PAYOFF-BIASED SOCIAL LEARNING UNDERLIES THE DIFFUSION OF NOVEL EXTRACTIVE FORAGING TRADITIONS IN A WILD PRIMATE

BRENDAN J BARRETT[1,2], RICHARD L MCELREATH[1,2,3], AND SUSAN E PERRY[4,5]

Abstract. The type and variety of learning strategies used by individuals to acquire behaviours in the wild are poorly understood, despite the taxonomic prevalence of behavioural traditions. Social learning strategies such as conformity can be broadly adaptive, but may also retard the spread of adaptive innovations. Strategies like payoff-biased social learning, in contrast, are effective at diffusing new behaviour but may perform poorly when adaptive behaviour is common. We present a field experiment in a wild primate, *Cebus capucinus*, that introduced a novel food item and documented the innovation and diffusion of successful extraction techniques. We develop a multilevel, Bayesian statistical analysis that allows us to quantify individual-level evidence for different social and individual learning strategies. We find that payoff-biased social learning and age-biased social learning are primarily responsible for the diffusion of the new techniques. We find no evidence of conformity; instead rare techniques receive slightly increased attention. We also find substantial and important variation in individual learning strategies that is patterned by age, with younger individuals being more influenced by both social information and their own individual experience. The aggregate cultural dynamics in turn depend upon the variation in learning strategies and the age structure of the wild population.

Keywords: payoff-bias, social learning, behavioural traditions, *Cebus*, cultural transmission, extractive foraging

[1]University of California, Davis; Animal Behavior Graduate Group
[2]University of California, Davis; Department of Anthropology
[3]Max Planck Institute for Evolutionary Anthropology
[4]University of California, Los Angeles; Department of Anthropology
[5]University of California, Los Angeles; Center for Behavior, Evolution, and Culture
E-mail address: bjbarrett@ucdavis.edu.

1. Introduction

The existence of culture or behavioural traditions [1] in non-human animals has been a topic of intrigue to evolutionary biologists and ethologists for centuries [2–4]. Recently, research interest in animal cultures has soared, partially driven by findings from long-term cross-site collaborations within primatology [5–7] and cetaceology [8;9] in the early 21st century. As the diversity of taxa in which social learning is studied grows, it appears that traditions might be more widespread and ecologically meaningful than was previously appreciated.

As evidence accumulates, the study of cultural mechanisms has shifted focus from asking "can animals learn socially?" to "how and under what conditions do animals learn socially?" The ecological drivers that favor social learning are theoretically well explored [10]. The mechanistic details and evolutionary and ecological consequences of social learning are less well understood. From an individual's perspective, it may be difficult to know whom or exactly what to copy. To cope with these difficulties, organisms use heuristics and strategies [10–12] to minimize the costs and increase the efficiency of social learning. Variation in learning strategy, whether between individuals or over the life course, can be equally important [13–15].

Different strategies have different advantages. Two families of social learning strategies that have received both theoretical and empirical attention are conformity and payoff-bias [10;16;17]. Conformist transmission, or positive frequency dependence, can be adaptive [10;18;19]. However, unless it is combined with other, flexible strategies, conformity may reduce the spread of adaptive innovations or cause population collapse [20]. In contrast to conformity, payoff-biased social learning is very effective at spreading novel adaptations. Payoff-biased social learning attends to behaviour that is associated with higher payoffs. However, it can be outperformed by conformity, once adaptive behaviour is common [21].

There is empirical evidence for both conformist and payoff-biased social learning in humans [17]. In other animals, conformity [22;23] has been studied more extensively than payoff-bias. To our knowledge, no non-human study has directly compared the explanatory power of conformity and payoff-biased social learning.

Here we report results from a field experiment with white-faced capuchin monkeys (*Cebus capucinus*) that is capable of distinguishing conformist and payoff-biased social learning. Capuchins provide an ideal study system for understanding social learning and traditions. They are tolerant of foraging in proximity with conspecifics [24], independently evolved many brain correlates associated with intelligence [25;26] and display the largest recorded repertoire of candidate behavioural traditions of any platyrrhine: social conventions [7], interspecific interactions [27] and extractive foraging techniques

[28–31]. Their reliance on social learning, frequency of innovation, and complexity of social interactions exemplifies what is predicted for long-lived animals with a slow life history strategy [32]. We investigated the innovation and transmission of extractive foraging techniques used to access the protected seeds of the *Sterculia apetala* fruit. This fruit occurs sporadically over the range of *C. capucinus*. Only some groups are experienced with it. By introducing the fruit to a naive group in controlled settings, we observed the rise and spread of new foraging traditions. We then inferred which social learning strategies best predict individual behaviour and how they influence the origins and maintenance of traditions.

The statistical analysis employs a multilevel dynamic learning model, of the form developed by [17], and inference is based upon samples from the full posterior distribution, using Hamiltonian Monte Carlo [33]. This model allows estimation of unique social and individual learning strategies for each individual in the sample. The analysis utilizes dynamic social network data, which were available during each field experimental session. It also permits examination of the relationship between any individual state (i.e. age, rank) and learning strategy. The multilevel approach makes it possible to apply these models to field data that lack precise balance. We provide all the code needed to replicate our results and to apply this same approach to any group time series of behaviour.

We document that the capuchins innovated a number of successful techniques. However, these techniques vary in their physical and time requirements. The statistical analysis suggests that payoff-biased social learning was responsible for this spread of the quickest, most successful techniques through the group. We find no evidence of conformity, but do find evidence of weak anti-conformity—rare techniques attracted more attention. We also find evidence of an age bias in social learning, in which older individuals were more likely to transmit their behaviour. Individuals varied in how they made use of social cues and individual experience, and age was a strong predictor. Our results comprise the first application of multilevel, dynamic social learning models to a study of wild primates and suggest that payoffs to behaviour can have important and different influences on social and individual learning. Methodologically, the approach we have developed is flexible, practical, and allows for a stronger connection between theoretical models of learning and the statistical models used to analyze data.

2. Study Design

2.1. **Study system.** This study was conducted between 2013 and 2015 on a group of habituated white-faced capuchin monkeys in and near Reserva Biológica Lomas Barbudal (RBLB) in northwest Costa Rica, during the months of December–February. See supplemental for additional information about the field site [34].

