The complexity of solutions observed in groups did not surpass those discovered by chimpanzees tested individually indicating a lack of CCE (not meeting criterion (iv)). The differences in findings across studies maybe be explained by the type of tool modifications involved (i.e., construction/deconstruction), a point we discuss in more detail below.

Chimpanzee invention

Whether the complex behaviors observed in wild populations are cumulative has also been tested by exposing naïve, captive populations to the ecological conditions thought to be necessary for the behavior to occur. The logic is that if animals discover the same behaviors as their wild counterparts, we can conclude that (i) action copying was not necessarily required for the inception of the behavior and (ii) it does not surpass an individual’s capacity to develop through personal exposure to the right conditions. This work, by Tennie and colleagues, has focused on several wild chimpanzee tool use behaviors, including pestle pounding, algae scooping, and nut-cracking (Bandini & Tennie, 2017, 2019, 2020; Neadle et al., 2020). In all cases except nut cracking, chimpanzees were able to re-invent the behavior under the conditions provided in captive settings, suggesting they are within an individual’s capability to invent (their ‘Zone of Latent Solutions’ or ‘ZLS’, Tennie et al., 2009, 2016). There are several challenges to the ZLS, however. It is difficult to be certain that these animals, some of whom were wild born, were truly naïve to these behaviors - the researchers rely upon keeper and caretaker testimony to document the animals’ experience (Bandini & Tennie, 2017). A further concern is the omittance of complex components of these behaviors that are typically needed in the wild. For example, the processes of finding, selecting, and fashioning the correct tool for the task are not typically required in captive tasks, where the chimpanzees are instead provisioned with useable tools from the outset (Bandini & Tennie, 2017). Notably, the objective of the task was to re-create the act, as a test for the role of form-copying in the examples from wild chimpanzees and did not claim to be re-creating the entire behavior. For example, in an ‘algae scooping’ study, chimpanzees were provisioned with a stick to fish a floating piece of bread out of water, whereas in the wild, chimpanzees must first identify an appropriate stick (smooth or hooked) of adequate length, detach it, and even modify it by the removal of smaller branches or fraying the ends (Boesch et al., 2017). Thus, while these studies suggest that at least some components of these behaviors exist in a ZLS, we encourage future studies to examine whether whole behavioral sequences can be re-invented, rather than just part of the action.

Research on CCE in nonhuman animals both from the wild and captivity is suggestive, however, it is not conclusive evidence of CCE. At present, notwithstanding evidence in non-primate species that we discuss further below, one of the most convincing cases from wild
data is chimpanzee nut-cracking. This is because it can take years to master (Inoue-Nakamura & Matsuzawa, 1997), is not easily re-invented by naïve individuals (Neadle et al., 2020), and involves social learning (Marshall-Pescini & Whiten, 2008a). However, chimpanzees have been cracking nuts for thousands of years (Mercader et al., 2007) which raises questions over whether improvement has been made to this behavior - over millennia. Nevertheless, the behavioural change that might have predated the available archaeological record is also unknown (Whiten, 2021). Researchers working with captive populations, or performing field experiments, have documented improvements in behavioral efficiency across transmission chains of multiple participants (Chimento et al., 2021; Claidière et al., 2014; Sasaki & Biro, 2017). Open diffusion experiments have also begun to show promise as a way to assess CCE, revealing modest increases in behavioral complexity (Vale et al., 2017, but see Vale et al. 2020), and have created the opportunity to isolate some of the socio-cognitive factors that may play a role in species’ extent for CCE (Davis et al., 2016; Dean et al., 2012). Many of these studies have also highlighted some of the difficulties of studying CCE and associated methodological limitations they can impose in nonhumans (a subject discussed later), including assessment of CCE in non-primates.

Cultural evolution and CCE in animal song

Aside from primates, there is growing evidence for cultural evolution, and perhaps CCE, particularly in the vocal displays of other animals. Though cultural transmission of vocalizations has been reported in diverse species (reviewed in Garland & McGregor, 2020), here we focus two pertinent cases of potential CCE, in humpback whales and zebra finches, that show largescale cultural shifts over time or change across laboratory learning generations.

Among whales, the songs produced by humpback males (Megaptera novaeangliae) are one of the most complex. These songs are long and hierarchically structured vocalizations, and populations of whales develop their own variants, or ‘local dialects’, through social learning (Noad et al., 2000; see Barker et al., 2021 for a recent example in naked mole rats and Watson et al., 2015 for a case in chimpanzees). Although a variety of animals display local dialects, few studies have explored whether animal vocalizations change over time. An exception is the songs of humpback whales that have now been subject to decades of study. Detailed recordings in the South Pacific have now revealed that their dialects evolve and change in ‘waves’, as song variants socially transmit (criteria (i & ii)) eastward through neighbouring populations (Garland et al., 2011). The scale of this cultural evolution is also vast, occurring across the Pacific Ocean basin, with songs repeatedly changing every few years (Garland et al., 2011; Noad et al., 2000). There is some indication that song complexity also changes, increasing as the songs evolve (criterion (iii)), but also decreasing in cases where complete song ‘revolutions’ (replacements) occur (Allen et al., 2018). Although we do not yet know the complexity of humpback whale song in the absence of social information (uncertainty of criterion (iv)), which can be difficult to test in such large mammals, the repeated cultural change that propagates through multiple populations every few years (Garland et al., 2011) shows striking parallels to human CCE.

