# 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 presence of behavioural traditions in diverse taxa. Social learning strategies such as conformity can be broadly adaptive, but may also retard the spread of adaptive innovations. Strategies like payoff-biased 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 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

## 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,
5 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.

10 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 favour social learning are theoretically well explored [10]. The mechanistic details and evolutionary and ecological consequences of social learning are less well
15 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, may also be important [13–15].

20 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 especially in spatially heterogeneous environments [10;18;19]. However, unless it is combined with other, flexible
25 strategies, conformity may prohibit more adaptive behaviours from spreading [18;20] or cause population collapse [21]. 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 and presumably increase fitness. However, it can be outper-
30 formed by conformity, once adaptive behaviour is common [22].

There is empirical evidence for both conformist and payoff-biased social learning in humans [17]. In other animals, conformity [23;24] 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-
35 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 are an excellent study system for understanding social learning and traditions. They are tolerant of foraging
40 in proximity with conspecifics [25], independently evolved many brain correlates associated with intelligence [26;27] and display the largest recorded repertoire of candidate behavioural traditions of any platyrrhine: social conventions [7], interspecific interactions [28] and extractive foraging techniques

[29–32]. Their reliance on social learning, frequency of innovation, and com-plexity of social interactions exemplifies what is predicted for long-lived an-imals with a slow life history strategy [33]. We investigated the innovation and transmission of extractive foraging techniques used to access the pro-tected 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 (aka hierarchical or varying effects) dynamic learning model, of the form developed by [17], and inference is based upon samples from the full posterior distribution, using Hamilton-ian Monte Carlo [34]. This model allows estimation of unique social and individual learning strategies for each individual in the sample. The anal-ysis 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 and repeatedly sample individuals. We provide all 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 tech-niques. However, these techniques vary in their physical and time require-ments. 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 be copied. 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 learn-ing. 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 analyse 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 electronic supplemental material and [35;36] for additional information about field site.

Capuchins heavily rely on extractive foraging to exploit difficult to access resources; this makes them an excellent comparative study system for understanding the evolution of extractive foraging in humans [26]. 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 [37]. Many of the techniques required to access protected foods are candidate behavioural traditions [29].

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 [37]. The fruits are *empanada* shaped and the fatty, protein rich seeds within are protected by a hardened outer husk and stinging hairs [38]. Instead of waiting for fruits to dehisce, capuchins will open closed fruits and work around their the structural defences, 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 [37]. Panamá processing techniques are also observed to vary between groups at RBLB and other field sites in the area [29], suggesting they are socially-learned traditions. Wild capuchins without prior exposure to panamá fruits can't initially open them [38], suggesting that personal experience and/or social influence are important.

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 [30].

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 as they typically grow in evergreen, primary forests. 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 two natal Flakes adult males (old enough to 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.

Five adults in the group (two females, three 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 two migrant

males from non-study groups, it is unknown if they previously learned to process panamá, but this seems likely as evidenced by their skill. These individuals acted as models for different behaviours, as they differed in the
135  primary panamá processing techniques they presumably acquired in their natal groups. 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.

140  2.2. **Data Collection.** We collected panamá fruits from areas near RBLB for our experiment. 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
145  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, they choose from multiple available fruits in a
150  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
155  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
160  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
165  observed them. Further information about fruit collection, data collection, and observer training can be found in the electronic supplemental material text and video, in addition to descriptions (Table S2) and video of panamá processing techniques.

## 3. Statistical Analyses

170  We analysed these data using multilevel experience-weighted attraction (EWA) models [39;40]. EWA models are a family of models that link individual learning rules and social information use to population-level dynamics by fitting existing mathematical models of learning as statistical models [16;17;41].

175  3.1. **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 and how the modelling framework incorporates them.

180  *Payoff-biased learning.* Copying the behaviour with the highest observable payoff is a useful social learning strategy [22;42]. 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 [43] while wild tufted capuchins bias attention toward the most
185  efficient tool-users [44]. Cues of payoff may be noisy, however, and different individuals may require different techniques.