Capuchins heavily rely on extractive foraging to exploit difficult to access resources; this makes them an ideal comparative study system for understanding the evolution of extractive foraging in humans [25]. In neotropical dry forests, capuchins increase their reliance on extractive foraging during seasonal transitions when resources are limited. Capuchins receive more close, directed attention from conspecifics when they are foraging on large, structurally protected foods [35]. Many of the techniques required to access protected foods are candidate behavioural traditions [28].

Panamá fruits, *Sterculia apetala*, are a dietary staple of capuchins at RBLB; they comprise 8% of the diet of most groups in the early dry season [35]. The fruits are *empanada* shaped and the fatty, protein rich seeds within are protected by a hardened outer husk and stinging hairs [36]. Instead of waiting for fruits to dehisce, capuchins will open closed fruits and work around their the structural defenses, thus reducing competition with other organisms. Panamá fruits require multiple steps to effectively open, process, and consume, and panamá foraging generates the second highest level of close-range observation from conspecifics at RBLB [35]. Panamá processing techniques are also observed to vary between groups at RBLB and other field sites in the area [28], suggesting they are socially-learned traditions. Interestingly, some wild-caught capuchins are unable to open panamá fruits [36].

Panamá processing techniques differ in efficiency, measured by the average time it takes to open a fruit. Techniques also differ in efficacy, both in their probability of being successful and due to costs incurred by encountering stinging hairs. This contrasts with other extractive foraging traditions that show no difference in efficiency or efficacy [29]. Near RBLB, panamá trees grow in riparian and/or evergreen habitat in primary forests where they are dominant canopy species.

The focal group of this study, Flakes group (N=25), fissioned from the original study group in 2003. They migrated to a previously unoccupied patch of secondary agricultural and cattle ranching land characterized by riparian forest, pasture and neotropical oak woodland, where panamá trees are almost non-existent. Group scan data collected on foraging capuchins at RBLB from 2003–2011 show that Flakes was never observed foraging panamá, whereas other groups spent up to 1.21% of their annual foraging time eating panamá (Table S1). Two trees were found in the territory during phenological surveys, but are at the periphery, have small crowns, and are in areas of the habitat shared with other capuchin groups. When this study was designed, veterans of the field site had no recollection of observing Flakes foraging for panamá. Observations of 2 natal Flakes adult males (who would be expert panamá foragers in any other group) found outside of their territory migrating, suggest that they had little or no experience with panamá fruits. Efficiency at foraging for panamá markedly increased over the 3 years this experiment was conducted.

Several adults in the group (2 females, 3 males) grew up in different natal groups whose territories contained large numbers of panamá trees and whose groups exhibited higher rates of panamá foraging. For 2 migrant males from non-study groups, it is unknown if they previously learned to process panamá fruits, but this seems likely as evidenced by their skill. These individuals also differed in the primary processing techniques they used to process panamá that they presumably acquired in their natal group. By providing panamá fruits to both naïve/inexperienced juveniles and to knowledgeable adult demonstrators who differ in processing techniques, we collected fine-grained data showing how inexperienced capuchins learn a natural behaviour.

2.2. **Data Collection.** We collected panamá fruits from areas near RBLB to use in our experiment (see Supplemental). Observers were trained for at least one month on monkey identification in the field using facial recognition, size, and unique marks, and also memorized an ethogram of panamá foraging and social behaviour. Fruits were placed on a 25 cm diameter wooden platform which provided visual contrast of the fruits against the ground as fruits blended with the leaf litter, and so the capuchins had some sort of naturalistic spatial cue to associate with panamá fruits. Two fruits were placed on 1-2 platforms in each experimental bout. This permitted 1-4 capuchins to forage at a given time, and 2 fruits per platform was the maximum number on which a single human observer could reliably collect data.

We placed multiple fruits for two reasons. First, when individuals are naturally foraging for panamá, there choose from multiple available fruits in a tree. Second, we wanted to see whom they bias their attention toward when given a choice of multiple potential demonstrators. While many learning experiments have one potential demonstrator to learn from in a foraging bout or assume that everyone observes that demonstrator, we believe that allowing them to choose a potential learning model is more representative of how wild animals learn.

Fruits were placed on platforms under a poncho to obscure the monkey's view of us handling fruits. As ponchos were worn regularly when not experimenting, monkeys were unlikely to associate their presence with panamá platforms. When monkeys were not looking, we uncovered the fruits and walked to an observation area away from the platform so the monkeys could forage unimpeded. On digital audio recorders, we recorded if or when individuals saw, handled, processed, opened, ingested seeds from, and dropped each fruit. We verbally described how they were processing each fruit using an ethogram of techniques and which audience members observed them. More information on data collection and videos of processing can be found in the supplemental.

3. Results: Innovation and diffusion of techniques

We observed 7 types of predominant fruit processing techniques, which. varied in time required and the proportion of successful attempts (Table S2). Mean (median) duration ranged from 50 (29) to 330 (210) seconds. Proportion of successful attempts ranged from 0.38 to 0.89.

The technique frequencies changed over time, in the group and in most individuals (Figure 1). The most efficient technique, canine seam, went from non-existent in the group to the most common technique. It was introduced by an immigrant adult male (NP). Two knowledgeable adults, an adult female (ME) and the alpha male (QJ), switched to the canine seam technique. All others born after 2009 tried it at least once. However, canine seam never reached fixation in the population.

4. Results: Learning strategies

We analyzed these data using a hierarchical experience weighted attraction (EWA) model [37;38]. EWA models are a family of models that link individual learning strategies to records of group behavior [16;17;39].

4.1. **Social learning strategies.**

Our main focus is the contrast between two well-studied types of social learning, conformity and payoff-bias. However, we also investigate other, plausible strategies. We quickly describe the background of these strategies. We then describe how the modeling framework incorporates them.

Payoff-biased learning. Copying the behaviour with the highest observable payoff is useful social learning strategy [21;40]. In a foraging context, selectively copying rate-maximizing behaviour can increase the efficiency of diet and resource acquisition. Guppies choose food patches with higher return rates [41] while wild tufted capuchins bias attention toward the most efficient tool-users [42]. Cues of payoff may be noisy, however, and different individuals may require different techniques.

Model-biased learning. Where the content of behaviour cannot be evaluated, individuals might bias attention towards particular demonstrators or "models." Model biases [43] are efficient shortcuts to acquiring behaviour. Cues such as health, fertility, or rank may be correlated with adaptive behavior.