Song development in zebra finches has been tracked using similar experimental methods used to study CCE in humans. Typically, young male zebra finches learn their song by copying adult males (criterion (ii)), developing wild-type, local variants, with some
individual variability (indicative of criterion (i)). When deprived of this experience, however, young males develop a distinguishable, relatively unstructured ‘isolate’ variant, allowing song ‘recovery,’ or development, to be traced. Fehér et al. (2009) paired juveniles finches with isolate ‘tutor’ birds, before placing learners in transmission chains (pupils became the tutors for the next generation and so forth). Similar to whale song, finches’ songs changed over learning events, recovering from isolate versions as similarity to the wild variant increased in just a few generations (suggestive of (iii), but unknown if wild-type song represents an improvement). Furthermore, songs of later generations differed from the isolate songs, which is indicative of CCE (criterion (iv)). These findings mirror results from transmission chain studies with human spoken language, where learnability and structure improve with iterated learning events between individuals (Kirby et al., 2008).

As these two examples illustrate, animal songs can change over time through repeated learning events between individuals, constituting interesting examples of cultural evolution, if not cumulative culture. Indeed, these, along with the cases described in baboons (Claidière et al., 2014) and pigeons (Sasaki & Biro, 2017) meet most or all of our criteria for CCE.
<table>
<thead>
<tr>
<th>Study</th>
<th>Species</th>
<th>Observed behavior</th>
<th>Wild/captive/field experiment</th>
<th>Invention (Criterion (i))</th>
<th>Social transmission (ii)</th>
<th>Measurable improvement (iii) through (i &amp; ii)</th>
<th>Exceeds individuals’ discoveries (iv)</th>
<th>Repeated over generations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allen et al. (2018)</td>
<td>Humpback Whales (<em>Megaptera novaeangliae</em>)</td>
<td>Song structure</td>
<td>Wild</td>
<td>✓</td>
<td>C - Probable but no direct evidence</td>
<td>✓ - Cycles of increases and decreases in song complexity</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td>Boesch et al. (2020)</td>
<td>Chimpanzees (<em>Pan troglodytes</em>)</td>
<td>Termite fishing</td>
<td>Wild</td>
<td>C - Circumstantial evidence through newly documented behavioral elements</td>
<td>C - Probable but no direct evidence</td>
<td>C - Probable but no direct evidence</td>
<td>x</td>
<td>C - Circumstantial evidence documenting behavior in multiple generations</td>
</tr>
<tr>
<td>Chimento et al (2021)</td>
<td>Great tits (<em>Parus major</em>)</td>
<td>Puzzlebox: two levels of efficiency</td>
<td>Captive</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td>Claidiere et al. (2014)</td>
<td>Baboons (<em>Papio papio</em>)</td>
<td>Memorizing patterns on touchscreens</td>
<td>Captive</td>
<td>✓ - Change through transmission error</td>
<td>✓ - No social interaction, but met social learning definition of learning from the products of others behavior</td>
<td>✓ - If invention is a group process and social learning can be indirect</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Study</td>
<td>Species</td>
<td>Title</td>
<td>Complexity</td>
<td>Captive</td>
<td>Invention model seeded</td>
<td>✓?</td>
<td>Notes</td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Davis et al. (2016)</td>
<td>Chimpanzees <em>(Pan troglodytes)</em></td>
<td>Puzzlebox: two levels of increasing efficiency</td>
<td>Captive</td>
<td>Invention model seeded</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Davis et al. (2019). Study 2.1</td>
<td>Chimpanzees <em>(Pan troglodytes)</em></td>
<td>Puzzlebox: two levels of increasing complexity</td>
<td>Captive</td>
<td>Invention model seeded</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dean et al. (2012)</td>
<td>Children, chimpanzees &amp; capuchins <em>(Homo sapiens, Pan troglodytes, Cebus apella)</em></td>
<td>Puzzlebox: three levels of increasing complexity</td>
<td>Captive</td>
<td>✓</td>
<td>?</td>
<td>✓ - But only children reached hardest level</td>
<td>x - Reindl et al. (2020)</td>
<td></td>
</tr>
<tr>
<td>Feher et al. (2009)</td>
<td>Zebra Finches <em>(Taenyopygia guttata)</em></td>
<td>Song structure</td>
<td>Captive</td>
<td>✓</td>
<td>✓</td>
<td>C – Change toward wild-type song. Unknown if change represents improvement</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Garland et al. (2011)</td>
<td>Humpback Whales <em>(Megaptera novaeangliae)</em></td>
<td>Song structure</td>
<td>Wild</td>
<td>C - Indirect evidence through changing song</td>
<td>✓</td>
<td>C – Unknown if changes represent improvement (though Allen et al. 2018 document changes in complexity)</td>
<td>x</td>
<td>Repeated change over 11 years</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Study</th>
<th>Species/Mammal</th>
<th>Behavior</th>
<th>Location</th>
<th>Evidence Type</th>
<th>Wild Evidence</th>
<th>Captive Evidence</th>
<th>Historical Evidence</th>
<th>Circumstantial Evidence</th>
<th>Other Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Garland et al. (2017)</td>
<td>Humpback Whales (Megaptera novaeangliae)</td>
<td>Song structure</td>
<td>Wild</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hunt &amp; Gray (2003)</td>
<td>New Caledonian crows (Corvus moneduloides)</td>
<td>Tool use</td>
<td>Wild</td>
<td>C - Indirect evidence of diverse tools</td>