*Model-biased learning.* Sometimes evaluating the content of a behaviour is costly or impossible. In these circumstances, it may be an adaptive heuristic to bias attention toward particular demonstrators or "models," who display
190  cues (i.e. rank, health, fertility) that are likely to be correlated with adaptive behaviour.
   Prestige-biased learning is a popular example of model bias in humans [45]. While animals may lack the concept of prestige, they have analogues. Captive chimpanzees have been found to be more likely to copy dominant
195  individuals [41;46], while vervets copy same-sex high-ranking individuals [47].
   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 [48]. *Luehea* processing techniques of
200  capuchins at RBLB were predicted by both the technique their mother used and the technique they saw performed most often [30]. Kin-biased learning has been found in carnivores [49–51], 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
205  strength, size, or cognitive ability, it might be beneficial for learners to copy those who are most similar to them. Sex-biased learning has been found in several primate species [30;47].

*Frequency-dependent learning.* Frequency-dependent social learning occurs when frequency among demonstrators or frequency of demonstration influ-
210  ences 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 the fixation and main-
215  tain the stability of a cultural trait [10;18]. Experiments in many captive [20;52–55] and some wild [23;24] animals have found evidence of conformist learning.

3.2. **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 the probability of each option being chosen. 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 $\phi_j$ 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 $\lambda$ controls how strongly differences in attraction influence choice. When $\lambda$ 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 $T_{\text{open}}$, this ensures that payoffs decline as $T_{\text{open}}$ increases, but with the steepest declines early on.

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

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

where $\gamma_j$ 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:

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

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 $\beta_k$ 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. $\kappa$ 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 $\gamma_j$ and $\phi_j$ 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 $\phi_j$ for an individual $j$ is defined as:

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

where $\alpha_j$ is a varying intercept per individual and $\mu_\phi$ 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 of observed foraging bouts. This allows new social information to be used, while old information is discarded. We tested the sensitivity of the time window used to calculate social cues and found our results were robust to variations in window width (7, 14, 21, 28 days) (Table S4). Attempts to parametrise window width fit poorly. 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 [34], in R version 3.3.2 [56]. We compared models using WAIC [57]. To check our approach, we simulated the hypothesized data generating process and payoff structure 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.

## 4. RESULTS: INNOVATION AND DIFFUSION OF TECHNIQUES

Of the 25 individuals in the group, 23 tried to process panamá and 21 were successful at least once over 75 experimental days. We observed 7 types of predominant fruit processing techniques on 1441 fruits, which varied in time required and the proportion of successful attempts (Table S2). Mean

FIGURE 1. Adult male NP exhibits the canine seam technique.

(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 (figures 2,S3,S4). 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 (figure S4). However, canine seam never reached fixation in the population.

4.1. **Results of EWA models.** There was overwhelming support for some mix of individual and social learning over individual learning alone (Table S3). The highest ranked model was the global model containing all strategies and age effects on learning parameters, which received 94% 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.

Marginal posterior distributions of each parameter are displayed in Table 1 and visualized in figure S1. Note that the marginal posterior distribution of each parameter cannot be directly interpreted as the importance of each factor in the total diffusion of behaviour. The weight of social information ($\gamma$), for example, can be relatively small at each instantaneous choice but still be decisive in determining which behaviour spreads, because individual discovery rates may be even smaller. As each individual's behaviour is unique to their observed social information, personal experience, and estimated individual-level parameters, we encourage readers to view marginal predictions with visualizations of implied individual behaviour, using posterior predictive distributions in the electronic supplementary material (figure S3) .
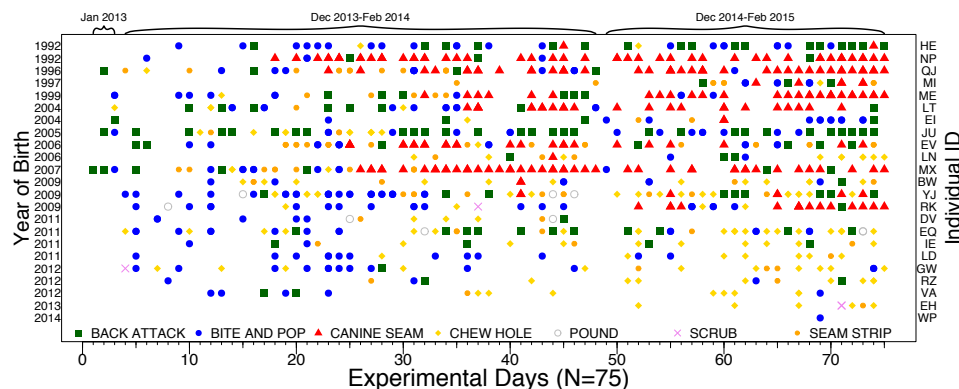
FIGURE 2. Techniques observed during experiment. Rows are unique individuals, from oldest (top) to youngest (bottom). X-axis is sequential order of experimental days. Each color/shape represents most common technique used by an individual on that day, no point indicates days of no processing. The most successful technique indicated by red triangles (canine seam) diffused to older members of the population. Younger individuals did not use canine seam.