Prestige-biased learning is a popular example of model bias in humans [44]. While animals may lack the concept of prestige, they have analogues. Captive chimpanzees have been found to be more likely to copy dominant individuals [39;45], while vervets copied same-sex high-ranking individuals [46].

Copying the behaviour of one's parents is another option. If a parent can survive and successfully reproduce, its offspring's existence serves as a cue that her parents are successful [47]. *Luehea* processing techniques of capuchins at RBLB were predicted by both the technique their mother used

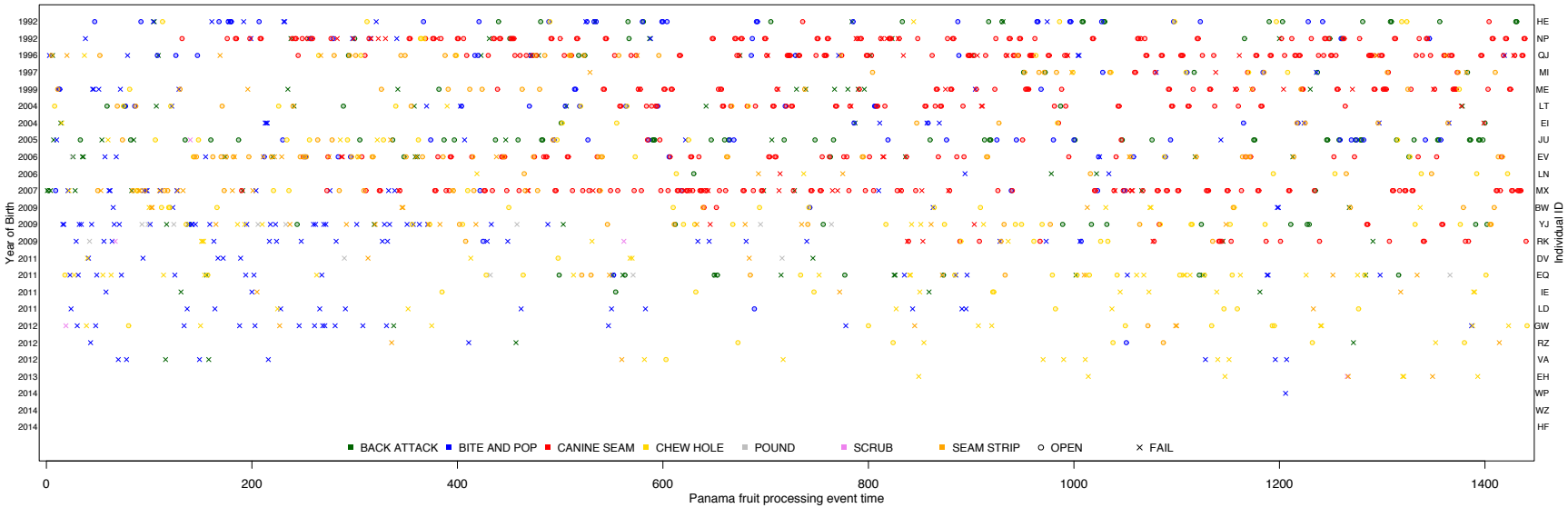


FIGURE 1. Techniques observed during experiment. Rows are unique individuals, from older (top) to younger (bottom). X-axis is event time. Each color represents a processing technique. ∘ indicate successes. × indicate failures. The most successful technique in red (canine seam) diffused to older members of the population. Younger individuals did not use canine seam, had more failed attempts, and fewer processing attempts.

and the technique they saw performed most often [29]. Kin-biased learning has been found in carnivores[48–50], but it is unclear whether this is due to cognition or is a consequence of family-unit social systems.

Copying similar individuals can be adaptive. Where individuals differ in strength, size, or cognitive ability, it might be beneficial for learners to copy those who are most similar to them. Great tits preferentially copied age-mates when learning to remove milk caps from bottles [51], while sex-biased learning has been found in several primate species [29;46].

Frequency-dependent learning. Frequency-dependent social learning occurs when frequency among demonstrators or frequency of demonstration influences adoption. It includes negative and positive frequency-dependence. Negative frequency dependence, or anti-conformity, is preferentially copying rare behaviour. It may be a form of neophilia. Positive frequency dependence, known also as conformity or majority-rule, is preferentially copying the most common behaviour. Conformity can lead to fixation of a cultural trait, maintain the stability of that trait [10;18]. Experiments in many captive [52–56] and some wild [22;23] animals have found evidence of conformist learning.

4.2. **Model design.** An EWA model comprises two parts: a set of expressions that specify how individuals accumulate experience and a second set of expressions that specify probability distributions over choices. Accumulated experience is represented by *attraction* scores, $A_{ij,t}$, unique to each behaviour i, individual j, and time t. A common formulation is to update $A_{ij,t}$ with an observed payoff $\pi_{ij,t}$:

$$A_{ij,t+1} = (1 - \phi_j)A_{ij,t} + \phi_j \pi_{ij,t} \tag{1}$$

The parameter ϕ_j controls the importance of recent payoffs in influencing attraction scores. This parameter is unique to individual j, and so can vary by age or any other feature.

To turn these attraction scores into behavioural choice, some function that defines a probability for each possible choice is needed. The conventional choice is a standard multinomial logistic, or *soft-max*, choice rule:

$$\Pr(i|A_{ijt}, \lambda) = \frac{\exp(\lambda A_{ij,t})}{\sum_k \exp(\lambda A_{kj,t})} = I_{ij} \tag{2}$$

The parameter λ controls how strongly differences in attraction influence choice. When λ is very large, the choice with the largest attraction score is nearly always selected. When $\lambda = 0$, choice is random with respect to attraction score. Individuals were assigned a payoff of zero, $\pi_{ij,t} = 0$, if they failed to open a panamá fruit. If they were successful, payoff was the inverse-log amount of time it took to open the fruit, $\pi_{ij,t} = \log(T_{\text{open}})^{-1}$. For the observed times T_{open}, this ensures that payoffs decline as T_{open} increases, but with the steepest declines early on.

Following previous work, social learning may influence choice directly and distinctly from individual learning. Let $S_{ij} = S(i|\mathbf{\Theta}_j)$ be the probability an individual j chooses behaviour i on the basis of a set of social cues and parameters $\mathbf{\Theta}_j$. Realized choice is given by:

$$\Pr(i|A_{ij,t}, \mathbf{\Theta}_j) = (1-\gamma_j)I_{ij,t} + \gamma_j S_{ij,t} \tag{3}$$

where γ_j is the weight, between 0 and 1, assigned to social cues. Under this formulation, social cues influence choice directly but attraction scores indirectly, only via the payoffs choice exposes an individual to.