<td>C - Probable but no direct evidence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jesmer et al. (2018)</td>
<td>Bighorn Sheep (Ovis canadensis)</td>
<td>Migration routes</td>
<td>Wild (historical data)</td>
<td>x</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>C - Circumstantial evidence documenting behavior in multiple generations</td>
</tr>
<tr>
<td>Loukola et al. (2017)</td>
<td>Bumblebees (Bombus spp.)</td>
<td>Transporting objects</td>
<td>Captive</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Luncz et al. (2018)</td>
<td>Chimpanzees (Pan troglodytes)</td>
<td>Nut-cracking</td>
<td>Wild</td>
<td>C - Original invention not observed</td>
<td>C - Indirect evidence of variations between neighboring groups.</td>
<td>x - Differences in efficiency documented for neighboring groups, but no improvements to a single trait</td>
<td>x - But some evidence that this behavior is not reinvented by naive chimpanzees (Bandini &amp; Tennie, 2020)</td>
<td>x - But evidence is reported elsewhere (e.g., Mercader et al., 2007)</td>
<td></td>
</tr>
<tr>
<td>Luther &amp; Derryberry (2012)</td>
<td>White-crowned sparrow (Zonotrichia leucophrys)</td>
<td>Song frequency</td>
<td>Wild</td>
<td>✓</td>
<td>C - Probable but no direct evidence</td>
<td>✓</td>
<td></td>
<td></td>
<td>C - Circumstantial evidence documenting behavior in multiple generations</td>
</tr>
<tr>
<td>Study</td>
<td>Species</td>
<td>Behaviors</td>
<td>Location</td>
<td>Model Seeded</td>
<td>Direct Evidence</td>
<td>Replicated</td>
<td>Published</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------</td>
<td>--------------------------</td>
<td>--------------------------------------------</td>
<td>------------</td>
<td>--------------</td>
<td>----------------</td>
<td>------------</td>
<td>------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marshall-Pescini &amp; Whiten (2008a)</td>
<td>Chimpanzees (Pan troglodytes)</td>
<td>Nut-cracking</td>
<td>Captive</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marshall-Pescini &amp; Whiten (2008b)</td>
<td>Chimpanzees (Pan troglodytes)</td>
<td>Puzzlebox: two levels of increasing complexity</td>
<td>Captive</td>
<td>✓</td>
<td>✓ - Simple behavior only</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Perry et al. (2003)</td>
<td>Capuchin monkeys (Cebus apella)</td>
<td>Social rituals</td>
<td>Wild</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Price et al. (2009)</td>
<td>Chimpanzees (Pan troglodytes)</td>
<td>Tool combining task to reach rewards</td>
<td>Captive</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Sanz et al. (2009)</td>
<td>Chimpanzees (Pan troglodytes)</td>
<td>Chimpanzee termite fishing</td>
<td>Wild</td>
<td>x</td>
<td>C - Probable but no direct evidence</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Sasaki &amp; Biro (2017)</td>
<td>Homing pigeons (Columba livia)</td>
<td>Homing routes</td>
<td>Field experiment</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Schofield et al. (2018)</td>
<td>Japanese macaques (Macaca fuscata)</td>
<td>Food washing behaviors</td>
<td>Wild (historical data)</td>
<td>✓</td>
<td>C - Probable but no direct evidence</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Species</td>
<td>Task Description</td>
<td>Ethology</td>
<td>Meeting Criterion</td>
<td>Transmission Type</td>
<td>Within Group Experiments</td>
<td>Between Group Experiments</td>
<td>Notes</td>
<td></td>
</tr>
<tr>
<td>------------------------------</td>
<td>----------------------------------</td>
<td>---------------------------------------</td>
<td>-----------</td>
<td>-------------------</td>
<td>-------------------------------------------------------------</td>
<td>---------------------------</td>
<td>---------------------------</td>
<td>--------------------------------</td>
<td></td>
</tr>
<tr>
<td>Vale et al. (2020)</td>
<td>Chimpanzees (Pan troglodytes)</td>
<td>Tool use and tool modification task</td>
<td>Captive</td>
<td>✓</td>
<td>C - Probable transmission of simple, but not complex techniques</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Vale et al. (2017)</td>
<td>Chimpanzees (Pan troglodytes)</td>
<td>Complex tool behavioral sequence</td>
<td>Captive</td>
<td>✓</td>
<td>Invention model seeded &amp; invention in a non-seeded group</td>
<td>✓</td>
<td>✓</td>
<td>✓ - But partial discovery in asocial control individual</td>
<td>x</td>
</tr>
<tr>
<td>Williams et al. (2013)</td>
<td>Savannah Sparrows (Passerculus sandwichiensis)</td>
<td>Song segments</td>
<td>Wild</td>
<td>x</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Yamamoto et al. (2013)</td>
<td>Chimpanzees (Pan troglodytes)</td>
<td>Extracting juice from container</td>
<td>Captive</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Studies of potential CCE in nonhuman animals against our proposed criteria. ✓ denotes meeting the criterion, x denotes failing to meet criterion and C denotes circumstantial or indirect evidence of meeting the given criterion.
Evaluation of the methods and advancements

Despite the significant advances made over the past two decades, limitations remain for comparative research in terms of what can or cannot be assessed with nonhuman animals and their implications for CCE research. Specifically, issues concerning sample sizes, species comparability of motor and cognitive abilities, as well as participant demographics which may require addressing for the field to continue to move forward. We next describe these concerns and provide some suggestions for their resolution.