**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— a bias toward copying rare behaviours. The marginal payoff-bias coefficient is strongly positive, indicating attraction to high-payoff actions. Figure 3 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 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 ($\phi$):** 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 4a). This suggests that older individuals are more canalized than younger individuals.

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 | $\lambda$ | $\phi$ | $\gamma$ | $f$ | $\beta_{pay}$ | $\beta_{kin}$ | $\beta_{rank}$ | $\beta_{coho}$ | $\beta_{age}$ | $\mu_\phi$ | $\mu_\gamma$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 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 | | |

**Weight of social information ($\gamma$):** $\gamma$ estimates for individuals varied considerably, ranging from 0.07-0.39 ($\sigma_{individual} = 0.66$). $\gamma$ was also negatively related to age ($\mu_{age} = -0.10$; 89% CI $[-0.18, -0.03]$; Fig. 4b). This suggests that younger individuals rely more on social cues.

**Age bias ($\beta_{age}$):** 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 all 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 little 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 novel food processing techniques. 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 behavioural 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.
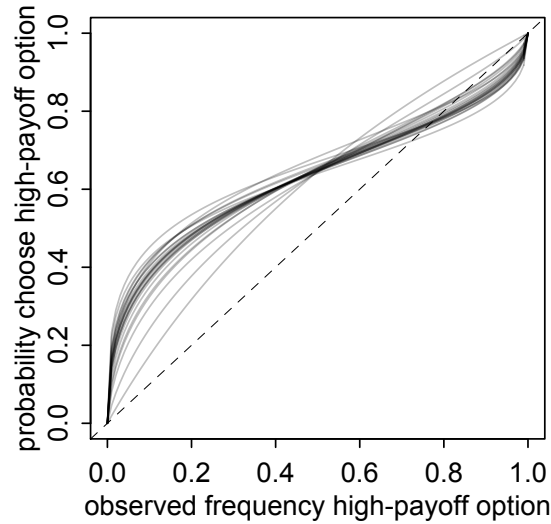
FIGURE 3. Posterior predictions of probabilities of choosing a socially observed option with payoff $\log(t_{\text{open}})^{-1} = 0.5$, relative to an observed option that was not successfully opened.
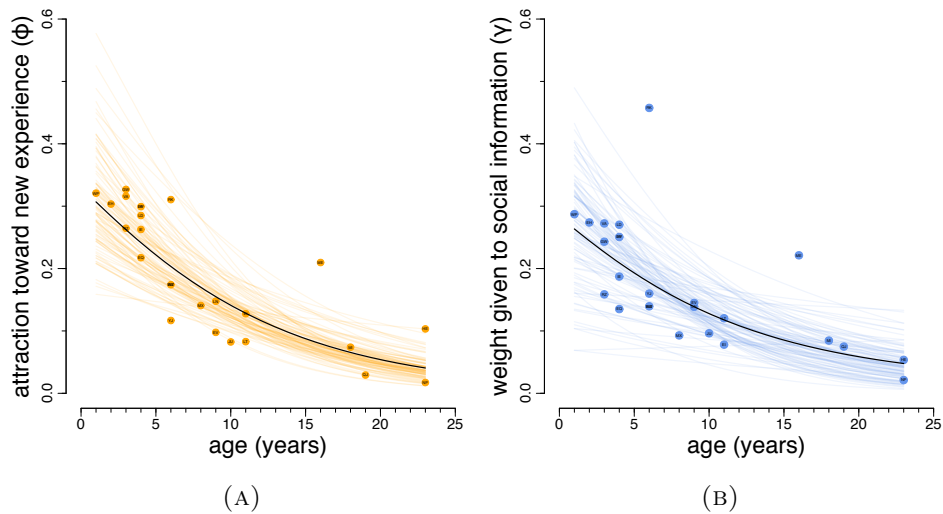


(A)                              (B)

FIGURE 4. Relationships between age and (A) attraction to new experience ($\phi$) and (B) influence of social information ($\gamma$). Black line represents posterior mean. Solid points are posterior means of individual varying effects. Lighter lines are 100 posterior samples.