We incorporate social cues into the term $S_{ij,t}$ by use of a multinomial probability expression with a log-linear component $B_{ij,t}$ that is an additive combination of cue frequencies. Specifically, the probability of each option i, as a function only of social cues, is:

$$S_{ij,t} = \frac{N^f_{ij,t} \exp B_{ij,t}}{\sum_m N^f_{mj,t} \exp B_{mj,t}} \tag{4}$$

This is easiest to understand in pieces. The $N_{ij,t}$ variables are the observed frequencies of each technique i at time t by individual j. The exponent f controls the amount and type of frequency dependence. When $f = 1$, social learning is unbiased by frequency and techniques influence choice in proportion to their occurrence. When $f > 1$, social learning is conformist. Other social cues, like payoff, are incorporated via the $B_{ij,t}$ term:

$$B_{ijt} = \sum_k \beta_k \kappa_{k,ijt} \tag{5}$$

This is the sum of the products of the influence parameters β_k and the cue values $\kappa_{k,ijt}$. We consider five cues.

(1) Payoff. $\kappa = \log(t_{\text{open}})^{-1}$ or, for failure, $\kappa = 0$.
(2) Demonstrator rank. $\kappa = 1$ for alpha rank, 0 otherwise.
(3) Matrilineal kinship. $\kappa = 1$ for matrilineal kin, 0 otherwise.
(4) Age similarity. κ is defined as the inverse absolute age difference: $(1 + |\text{age}_{\text{demonstrator}} - \text{age}_{\text{observer}}|)^{-1}$.
(5) Age bias. $\kappa = \text{age}_{\text{demonstrator}}$.

The final components needed are a way to make the individual-level parameters depend upon individual state and a way to define the window of attention for social cues at each time t. The parameters γ_j and ϕ_j control an individual j's use of social cues and rate of attraction updating, respectively. We model these parameters as logistic transforms of a linear combination of predictors. For example, the rate of updating ϕ_j for an individual j is defined as:

$$\text{logit}(\phi_j) = \alpha_j + \mu_\phi \times \text{age}_j \tag{6}$$

where α_j is a varying intercept per individual and μ_ϕ is the average influence of age on the log-odds of the updating rate. Social information available at each time step in the model was a moving window of the previous 14 days

TABLE 1. Posterior medians and standard deviations from the global model. Estimates of $\sigma_{individual}$ are the standard deviations of varying effects for that parameter across individuals. Posteriors visualized in Figures S1 and S2.

Parameter	λ	ϕ	γ	f_c	β_{pay}	β_{kin}	β_{rank}	β_{coho}	β_{age}	μ_ϕ	μ_γ
Posterior Med	20.97	0.15	0.14	0.38	1.02	0.19	-0.11	0.48	0.69	-0.11	-0.10
Posterior SD	1.11	0.03	0.03	0.28	0.84	0.93	0.91	0.93	0.92	0.03	0.05
$\sigma_{individual}$		0.66	0.69	1.29	0.28	0.25	0.26	0.26	0.25		

of observed foraging bouts. This allows new social information to be used, while old information is discarded. This choice is arbitrary. So to test the sensitivity of this time window, we also analyzed the data with moving time windows of 7, 21, and 28 days. The results we present are robust to these variations in the size of the window used to calculate social cues (Table S4).

To fit the model, we defined a global model incorporating all cues, using both parameter regularization and model comparison with sub-models to account for overfitting. Overall nine models were fit representing nine learning strategies. Models were fit using the Hamiltonian Monte Carlo engine Stan, version 2.14.1 [33], in R version 3.3.2 [57]. We compared models using WAIC [58]. To check our approach, we simulated the hypothesized data generating process and recovered data-generating values from our simulated data. We chose conservative, weakly informative priors for our estimated parameters. This made our models sceptical of large effects and helped ensure convergence. Data and code for models, simulations, and graphs are available at `https://github.com/bjbarrett/panama1`.

4.3. **Results of EWA models.** There was overwhelming support for some mix of individual and social learning over individual learning alone (see supplemental). The highest ranked model was the global model containing all strategies and age effects on learning parameters, which received 95% of the total model weight. We focus on this model, as it is both highest ranking and its parameter values agree with the weights assigned in the overall model set.

Individual marginal posterior distributions of each parameter are displayed in Table 1. Note that while the marginal posterior distribution of each parameter provides some information, the model is too complicated to interpret these parameters directly. For example, the weight of social information γ applies only at each choice. It is not a partitioning of the importance of social versus individual information in the diffusion of traditions. The overall influence of social information cannot be partitioned, like in an analysis of variance. Therefore we supplement these marginals with visualizations of implied individual behaviour, using posterior predictive distributions (Figure S3) .

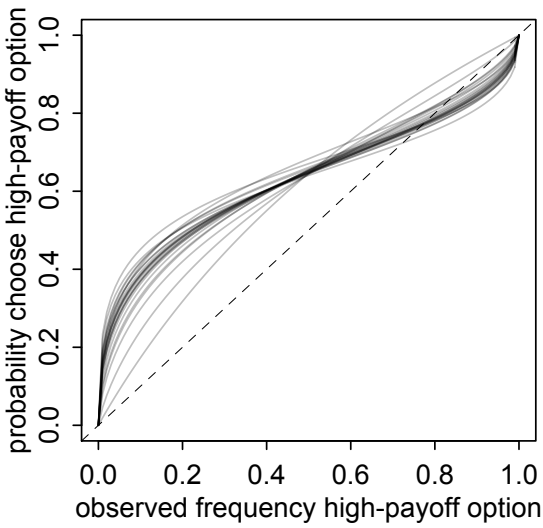
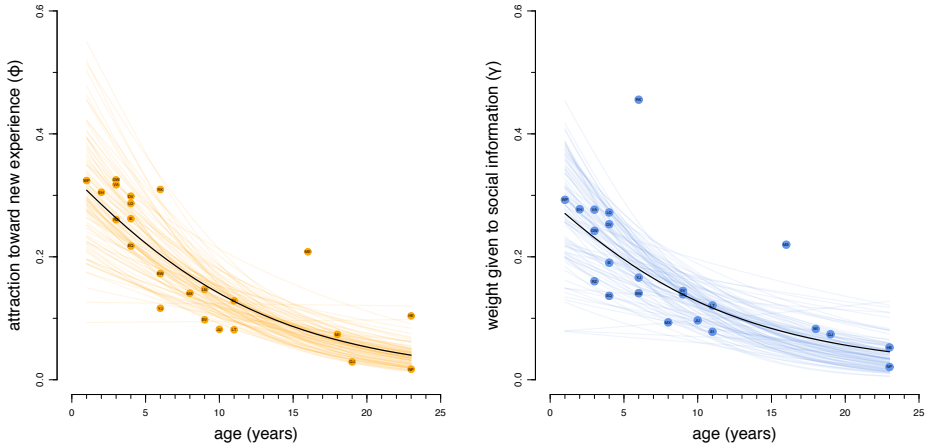