Sample sizes

Accessing sufficient sample sizes (comparable to human research) is difficult when studying many nonhuman animal species (Sjoberg, 2017). Research facilities, sanctuaries, and zoos often have limited physical spaces and/or facilities, and ethical guidelines, correctly, encourage reducing the number of tested nonhuman animals to the smallest possible number. This means the number of social groups, as well as the size of social and asocial samples are limited to relatively small numbers, and often participants have been subject to similar studies before (meaning they are not experimentally naive). In no small part because humans are detrimentally impacting nonhuman primate population numbers and behavioural diversity (Estrada et al., 2017; Kühl et al., 2019), there are similar sample size issues in the wild, where accessing and/or accurately tracking animals (or species) can be logistically difficult. The power to statistically detect effects is therefore difficult - particularly when using multiple comparisons - and generalizability of findings is weakened. Although concerns regarding replicability have led to many disciplines addressing issues of sample sizes (Loken & Gelman, 2017), tackling this in some comparative research is not easily solved given generational and/or multi-group studies demands relatively large sample sizes and many species of interest are not available in such numbers.

A potential way of increasing overall sample sizes and the numbers of groups is to pool data from multiple study sites, as has been done elsewhere (Altschul et al., 2019; Hopper, 2017; MacLean et al., 2014; Van Leeuwen et al., 2020; Weiss et al., 2007; Whiten et al., 1999). When taking this approach researchers should account for variables that may vary across sites, including exposure to humans, participant ages, testing experience, enrichment exposure, and group demographics (Altschul et al., 2019, discussed further below). An alternative option is to maximize individual-level data. This can be done using aggregated data over repeated trials (or ‘generations’) within individuals, which would, theoretically, significantly reduce the number of participants required (Caldwell et al., 2020; Claidière et al., 2018). Assessment of CCE here would involve exposing individuals to social information of varying success levels and documenting evidence of improved performance over trials. However, we also note that this approach, which may be better suited to certain animal species, requires careful study design, and researchers should consider task exposure effects (Caldwell et al., 2020).

Motor and cognitive abilities

A key methodological concern for any comparative study is to develop tasks that capture appropriate contextual validity for the species being examined. Multiple factors need to be considered, including whether study species have appropriate motor and cognitive capacities concerning administered tasks. Designing experimental paradigms, tasks, and apparatuses
that allow direct comparisons across animal species means considering whether specific actions are more difficult for one species than for another.

Tool use and puzzlebox tasks for CCE research have allowed researchers to make direct (Dean et al., 2012; Tennie et al., 2009) and indirect (Davis et al., 2016; McGuigan & Whiten, 2009; Reindl et al., 2016; Whiten, 2017b) inferences regarding the socio-cognitive mechanisms underpinning CCE in humans and chimpanzees. An important consideration, however, is whether the level of dexterity required for task success is more or less challenging for a given species. Motor diversity is a key predictor of innovative and individual problem-solving abilities in birds, nonhuman primates, and children (Diquelou et al., 2016; Griffin et al., 2014; Griffin & Guez, 2014; Keen, 2011), and thus species differences in this domain may limit any potential comparative conclusions. Tasks that require fine motor skills, for example, may be better suited to humans than to other animals or may require a substantial period of training for nonhuman animals to reach required criteria (Davis et al., 2019; Vale, Davis, et al., 2017). If significant training is required, this may also weaken comparisons to human studies in which participants require far less training for similar tasks. Animals learning tasks can also be required to observe humans (i.e., a heterospecific) face to face, requiring copying mirror images, while children are often adjacent to (i.e., a conspecific) experimenters, meaning they can use motor imitation.

Relatively, the type of actions we ask of animals may have important implications for the study of CCE. For instance, deconstructing tools may be more species-appropriate, and ecologically valid than constructing ones for tool-using species (Bania et al., 2009). This is because in the wild animals’ tools are typically made and modified using broadly destructive, rather than constructive, actions, such as detaching probe sticks or leaf tools, and trimming them (e.g., chimpanzees: Boesch et al., 2017; bearded capuchin monkeys: Mannu & Ottoni, 2009; New Caledonian crows: Hunt & Gray, 2003). At our study site, the National Center for Chimpanzee Care (NCCC), individuals also fashion tools by detaching materials with greater ease than those that require combinatorial actions, suggesting that this is not simply an artifact of what is readily available in the wild. In the first tool construction task conducted at this site, fourteen of 50 chimpanzees fashioned elongated tools by inserting one tool component into another (Price et al., 2009). However, very few chimpanzees beyond these have since demonstrated such skills, despite testing nearly the entire colony on various construction tasks (Neldner, 2020; Vale et al., 2016; Vale et al., 2020, Vale et al., unpublished data). For example, only three of 20 naïve chimpanzees recently tested on tool construction learned how to combine tools and they failed to use them as a functional tool to gain out-of-reach rewards (Vale et al., 2016). Learning to deconstruct object components to make a functional tool, or for exchange for higher valued rewards, in contrast, have been relatively prevalent in our colony (Neldner, 2020; Vale, Davis, et al., 2017; Vale et al., unpublished data). Creating tasks that are species-appropriate is therefore essential to avoid concluding that species lack certain skills when the outcome may be an artifact of specific task conditions (F. B. M. de Waal, 2016; Leavens et al., 2017).