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 captive capuchins [26;58]. 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 capuchins, like other animals [59], may be able to acquire new, efficient behaviour in a matter of days or weeks if knowledgeable models are available. 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 2). Juveniles born before 2009 did not use the canine seam technique (figure S4 and 2), 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 S4). This result is consistent with other research [60] suggesting that social learning guides 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 associate with who did not use canine seam (HE and MI respectively) 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 4b) and that older individuals were less likely to observe conspecifics (figure S5).

We also observed that older individuals were less likely to update information and had a greater attraction to previous experiences (figure 4a). This might be due to older individuals being less exploratory than younger individuals. One alternative explanation is that older individuals' higher success rates at processing panamá provided them with higher quality personal information to discern between the efficiency of varied processing techniques (figure S4). 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 multilevel, 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 (i.e sinosoidal curves and conformity) [24;61]. In our approach any population-level patterns are consequences of inferred (and potentially different) strategies among individuals (visualized in figure S3); they are not themselves used to make inferences about learning.

Our approach is most similar to network-based diffusion analysis (NBDA) [62;63]. In principle, our framework and NBDA can be analogized, despite differences in the details of modelled strategies, because both are multinomial time series modelling 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 typically employed NBDA in that it: 1) uses a full dynamic time series for available social information rather than a static social network and 2) emphasizes modelling the entire behavioural sequence including and beyond the first putative instance of social transmission. There is no reason in principle that ordinary NBDA models could not make similar use of these data, and recent advances [59] utilize dynamic social networks.

It is important to note that successfully fitting these dynamic, multilevel models benefits from recent advances in Monte Carlo algorithms. We used

an implementation of Hamiltonian Monte Carlo (NUTS2) provided by Stan [34]. 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 param-
465 eters, 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
470 tradition. However, social learning may increase within-group homogeneity, while individual learning may act to decrease it [51]. 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
475 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.

480 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
485 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
490 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
495 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
500 evaluate who individuals choose to bias attention toward when in the proximity of potential demonstrators to see how positive assortment due to social preferences, rank, or food sharing 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 [64]. 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 juveniles engage in kin-biased learning [65], but only interact 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. Acknowledgements

## 8. Research Funds

## 9. Data Accessibility

Data and code for models, simulations, and graphs are available at `https://github.com/bjbarrett/panama1`.

## References

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

[2] Kawamura S. The process of sub-culture propagation among Japanese macaques. Primates. 1959;2(1):43–60.

[3] Marler P, Tamura M. Culturally transmitted patterns of vocal behavior in sparrows. Science (New York, NY). 1964 Dec;146(3650):1483–1486.

[4] Bonner JT. The Evolution of Culture in Animals. Reprint edition ed. Princeton, N.J.: Princeton University Press; 1983.

[5] Whiten A, Goodall J, McGrew WC, Nishida T, Reynolds V, Sugiyama Y, et al. Cultures in chimpanzees. Nature. 1999 Jun;399(6737):682–685.

[6] van Schaik CP, Ancrenaz M, Borgen G, Galdikas B, Knott CD, Singleton I, et al. Orangutan Cultures and the Evolution of Material Culture. Science. 2003 Jan;299(5603):102–105.

[7] Perry S, Baker M, Fedigan L, Gros-Louis J, Jack K, MacKinnon KC, et al. Social Conventions in Wild White-faced Capuchin Monkeys: Evidence for Traditions in a Neotropical Primate. Current Anthropology. 2003 Apr;44(2):241–268.

[8] Whitehead H. Cultural Selection and Genetic Diversity in Matrilineal Whales. Science. 1998 Nov;282(5394):1708–1711.

[9] Rendell LE, Whitehead H. Culture in whales and dolphins. Behavioral and Brain Sciences. 2001;24:309–382.

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

[11] Laland KN. Social learning strategies. Animal Learning & Behavior. 2004 Feb;32(1):4–14.

[12] Rendell L, Boyd R, Cownden D, Enquist M, Eriksson K, Feldman MW, et al. Why copy others? Insights from the social learning strategies tournament. Science. 2010;328(5975):208–213.