FIGURE 2. Posterior predictions of probabilities of choosing a socially observed option with payoff $\log(t_{\text{open}})^{-1} = 0.5$, relative to an observed option that was not successfully opened.

Influence of conformity and payoff-bias (f and $\beta_{\textbf{pay}}$): The raw marginal conformist exponent is below 1 on average, indicating mild anti-conformity. The marginal payoff-bias coefficient is strongly positive, indicating attraction to high-payoff actions. Figure 2 visualizes the individual social learning function S_{ijt} (Expression 4) implied when only conformity and payoff-bias are present. The horizontal axis is the observed frequency of a higher payoff option among demonstrators. The vertical axis is the probability an individual chooses the higher payoff option. Each curve in the figure represents the posterior mean of for an individual. The diagonal dashed line represents unbiased social learning. All individuals are strongly biased by payoff, resulting in a preference for the high-payoff option over most of the range of the horizontal axis. But most individuals also display weak anti-conformity, resulting in a preference for the rarer, low-payoff option in the upper right corner.

Weight of past experience (ϕ): On average, capuchins more heavily favor previous experiences over new ones ($\phi = 0.15; [0.11, 0.20]$ 89% credible interval) , Table 1). However, there is considerable individual variation in attraction to new experience ($\sigma_{individual} = 0.66$) ranging from 0.08 to 0.36, which was negatively predicted by age ($\mu_{age} = -0.11$; 89% CI $[-0.16, -0.06]$; Figure 3a). This suggests that older individuals are more canalized than younger individuals.



(A) (B)

FIGURE 3. Relationships between age and (A) attraction to new experience (ϕ) and (B) influence of social information (γ). Black line represents posterior mean. Solid points are posterior means of individual varying effects. Lighter lines are 100 posterior samples.

Weight of social information (γ): γ estimates for individuals varied considerably, ranging from 0.07-0.39 ($\sigma_{individual} = 0.66$). γ was also negatively related to age ($\mu_{age} = -0.10$; 89% CI $[-0.18, -0.03]$; Fig. 3b). This suggests that younger individuals rely more on social cues.

Age bias (β_{age}): Age bias contributed notably to social learning in our global model($\beta_{age} = 0.69$; 89% CI $[-0.79, 2.14]$; Table 1), suggesting that capuchins were more likely to copy older demonstrators.

Age similarity, kin, and rank biases. None of age similarity, matrilineal kin, or rank biases presented a strong or consistent effect (coho, kin, and rank in Table 1). While these strategies may have influenced some individuals and decisions, there is no evidence of general importance for these cues.

5. DISCUSSION

We set out to examine the roles of conformist and payoff-biased social learning among wild capuchin monkeys during the diffusion of a novel food processing traditions. We find no evidence of conformity, defined as positive frequency dependence. We do however find strong evidence of payoff-biased learning.

Little work has examined whether animals use payoff-biased social learning. We do not know how common such strategies are in nature. It is common to experimentally examine payoff-equivalent options, shedding no light

on payoff-bias. The common exclusion approach to identifying animal culture accidentally excludes payoff-bias, by diagnosing ecologically correlated behavioral differences as non-cultural [5]. This may result in overlooking adaptive socially-learned behaviour. If payoff-bias is common, this makes the problem of identifying animal traditions more subtle.

We also found evidence that other social cues, such as age, influence social learning. Age also modulated underlying learning parameters. In combination, these influences are sufficient to describe the diffusion and retention of successful foraging techniques within the group. In the remainder of the discussion, we elaborate on the findings and summarize some of the advantages and disadvantages of our approach.

5.1. **Wild capuchins acquire extractive foraging techniques quickly via social learning.** This study shows that one group of wild capuchin monkeys socially learn extractive foraging techniques from conspecifics and supports claims that food processing techniques are socially learned traditions. It has been challenging to find experimental evidence for social learning of object manipulation tasks in capuchins [25;59]. Better evidence for social learning might be found across a broader range of taxa, if more ecologically valid behaviours are studied in the wild. This study also demonstrates that animals may be able to acquire new, efficient behaviour in a matter of a days or weeks. This rapid pace of social transmission suggests that learning can act to rapidly facilitate behavioural responses to environmental change [12].

We found that payoff-biased learning and negative frequency dependence guided diffusion of panamá processing techniques in this group (Table 1). These strategies are consistent with the observation that the rarest and most efficient panamá processing technique, canine seam, eventually became the most common. This was the case for most, but not all, naive and knowledgeable adults and subadults born after 2009 (Figure 1). Juveniles born before 2009 did not use the canine seam technique (Figure 1), likely because their mouths were not sufficiently large and strong.

Payoff-bias had the largest effect on the probability of choosing a behaviour, while negative frequency dependence may have prevented it from ever reaching fixation. Experimental evidence of wild animals using payoff-biased learning has not been previously reported. Our finding of negative frequency-dependent learning suggests that capuchins bias their attention towards rare or novel behaviours—a type of neophilia.

While all adult individuals tried the canine seam technique, they typically settled on the technique(s) that was most successful for them. Individuals who settled on the canine seam technique also sporadically tried other behaviours (Figure 1). This result is consistent with the possibility that social learning is guiding exploration but personal experience strongly influences adoption.

While we found the strongest support for payoff-biased learning, our modelling suggests that animals use multiple social learning strategies simultaneously or that social biases and content biases might be equifinal. Age-biased learning also had support in the global model (Table 1). This might be due to older individuals' increased likelihood of being efficient panamá processors compared to juveniles, but the preferences for some individuals (JU and LN) to copy the techniques of the adults they commonly associates with who did not use canine seam (HE and MI accordingly) suggests otherwise.