Breeding in captive chimpanzee populations has been banned in the U.S. (and several other countries) since 2007 (Knight, 2008), meaning populations in many locations are aging and there are very few juveniles or young chimpanzees available for testing. When dealing with many aging captive nonhuman primate populations, we must consider the effects of their cognitive abilities, motivation, and participation in experiments that vary across the lifespan, as these will ultimately have consequences on the conclusions we draw. Whether animals attempt and persist at tasks, for example, can decline with age (Barbary macaques, Macaca
Sylvanus: Rathke & Fischer, 2020), while perseveration with known solutions or strategies can increase (rhesus macaques: Lai et al., 1995: for examples of aging and cognitive decline in other species, see Chapagain et al., 2020 for dogs; Kapellusch et al., 2018 for rats and Kwapis et al., 2020 for mice). Openness, linked with cognitive performance, also changes over the chimpanzee lifespan – with males in particular decreasing over adulthood (Rawlings et al., 2020). Candidate CCE behaviors such as nut cracking are cognitively demanding and can take years to master, and if aging populations are showing cognitive decline these factors will have significant implications for studies of animal learning. CCE requires behavioral change, individual or group level behavioral flexibility and the motivation to learn often complex new skills, and this may lead researchers to underestimate species’ CCE abilities. Sampling biases

A decade ago, psychologists acknowledged an overreliance on so-called WEIRD (Western, Educated, Industrialized, Rich, and Democratic) (human) samples, who often perform at the extreme ends of continuums of psychological studies (Henrich et al., 2010). The pace of human cross-cultural research has rapidly increased over recent years - including within the field of CCE. As a result, the field has made and continues to make, significant conceptual, theoretical, and methodological improvements through the assessment of diverse populations (Broesch et al., 2020; Hruschka et al., 2018; Kline et al., 2018; Nielsen et al., 2017). A similar argument regarding nonhuman animal samples was raised at the same time where individuals raised in captivity (i.e. Barren, Institutional, Zoo, And other Rare Rearing Environments - BIZARRE) may not represent their wild counterparts (Leavens et al., 2010). Yet, comparative research has yet to fully address the issue of sampling biases (Webster & Rutz, 2020), where individual and group level factors impact which species and participants engage in behavioral research. Despite extensive and valuable investment in studying wild populations, sampling biases have important implications for the generalizability of data and conclusions that can be drawn from CCE studies.

At the individual level, factors such as personality, dominance status, prior testing experience, social status, and rearing history are important causes of selection bias (Altschul et al., 2017; Brosnan et al., 2015; Herrelko et al., 2012; Hopper et al., 2014; Morton et al., 2013; Rawlings et al., 2020). Much of primate behavioral research is based on voluntary participation, for good reason, but this also means that individuals who enjoy testing and/or are high ranking enough are most likely to participate, and indeed are those as trained demonstrators – which impacts subsequent learning of behaviours in groups (Kendal et al., 2015; Vale et al., 2020; Watson et al., 2017). At the group level, nonhuman animals raised in captive contexts may not be representative of the species at large. Frequent caregiver interaction and exposure to enrichment activities or behavioral testing may skew behaviour and task performance, as does interaction with tasks through bars, making comparisons with wild or even other captive populations and children difficult (Haslam, 2013; Leavens et al., 2017). Conversely, captive environments are comparatively impoverished as a function of less stimulation, contextual diversity, and fewer group members, compared to natural conditions (Boesch, 2007). Further, in many research sites, nonhuman primate groups have been rearranged or moved locations, which can have a significant impact on social bonds and social-based research (Dufour et al., 2011; Schel et al., 2013), as can environmental uncertainty (Galef & Whiskin, 2004). Implementing steps to overcome selection bias is crucial for a fairer representation of the population intended to be examined (Morton et al., 2013; Webster & Rutz, 2020).
Here we reiterate the recent STRANGE framework proposed by Webster and Rutz (2020) which provides an assessment tool that researchers can use to identify if their sample is representative of the larger population of interest. Seven categories are identified, namely subject’s: Social background; Trappability and self-selection; Rearing history; Acclimation and habituation; Natural changes in responsiveness; Genetic make-up; and Experience (i.e., STRANGE). These categories highlight, for example, that some individuals may be more motivated to participate than others (‘Trappability and self-selection’) and more or less likely to solve particular tasks depending on their previous ‘Experience’ and ‘Rearing’ conditions. Researchers interested in CCE may find this a useful framework to assess samples’ representativeness and how it may be improved. For example, self-selection biases may be reduced by testing social groups, and for extended periods to overcome potential task monopolization, or by introducing multiple tasks or stimuli so that more than one subject can participate at any given time. Variation in subject rearing histories and experimental experience, if known, can also be controlled for either statistically or in the research design (Bandini, 2021; Neadle et al., 2020; Vale et al., 2020). Implementing comparable testing setups between species will also allow for fairer comparisons (McGuigan et al., 2017; Neldner, 2020).

The STRANGE framework is therefore a useful tool for researchers to present information about their study group which allows fairer conclusions on the generalizability of the findings. We also fully acknowledge it is not always possible to control the fact that a sample is STRANGE. Work on such samples has provided a wealth of key knowledge on the evolutionary basis of CCE, and with breeding bans across many institutions, nonhuman primate researchers are running out of opportunities to collect such data. We suggest in such cases that comparative researchers fully and openly acknowledge the sample biases. Doing so will avoid the extreme data collected from STRANGE samples to become interpreted as the default of the entire species.