[13] Mesoudi A, Chang L, Dall SRX, Thornton A. The Evolution of Individual and Cultural Variation in Social Learning. Trends in Ecology & Evolution. 2016 Mar;31(3):215–225.

[14] Farine DR, Spencer KA, Boogert NJ. Early-Life Stress Triggers Juvenile Zebra Finches to Switch Social Learning Strategies. Current Biology. 2015 Aug;25(16):2184–2188.

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

[16] McElreath R, Lubell M, Richerson PJ, Waring TM, Baum W, Edsten E, et al. Applying evolutionary models to the laboratory study of social learning. Evolution and Human Behavior. 2005 Nov;26(6):483–508.

[17] McElreath R, Bell AV, Efferson C, Lubell M, Richerson PJ, Waring T. Beyond existence and aiming outside the laboratory: estimating frequency-dependent and pay-off-biased social learning strategies. Philosophical Transactions of the Royal Society B: Biological Sciences. 2008;363(1509):3515–3528.

[18] Henrich J, Boyd R. The Evolution of Conformist Transmission and the Emergence of Between-Group Differences. Evolution and Human Behavior. 1998 Jul;19(4):215–241.

[19] Nakahashi W, Wakano JY, Henrich J. Adaptive Social Learning Strategies in Temporally and Spatially Varying Environments. Human Nature. 2012 Aug;23(4):386–418.

[20] Day RL, MacDonald T, Brown C, Laland KN, Reader SM. Interactions between shoal size and conformity in guppy social foraging. Animal Behaviour. 2001 Nov;62(5):917–925.

[21] Whitehead H, Richerson PJ. The evolution of conformist social learning can cause population collapse in realistically variable environments. Evolution and Human Behavior. 2009;30(4):261 – 273.

[22] Baldini R. Two success-biased social learning strategies. Theoretical population biology. 2013 Jun;86:43–49.

[23] van de Waal E, Borgeaud C, Whiten A. Potent Social Learning and Conformity Shape a Wild Primate's Foraging Decisions. Science. 2013 Apr;340(6131):483–485.

[24] Aplin LM, Farine DR, Morand-Ferron J, Cockburn A, Thornton A, Sheldon BC. Experimentally induced innovations lead to persistent culture via conformity in wild birds. Nature. 2015 Feb;518(7540):538–541.

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

[26] Fragaszy DM, Visalberghi E, Fedigan LM. The complete capuchin: the biology of the genus Cebus. Cambridge; New York: Cambridge University Press; 2004.

[27] Jerison H. Evolution of the Brain and Intelligence. Academic Press; 1973.

[28] Rose LM, Perry S, Panger MA, Jack K, Manson JH, Gros-Louis J, et al. Interspecific Interactions Between Cebus capucinus and Other Species: Data from Three Costa Rican Sites. International Journal of Primatology. 2003 Aug;24(4):759–796.

[29] Panger MA, Perry S, Rose LM, Gros-Louis J, Vogel E, MacKinonn KC, et al. Cross-site differences in foraging behavior of white-faced capuchins (Cebus capucinus). American Journal of Physical Anthropology. 2002;119(1):52–66.

[30] Perry S. Conformism in the food processing techniques of white-faced capuchin monkeys (Cebus capucinus). Animal Cognition. 2009 Sep;12(5):705–716.

[31] O'Malley RC, Fedigan LM. Evaluating social influences on food-processing behavior in white-faced capuchins (Cebus capucinus). American Journal of Physical Anthropology. 2005;127(4):481–491.

[32] Perry S. Social traditions and social learning in capuchin monkeys (Cebus). Philosophical Transactions of the Royal Society B: Biological Sciences. 2011 Apr;366(1567):988–996.

[33] Sol D, Sayol F, Ducatez S, Lefebvre L. The life-history basis of behavioural innovations. Phil Trans R Soc B. 2016;371(1690):20150187.

[34] Stan Development Team. RStan: the R interface to Stan; 2016. V. 2.14.0. http://mc-stan.org/rstan.html.

[35] Perry S, Godoy I, Lammers W. The Lomas Barbudal Monkey Project: Two Decades of Research on Cebus capucinus. In: Kappeler PM, Watts DP, editors. Long-Term Field Studies of Primates. Springer Berlin Heidelberg; 2012. p. 141–163.