Nevertheless, observational studies are always limited in their ability to distinguish some mechanisms from others. We believe that long-term field studies, field experiments, and controlled captive experiments all have important and complementary roles to play.

5.2. **Age predicts individual variation in social and individual learning.** Individual variation in social learning may have meaningful evolutionary and social implications, yet remains poorly studied [13]. We found that younger individuals more heavily relied on social learning than older individuals (Figure 3b) and that older individuals were less likely to observe conspecifics.

We also observed that older individuals were less likely to update information and had a greater attraction to previous experiences (Fig. 3a). This might be due to older individuals being less exploratory than younger individuals. But an alternative and likely explanation is that older individuals were more capable of discerning between the efficiency of different techniques. Older individuals processed successfully more frequently and had more opportunities to evaluate higher quality personal information (Figure 1).

This age structure in proclivity to learn socially suggests flexible learning strategies that change over development. Theory predicting and explaining such flexible variation waits to be constructed.

5.3. **Statistical approach.** Our analytical approach was designed around three important principles. First, it allows us to evaluate the possible influence of several different, theoretically plausible, social learning biases. Second, the framework combines social learning biases with a dynamic reinforcement model in which individuals remember and are influenced by past experience with different techniques. Third, the approach is fully hierarchical, with each individual possessing its own parameters for relative use of each learning strategy. This allows us to evaluate heterogeneity and its contribution to population dynamics.

Our approach is distinct from looking for evidence of population-level learning dynamics consistent with the hypothesized learning strategy [23; 60]. In our approach, population level patterns are consequences of inferred strategies. Such patterns are not themselves used to make inferences about learning.

Our approach is most similar to network-based diffusion analysis (NBDA) [51;61;62]. In principle, our framework and NBDA can be analogized, despite differences in the details of modeled strategies, because both are multinomial time series modeling frameworks that can be treated as both survival (time-to-event) or event history analyses. There are some notable differences in practice. Our approach differs from typical NBDA in that it: 1) uses a full dynamic time series for available social information rather than a static social network and 2) emphasizes modeling the entire behavioural sequence. There is no reason in principle that ordinary NBDA models could not make similar use of these data.

It is important to note that successfully fitting these dynamic, hierarchical models benefits from recent advances in Monte Carlo algorithms. We used an implementation of Hamiltonian Monte Carlo (NUTS2) provided by Stan [33]. Our global model contains 231 parameters and would prove very challenging for older algorithms like Gibbs Sampling. Hamiltonian Monte Carlo not only excels at high-dimension models, even with thousands of parameters, but it also provides greatly improved mixing diagnostics that allow us to have greater confidence in the correctness of the results, regardless of model complexity.

5.4. **Implications for the origins and maintenance of traditions.** This model suggests that payoff-biased learning can cause the spread of a tradition. However, social learning may increase within-group homogeneity, while individual learning may act to decrease it [50]. Our findings are consistent with this idea. Limited transfer of individuals in xenophobic species like *Cebus* is exceptionally important in maintaining group specific traditions for behaviours that differ in payoff. However, this likely acts concordant with transmission biases. Variation might also be maintained due to biases for copying particular subsets of individuals (e.g. a particular age-class or kin group) in a stable social system. Migration of new individuals with more efficient behaviours could seed a new tradition in the group, the diffusion of which may be due to payoff-biased learning.

5.5. **Future Directions.** We have noted that equifinality might exist between learning strategies. On average, older individuals were better at opening panamá fruit. Perhaps individuals are biasing learning toward older individuals and acquiring the efficient techniques indirectly instead of turning attention toward the content of the behaviour. While we think this is likely not the case based on the evidence considered in this study, it is a possibility in all learning studies. In many cases, where we are interested in predicting the population dynamics of learning in a given context, the exact social learning strategy might not matter if it has the same dynamics and leads to the same frequency in a population. Many learning strategies are likely equifinal under the right social conditions. However, the exact nature of the cognitive mechanisms of the learning strategies organisms employ, and the social factors which indirectly structure learning, become important when

we wish to use social learning in applied contexts. Further theoretical and empirical explorations of social learning need to address that learning is a two stage process: one of assortment and one of information use.

An important aspect of learning that we have neglected is the endogeneity of social information. Our statistical models evaluated how individuals use information they observed. However, before individuals acquire social information, they make the decision to observe others. Future analyses will evaluate who individuals choose to bias attention toward when in the proximity of potential demonstrators to see how positive assortment might structure opportunities for social learning and affect the establishment and maintenance of traditions.

Most models of social learning in the evolutionary anthropology and animal behaviour literature assume a randomly assorted population. However, non-random assortment occurs before information is acquired in a population, and it can drastically affect social learning and cultural dynamics. Sometimes this assortment may be an adaptive heuristic, such as deciding to bias attention. Other times it may be an indirect consequence of social behaviour, such as avoidance of a potentially dangerous demonstrator [15]. Asymmetrical age structure in a population may also make the behavioural variants in the population non-random when learning abilities are constrained by skill and developing cognition [63]. Social networks can also change drastically over development, opening up avenues for new possible learning strategies. Some learning strategies might be difficult to tease apart in small, non-diverse social systems. If a juvenile engages in kin-biased learning [64], but only interacts with their kin group, how are we to discern kin-biased learning from linear imitation or conformity, and under what conditions does this distinction matter?

6. Authors' Contributions

BB designed study, collected data, carried out analysis, and drafted manuscript; RM participated in analysis and helped draft manuscript; SP established field site, collected data, and helped draft manuscript.

7. Acknowledgments

Thanks to: M. Grote, M. Crofoot,and the UC Davis CE/HBE lab for useful feedback, A. Cobden, B. Davis, E. Seabright, M. White, M. Ziegler, C. Angyal, for assistance with data collection, W. Lammers, J.C Ordoñez Jiménez, I. Godoy, and K. Kajokaite, for assistance with field logistics, and MINAET and SINAC, Hacienda Pelon de la Bajura, Hacienda Brin D'Amor for permission to work on their land. This research was performed in compliance with Costa Rican law, and followed approved UCLA (ARC #1996-122 and 2005-084 plus renewals) and UC Davis (IAUCUC permit #17297) animal care and use protocols.