Likewise, comparisons between the cultural learning strategies of human children and nonhuman animals have been a valuable line of research for establishing the ontogeny of CCE (Dean et al., 2012, 2014; Tennie et al., 2009; Vale, Flynn, et al., 2017). Yet, as with primates, while the field is beginning to assess diverse populations, the vast majority of experimental CCE research on children has been on WEIRD populations. This bias presents issues with conclusions that can be drawn when comparing children’s performance with other species, and researchers should be mindful of the human population in which they study. Tools such as STRANGE frameworks could usefully be applied to developmental research to improve assessment of the generalisability of findings, which may also have similar biases within populations. Children who have caregivers who actively consent to participate in scientific studies, and children who willingly participate in research, for example, can generate selection biases (Anderman et al., 1995). As with nonhuman animal studies, we encourage CCE researchers interested in development to evaluate, report and improve the representativeness of their samples.

The value of observational and experimental approaches

While much of our focus has been on experimental work, our intention is not to prioritize this method over observational approaches. Experiments are not feasible with certain species (e.g., large animals such as whales, Janson & Brosnan, 2013), and can lack external validity. Observational approaches, by contrast, afford documentation and evaluation of naturally occurring candidate CCE behaviours (Henrich, 2015; Noad et al., 2000; Reyes-Garcia et al.,...
2016), which can be highly informative of species’ natural behaviors as well as for planning experiments. The use of both approaches is required to make crucial, and complementary, contributions to the field of comparative CCE.

Bridging the gap between lab experimental work and observational research are field experiments, which have become an important tool for understanding how animal cultures emerge and spread within groups, in natural settings. Field experiments involve using experiments in wild populations of animals, often to study naturally occurring behaviours (Aplin et al., 2015; Biro et al., 2003; Gruber et al., 2009; Sasaki & Biro, 2017; Van de Waal et al., 2014). The core value of field experiments is thus that they provide experimental control while studying animals in their natural habitat. Indeed, arguably the strongest evidence for nonhuman CCE, based on our criteria, is the field experiment of navigational routes with homing pigeons (Sasaki & Biro, 2017). In many cases, they exceed the capacity of lab experiments because the question can be studied under a broader range of relevant contexts. Likewise, field experiments can also exceed the capacity of observational research due to the experimental control they provide. However, it is important to note that lab research typically allows more control and can usually offer additional manipulations. Field experiments can only be used to study what we know to examine, so we need observations to document behaviours occurring in their natural environment, and how behaviours we observe in the lab manifest in the wild.

The continued use of field experiments, alongside other experimental and observational approaches, will be crucial to documenting the scope of CCE in the animal kingdom in ways not previously possible. The development of sophisticated technology and advanced methodological approaches are providing new methods to collect rich data on CCE in diverse species. Lightweight trackers and camera traps have allowed researchers to detail migration routes and improvements in spatial navigation in wild animals. Advanced statistical approaches, such as network-based diffusion analysis (NBDA), mean that scientists can document the advent and spread of innovations across dynamic social networks (Allen et al., 2013; Hobaiter et al., 2014; Migliano et al., 2020; Wild et al., 2019). Such tools will be valuable for understanding how the social environment impacts CCE in natural conditions.

In humans, the vast majority of evidence for CCE in non-western populations remains based on ethnographic data (Henrich, 2004; Reyes-García et al., 2016; Salali et al., 2016), which is problematic from the perspective of understanding other species’ CCE because these results are not comparable to those generated in lab and field experiments or observational research in non-human species. Fortunately, the recent growth of cross-cultural research means that mechanisms such as social learning and innovation are being studied experimentally in a broader range of populations (Berl & Hewlett, 2015; Clegg et al., 2017; Legare, 2017; Neldner et al., 2017; Nielsen et al., 2014; Rawlings et al., 2019). We encourage cross-cultural research to complement ethnographic studies with explicitly experimentally CCE research in diverse populations, using ethical and equitable approaches (Broesch et al., 2020; Urassa et al., 2021). Additionally, we simultaneously encourage cross-cultural researchers to validate experimental tasks and paradigms to improve construct validity and thus quality of explanations (Broesch et al., 2020; Hruschka et al., 2018; Kline et al., 2018; Lew-Levy et al., 2020). These steps will be essential to for improving our understanding of how cultural diversity shapes the remarkable diversity of human CCE and how CCE evolved in humans and other species.

Directions for future research
We next discuss two key areas we believe are particularly pertinent to continuing to move the field of cumulative cultural evolution forward. We suggest that refining asocial control testing methods and continuing to broaden study species will be particularly beneficial for improving theory in future research.

Asocial control testing methods

Crucial to advancements in the field of CCE include the addition of asocial controls that examine an individual’s invention abilities, to provide a comparator for group-level achievements and transmission chain designs that allow ‘generational’ changes to occur across participants (Miton & Charbonneau, 2018; Reindl et al., 2020). For nonhuman animal research, however, these require individuals to be separated from their group, which may be undesirable for a social species and is not always possible for many species (particularly into individual isolation). Individuals may be reluctant to separate or testing facilities may have regulations on whether, and for how long individuals can be separated, and isolation can induce unnecessary stress upon participants (which can also reduce their willingness to participate or change their behavior). Transmission chain studies, which involve dyadic testing, require careful consideration of differences in dominance statuses between pairs of individuals. Moreover, researchers rarely consider wider contextual issues when comparing asocial to group performance. Individuals in asocial conditions may show increased vigilance compared to those in groups, and they lack opportunities for social facilitation that enhances activity based on the mere presence of others (Zajonc, 1965). Recent research shows that chimpanzees tested in social groups exhibit more diverse behavioral repertoires and had greater task success than participants tested in the absence of conspecifics (Vale et al., 2020). This suggests that social facilitatory effects and/or a larger pool of skills and motivations in groups can influence study results and that social settings should be considered by researchers as they design control conditions (see also Finestone et al., 2014).

These issues can prohibit researchers from using certain techniques or result in unbalanced study designs, where, for example, few individuals form the asocial controls (or come from smaller groups) compared to much larger social groups. Though this is often necessary, it is problematic as we often do not know if individuals would reach the same solutions as groups can if given the same cumulative exposure time or if more control individuals were tested (see Table 1).

One potential solution is to test individuals in the presence of others (either in dyads or groups, although dyadic testing faces similar separation and dominance issues as transmission chain studies). Similar to asocial controls, such naïve groups can provide baselines for groups that are seeded with animals trained on complex or efficient solutions, although this method does not work as a comparator for unseeded groups. Dyads and control groups may also benefit from social facilitation, as well as reduced vigilance, thereby ruling this out as a potential explanation for any enhanced performance in experimental groups. Alternatively, the role of social learning can be assessed even when control individuals necessarily receive shorter task exposure than groups, by controlling for test times or the number of trials participants engage in. That is, if groups are tested for longer periods or receive more trials, their data can be capped so they are equivalent to asocial control durations. A note of caution is required with this later method, as researchers will need to ensure they provide a reasonable amount of time for subjects to develop the skill or behavior of interest.
Studies directly comparing humans and chimpanzees have been highly informative for our understanding of what socio-cognitive mechanisms may distinguish human culture, including CCE, from one of our closest extant species. However, to our knowledge, all such studies have involved comparisons between human children (typically young children) and adult chimpanzees (Dean et al., 2012; Haun et al., 2014; McGuigan et al., 2017; Tennie et al., 2009; Vale et al., 2020; Vale, Flynn, et al., 2017; Van Leeuwen et al., 2014). There are good reasons for this; by middle childhood, children begin to outperform chimpanzees on a range of cognitive and socio-cognitive tasks, and thus using young children allows us not to conflate species differences in CCE behaviours with inter-species cognitive differences (Herrmann et al., 2007; Nielsen, 2009; Wobber et al., 2014). Relatedly, the tasks typically used in comparative research are relatively simple ones, such as puzzleboxes, for which adults would perform at ceiling (if the outcome variables are success based).

Nonetheless, despite these rationales, comparing young children to adult chimpanzees is also problematic for multiple reasons. First, developmental experiences, such as rearing conditions and maternal stress or deprivation, shape innovation and social learning propensities in a range of species, meaning that comparing samples at different life history stages is problematic for interpretation (Bard & Leavens, 2014; Mesoudi et al., 2016). Second, there are also within-species age differences in children and chimpanzees' cultural learning strategies. In non-human primates, surveys suggest that adults are more innovative than nonadults (Reader & Laland, 2001). Observational studies have shown that infancy and juvenility maybe be a sensitive period chimpanzees to socially acquire complex cultural behaviours, such as nut cracking (Inoue-Nakamura & Matsuzawa, 1997; Lonsdorf, 2017; Lonsdorf et al., 2004). Likewise, young children are particularly reliant on copying others, but become more innovative and less reliant on social information as they age through childhood (Carr et al., 2015; Rawlings, 2018). If this is the case, we would expect juvenile humans to behave differently from adult non-human primates based on age alone, invalidating these purported species differences. Third, from middle age, chimpanzees show an age-related decline in performance on cognitive tasks (Hopkins et al., 2021), meaning that drawing meaningful conclusions from comparing young children to older chimpanzees on cognitive based tasks is difficult. This is particularly relevant in studies that use captive chimpanzees, as this population is heavily biased towards older individuals.

Given these challenges, it is important for researchers to include age-period matched samples across species when possible, or at least to discuss the potential confounds. Comparing young nonhuman primates with human children will allow for direct comparisons of how development shapes the ontogeny of CCE in both species. Given breeding bans, particularly in chimpanzees, this will be difficult in some locations (i.e., research sites in the US), but in situations in which it is possible, such as in zoos and in the wild, we can use these data to help interpret non-age matched studies. This is also another context in which initiatives such as the ManyPrimates project, in which researchers collaborative pool individuals from multiple sites, would be particularly beneficial (Altschul et al., 2019). Finally, few studies have compared adult humans and chimpanzees. As noted, running such studies can be difficult if the unit of measurement is simply success on solving a given task, but tasks can be designed to capture cumulative improvement in other ways. Documenting time to success, propensity to engage in certain behaviours, and the maximum level of complexity or efficiency reached would allow for adult comparisons of both species.
An issue within comparative psychology has been a focus on too few model species, and the study of CCE is no exception, with its heavy focus on primates, particularly chimpanzees. Although recent years have seen a widening range of species studied there is still much to learn about CCE in the wider animal kingdom. Indeed, promising findings from species such as bighorn sheep (Jesmer et al., 2018), Savannah sparrows (Williams et al., 2013), humpback whales (Allen et al., 2013; Garland et al., 2011, 2017), zebra finches (Fehér et al., 2009), new Caledonian crows (Hunt & Gray, 2003), and bumblebees (Loukola et al., 2017) are indicative that diversifying species in CCE studies is important for a richer understanding of its extent beyond humans and its evolutionary origins. For example, bumblebees (Bombus spp.) developed more efficient ways of solving a ball pulling task after receiving social information (Loukola et al., 2017), while the songs of a population of Savannah sparrows (Passerculus sandwichiensis) were found to vary over three decades in a manner that increased fitness (Williams et al., 2013).