8. Research Funds

Funding to BB was provided by the American Society of Primatologists and the ARCS Foundation. This material is based upon work supported by a NSF GRF (Grant No. 1650042) awarded to BB. SP recieved funding from the Max Planck Institute for Evolutionary Anthropology, the National Science Foundation (grants No. SBR-0613226 and BCS-0848360), L.S.B Leakey Foundation, National Geographic Society, and UCLA COR. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

[1] Fragaszy DM, Perry S, 2003 *The Biology of Traditions: Models and Evidence.* Cambridge University Press

[2] Kawamura S, 1959 The process of sub-culture propagation among Japanese macaques. *Primates* **2**, 43–60

[3] Marler P, Tamura M, 1964 Culturally transmitted patterns of vocal behavior in sparrows. *Sci. (New York, N.Y.)* **146**, 1483–1486

[4] Bonner JT, 1983 *The Evolution of Culture in Animals.* Princeton, N.J.: Princeton University Press, reprint edition edn.

[5] Whiten A, *et al.*, 1999 Cultures in chimpanzees. *Nat.* **399**, 682–685. doi:10.1038/21415

[6] van Schaik CP, *et al.*, 2003 Orangutan Cultures and the Evolution of Material Culture. *Sci.* **299**. doi:10.1126/science.1078004

[7] Perry S, *et al.*, 2003 Social conventions in wild white-faced capuchin monkeys: Evidence for traditions in a neotropical primate. *Curr. Anthropol.* **44**, 241–268. doi:10.1086/345825

[8] Whitehead H, 1998 Cultural Selection and Genetic Diversity in Matrilineal Whales. *Sci.* **282**, 1708–1711. doi:10.1126/science.282.5394.1708

[9] Rendell LE, Whitehead H, 2001 Culture in whales and dolphins. *Behav. Brain Sci.* **24**, 309–382

[10] Boyd R, Richerson PJ, 1985 *Culture and the Evolutionary Process.* University Of Chicago Press

[11] Laland KN, 2004 Social learning strategies. *Animal Learn. & Behav.* **32**, 4–14. doi:10.3758/BF03196002

[12] Rendell L, *et al.*, 2010 Why copy others? Insights from the social learning strategies tournament. *Sci.* **328**

[13] Mesoudi A, *et al.*, 2016 The Evolution of Individual and Cultural Variation in Social Learning. *Trends Ecol. & Evol.* **31**, 215–225. doi:10.1016/j.tree.2015.12.012

[14] Farine DR, *et al.*, 2015 Early-Life Stress Triggers Juvenile Zebra Finches to Switch Social Learning Strategies. *Curr. Biol.* **25**, 2184–2188. doi:10.1016/j.cub.2015.06.071

[15] Russon AE, 2002 Comparative developmental perspectives on culture: the great apes. In H Keller, YH Poortinga, A Schlmerich, eds., *Between Culture and Biology: Perspectives on Ontogenetic Development.* Cambridge University Press

[16] McElreath R, *et al.*, 2005 Applying evolutionary models to the laboratory study of social learning. *Evol. Hum. Behav.* **26**, 483–508. doi:10.1016/j.evolhumbehav.2005.04.003

[17] McElreath R, *et al.*, 2008 Beyond existence and aiming outside the laboratory: estimating frequency-dependent and pay-off-biased social learning strategies. *Philos. Transactions Royal Soc. B: Biol. Sci.* **363**, 3515–3528

[18] Henrich J, Boyd R, 1998 The Evolution of Conformist Transmission and the Emergence of Between-Group Differences. *Evol. Hum. Behav.* **19**, 215–241. doi:10.1016/S1090-5138(98)00018-X

[19] Nakahashi W, *et al.*, 2012 Adaptive Social Learning Strategies in Temporally and Spatially Varying Environments. *Hum. Nat.* **23**. doi:10.1007/s12110-012-9151-y

[20] Whitehead H, Richerson PJ, 2009 The evolution of conformist social learning can cause population collapse in realistically variable environments. *Evol. Hum. Behav.* **30**, 261 – 273

[21] Baldini R, 2013 Two success-biased social learning strategies. *Theor. population biology* **86**, 43–49. doi:10.1016/j.tpb.2013.03.005

[22] van de Waal E, *et al.*, 2013 Potent Social Learning and Conformity Shape a Wild Primate's Foraging Decisions. *Sci.* **340**, 483–485. doi:10.1126/science.1232769

[23] Aplin LM, *et al.*, 2015 Experimentally induced innovations lead to persistent culture via conformity in wild birds. *Nat.* **518**, 538–541. doi:10.1038/nature13998

[24] Oppenheimer JR, 1968 *Behavior and ecology of the white-faced monkey: Cebus capucinus, on Barro Colorado island, C.Z.* University of Illinois.

[25] Fragaszy DM, *et al.*, 2004 *The complete capuchin: the biology of the genus Cebus.* Cambridge; New York: Cambridge University Press

[26] Jerison H, 1973 *Evolution of the Brain and Intelligence.* Academic Press

[27] Rose LM, *et al.*, 2003 Interspecific Interactions Between Cebus capucinus and Other Species: Data from Three Costa Rican Sites. *Int. J. Primatol.* **24**. doi:10.1023/A:1024624721363

[28] Panger MA, *et al.*, 2002 Cross-site differences in foraging behavior of white-faced capuchins (Cebus capucinus). *Am. J. Phys. Anthropol.* **119**

[29] Perry S, 2009 Conformism in the food processing techniques of white-faced capuchin monkeys (Cebus capucinus). *Animal Cogn.* **12**. doi:10.1007/s10071-009-0230-3

[30] O'Malley RC, Fedigan LM, 2005 Evaluating social influences on food-processing behavior in white-faced capuchins (Cebus capucinus). *Am. J. Phys. Anthropol.* **127**, 481–491. doi:10.1002/ajpa.20095

[31] Perry S, 2011 Social traditions and social learning in capuchin monkeys (Cebus). *Philos. Transactions Royal Soc. B: Biol. Sci.* **366**, 988–996. doi:10.1098/rstb.2010.0317

[32] Sol D, *et al.*, 2016 The life-history basis of behavioural innovations. *Phil. Trans. R. Soc. B* **371**, 20150187

[33] Stan Development Team, 2016. RStan: the R interface to Stan. http://mc-stan.org/rstan.html

[34] Perry S, *et al.*, 2012 The Lomas Barbudal Monkey Project: Two Decades of Research on Cebus capucinus. In PM Kappeler, DP Watts, eds., *Long-Term*

Field Studies of Primates, 141–163. Springer Berlin Heidelberg

[35] Perry S, Ordoñez Jimenez JC, 2006 The efiects of food size, rarity, and processing complexity on white-faced capuchins' visual attention to foraging conspecifics. In G Hohmann, MM Robbins, C Boesch, eds., *Feeding Ecology in Apes and Other Primates*, vol. 48, 203–234. Cambridge University Press

[36] Janzen DH, 1972 Escape in Space by Sterculia Apetala Seeds from the Bug Dysdercus Fasciatus in a Costa Rican Deciduous Forest. *Ecol.* **53**. doi:10.2307/1934092

[37] Camerer C, Hua Ho T, 1999 Experience-weighted Attraction Learning in Normal Form Games. *Econom.* **67**, 827–874. doi:10.1111/1468-0262.00054

[38] Hoppitt W, Laland KN, 2013 *Social Learning: An Introduction to Mechanisms, Methods, and Models*. Princeton: Princeton University Press

[39] Kendal R, *et al.*, 2015 Chimpanzees copy dominant and knowledgeable individuals: implications for cultural diversity. *Evol. Hum. Behav.* **36**, 65–72

[40] Baldini R, 2012 Success-biased social learning: cultural and evolutionary dynamics. *Theor. population biology* **82**, 222–228. doi:10.1016/j.tpb.2012.06.005

[41] Pike TW, *et al.*, 2010 Learning by proportional observation in a species of fish. *Behav. Ecol.* **21**. doi:10.1093/beheco/arq025

[42] Ottoni EB, *et al.*, 2005 Watching the best nutcrackers: what capuchin monkeys (Cebus apella) know about others' tool-using skills. *Animal Cogn.* **8**, 215–219. doi:10.1007/s10071-004-0245-8

[43] Henrich J, McElreath R, 2003 The evolution of cultural evolution. *Evol. Anthropol. Issues, News, Rev.* **12**, 123–135. doi:10.1002/evan.10110

[44] Henrich J, Gil-White FJ, 2001 The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evol. Hum. Behav.* **22**, 165–196. doi:10.1016/S1090-5138(00)00071-4

[45] Horner V, *et al.*, 2010 Prestige Affects Cultural Learning in Chimpanzees. *PLOS ONE* **5**, e10625. doi:10.1371/journal.pone.0010625

[46] van de Waal E, *et al.*, 2010 Selective attention to philopatric models causes directed social learning in wild vervet monkeys. *Proc. Royal Soc. B: Biol. Sci.* **277**, 2105–2111

[47] McElreath R, Strimling P, 2008 When Natural Selection Favors Imitation of Parents. *Curr. Anthropol.* **49**, 307–316. doi:10.1086/524364

[48] Mazur R, Seher V, 2008 Socially learned foraging behaviour in wild black bears, Ursus americanus. *Animal Behav.* **75**, 1503–1508. doi:10.1016/j.anbehav.2007.10.027

[49] Müller CA, Cant MA, 2010 Imitation and Traditions in Wild Banded Mongooses. *Curr. Biol.* **20**. doi:10.1016/j.cub.2010.04.037

[50] Thornton A, Clutton-Brock T, 2011 Social learning and the development of individual and group behaviour in mammal societies. *Philos. Transactions Royal Soc. B: Biol. Sci.* **366**, 978–987. doi:10.1098/rstb.2010.0312

[51] Aplin LM, *et al.*, 2013 Milk bottles revisited: social learning and individual variation in the blue tit, Cyanistes caeruleus. *Animal Behav.* **85**, 1225–1232. doi:10.1016/j.anbehav.2013.03.009

[52] Chou LS, Richerson PJ, 1992 Multiple models in social transmission of food selection by Norway rats, Rattus norvegicus. *Animal Behav.* **44, Part 2**. doi:10.1016/0003-3472(92)90039-C

[53] Day RL, *et al.*, 2001 Interactions between shoal size and conformity in guppy social foraging. *Animal Behav.* **62**. doi:10.1006/anbe.2001.1820

[54] Whiten A, *et al.*, 2005 Conformity to cultural norms of tool use in chimpanzees. *Nat.* **437**, 737–740. doi:10.1038/nature04047

[55] Dindo M, *et al.*, 2009 In-Group Conformity Sustains Different Foraging Traditions in Capuchin Monkeys (Cebus apella). *PLoS ONE* **4**, e7858. doi:10.1371/journal.pone.0007858

[56] Pike TW, Laland KN, 2010 Conformist learning in nine-spined sticklebacks' foraging decisions. *Biol. Lett.* **6**. doi:10.1098/rsbl.2009.1014

[57] R Core Team, 2016. R: A Language and Environment for Statistical Computing

[58] Watanabe S, 2010 Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory. *J. Mach. Learn. Res.* **11**

[59] Fragaszy D, Visalberghi E, 2004 Socially biased learning in monkeys. *Animal Learn. & Behav.* **32**, 24–35. doi:10.3758/BF03196004

[60] Henrich J, 2001 Cultural transmission and the diffusion of innovations: Adoption dynamics indicate that biased cultural transmission is the predominate force in behavioral change. *Am. Anthropol.* **103**, 992–1013

[61] Hoppitt W, Laland KN, 2011 Detecting social learning using networks: a users guide. *Am. J. Primatol.* **73**. doi:10.1002/ajp.20920

[62] Allen J, *et al.*, 2013 Network-Based Diffusion Analysis Reveals Cultural Transmission of Lobtail Feeding in Humpback Whales. *Sci.* **340**, 485–488. doi:10.1126/science.1231976

[63] Russon AE, 2003 Developmental perspectives on great ape traditions. In DM Fragaszy, S Perry, eds., *The biology of traditions: models and evidence. Cambridge University Press, Cambridge*

[64] Kline MA, *et al.*, 2013 Teaching and the Life History of Cultural Transmission in Fijian Villages. *Hum. Nat.* **24**. doi:10.1007/s12110-013-9180-1

[65] Frankie GW, *et al.*, 1988 Nest Site and Habitat Preferences of Centris Bees in the Costa Rican Dry Forest. *Biotropica* **20**. doi:10.2307/2388320

[66] Perry S, *et al.*, 2003 Traditions in wild white-faced capuchin monkeys. In DM Fragaszy, SE Perry, eds., *The biology of traditions: Models and evidence*