Studying other species is critical to gain a richer understanding of nonhuman behavior, however, taxa such as birds or bumblebees also offer some methodological advantages over nonhuman primates. Logistically, it may be easier to recruit larger sample sizes and/or implement generational study in smaller-bodied and/or shorter-lived species. It may also be easier to make use of natural behaviors with certain species (such as songs or migration routes), which improves ecological validity. In addition, the greater variety of natural behaviors available across the entire animal kingdom increases the number of questions that can be answered relative to studying only primates. For instance, we can ask is CCE more common in material culture or communication? Is it more common in long-lived organisms in which individuals have extensive opportunities to learn from one another, or in shorter-lived organisms in which there may be particular advantages to learning from others? As we continue to diversify our study species these questions can be addressed, and in turn, more questions will emerge.

Relatedly, nonhuman primate CCE research has largely focussed on tool use behaviours. As the increasing evidence of forms of CCE in domains outside of tool use - and with species who do not use tools frequently - shows, such a focus may be impeding theoretical progress. While we acknowledge that studying, for example, social conventions and communication can be difficult, experimentally expanding the way in which we study CCE is key to advancing the field. Expanding longitudinal data collection (Jesmer et al., 2018; Schofield et al., 2018) to measure if and how CCE emerges in these domains is an important step to diversify our understanding on the contexts in which CCE occurs.

Perhaps most critically, throughout comparative psychology there is a bias towards assuming that multi-step and/or highly social behaviors, like CCE, are complex and therefore require large brains, an assumption that has repeatedly been proven false (e.g., empathetic responding in rats; Bartal et al., 2011; interspecific cooperation in moray eels and grouper; Bshary et al., 2006; and visual individual recognition in wasps; Tibbetts, 2002). Knowing which species show aspects of CCE, and how it may vary across organisms, is required to both fully understand what behavioral mechanisms are essential and to better understand extant variability. For example, there are clear differences in what can be accomplished by
humans with language, and chimpanzees without, but language obscures other less obvious
factors; studying the variability in other species helps identify how these factors shape CCE
as well.

**Concluding remarks**

The cumulative nature of human culture is unquestionably distinct from all other species, yet
the question of nonhuman animal CCE remains an open one. The marked increase in CCE
research, both within the primates and more broadly across species, over the last two decades
has significantly improved our understanding of its cognitive underpinnings, its ontogenetic
and evolutionary origins, and the extent to which it exists outside of humans. Comparative
research has been crucial to identifying cross-species similarities and differences in cultural
behaviors and, ultimately, to identify key mechanisms underpinning the uniqueness of human
cumulative cultural evolution. As the field has developed, however, researchers have used
different criteria to determine CCE, which has led to studies that are not directly comparable.
The field of primatology has critically improved our knowledge of CCE through the
collection of data from multiple sites both in captivity and in wild settings and longitudinally.
Yet it also suffers from small samples, making generalizability difficult, and primates are
long-lived organisms for which generational studies are nearly impossible. We propose that a
combination of stricter definitions, more robust methods and a broader phylogenetic scope
will allow us to more fully understand the evolution and development of CCE and, therefore,
better understand what, if any, aspects of it are unique to humans and how our behavior fits
into that of the rest of the animal kingdom.

**Acknowledgements**

BR and CL were supported by a National Science Foundation grant (award number 1730678)
and a Templeton World Charity Foundation grant (award number TWCF0312) grant awarded
to CL. BR and CL were also supported by a PRC P2C Center grant to the Population
Research Center at The University of Texas at Austin. SFB and GLV were supported by
National Science Foundation grants SES 1658867 and SES 1919305.
References


Garland, E. C., Goldizen, A. W., Rekdahl, M. L., Constantine, R., Garrigue, C., Hauser, N.


https://doi.org/10.1098/rstb.2019.0419


https://doi.org/10.3233/DEV-2012-11088


https://doi.org/10.1002/ajp.22259


https://doi.org/10.1073/pnas.1323533111


https://doi.org/10.1007/s10071-007-0135-y


https://doi.org/10.1016/j.cognition.2017.07.015


https://doi.org/10.4000/primatologie.254


https://doi.org/10.1086/345825

Perry, S. E., Barrett, B. J., & Godoy, I. (2017). Older, sociable capuchins (Cebus capucinus) invent more social behaviors, but younger monkeys innovate more in other contexts. Proceedings of the National Academy of Sciences of the United States of America, 114(30), 7806–7813. https://doi.org/10.1073/pnas.1620739114


https://doi.org/10.1098/rspb.2009.0640


Rawlings, B. (2018). Establishing predictors of learning strategies; an investigation of the development of, and evolutionary foundations of, intrinsic and extrinsic factors influencing when we learn from others and from whom we learn. Durham University.


https://doi.org/10.4324/9780429352065-2