[36] Frankie GW, Vinson SB, Newstrom LE, Barthell JF. Nest Site and Habitat Preferences of Centris Bees in the Costa Rican Dry Forest. Biotropica. 1988;20(4):301–310.

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

[38] Janzen DH. Escape in Space by Sterculia Apetala Seeds from the Bug Dysdercus Fasciatus in a Costa Rican Deciduous Forest. Ecology. 1972 Mar;53(2):350.

[39] Camerer C, Hua Ho T. Experience-weighted Attraction Learning in Normal Form Games. Econometrica. 1999;67(4):827–874.

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

[41] Kendal R, Hopper LM, Whiten A, Brosnan SF, Lambeth SP, Schapiro SJ, et al. Chimpanzees copy dominant and knowledgeable individuals: implications for cultural diversity. Evolution and Human Behavior. 2015;36(1):65–72.

[42] Baldini R. Success-biased social learning: cultural and evolutionary dynamics. Theoretical population biology. 2012 Nov;82(3):222–228.

[43] Pike TW, Kendal JR, Rendell LE, Laland KN. Learning by proportional observation in a species of fish. Behavioral Ecology. 2010 May;21(3):570–575.

[44] Ottoni EB, Resende BDd, Izar P. Watching the best nutcrackers: what capuchin monkeys (Cebus apella) know about others' tool-using skills. Animal Cognition. 2005 Oct;8(4):215–219.

[45] Henrich J, Gil-White FJ. The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. Evolution and Human Behavior. 2001 May;22(3):165–196.

[46] Horner V, Proctor D, Bonnie KE, Whiten A, Waal FBMd. Prestige Affects Cultural Learning in Chimpanzees. PLOS ONE. 2010 May;5(5):e10625.

[47] van de Waal E, Renevey N, Favre CM, Bshary R. Selective attention to philopatric models causes directed social learning in wild vervet monkeys. Proceedings of the Royal Society B: Biological Sciences. 2010;277(1691):2105–2111.

[48] McElreath R, Strimling P. When Natural Selection Favors Imitation of Parents. Current Anthropology. 2008 Apr;49(2):307–316.

[49] Mazur R, Seher V. Socially learned foraging behaviour in wild black bears, Ursus americanus. Animal Behaviour. 2008 Apr;75(4):1503–1508.

[50] Müller CA, Cant MA. Imitation and Traditions in Wild Banded Mongooses. Current Biology. 2010 Jul;20(13):1171–1175.

[51] Thornton A, Clutton-Brock T. Social learning and the development of individual and group behaviour in mammal societies. Philosophical Transactions of the Royal Society B: Biological Sciences. 2011 Apr;366(1567):978–987.

[52] Chou LS, Richerson PJ. Multiple models in social transmission of food selection by Norway rats, Rattus norvegicus. Animal Behaviour. 1992 Aug;44, Part 2:337–343.

[53] Whiten A, Horner V, de Waal FBM. Conformity to cultural norms of tool use in chimpanzees. Nature. 2005 Sep;437(7059):737–740.

[54] Dindo M, Whiten A, de Waal FBM. In-Group Conformity Sustains Different Foraging Traditions in Capuchin Monkeys (Cebus apella). PLoS ONE. 2009 Nov;4(11):e7858.

[55] Pike TW, Laland KN. Conformist learning in nine-spined sticklebacks' foraging decisions. Biology Letters. 2010 Aug;6(4):466–468.

[56] R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2016. V. 3.3.1.

[57] Watanabe S. Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory. J Mach Learn Res. 2010 Dec;11:3571–3594.

[58] Fragaszy D, Visalberghi E. Socially biased learning in monkeys. Animal Learning & Behavior. 2004 Feb;32(1):24–35.

[59] Hobaiter C, Poisot T, Zuberbühler K, Hoppitt W, Gruber T. Social network analysis shows direct evidence for social transmission of tool use in wild chimpanzees. PLoS Biol. 2014;12(9):e1001960.

[60] Galef BG, Whiskin EE. Interaction of social and individual learning in food preferences of Norway rats. Animal Behaviour. 2001;62(1):41–46.

[61] Henrich J. Cultural Transmission and the Diffusion of Innovations: Adoption Dynamics Indicate That Biased Cultural Transmission Is the Predominate Force in Behavioral Change. American Anthropologist. 2001;103(4):992–1013. Available from: http://www.jstor.org/stable/684125.

[62] Hoppitt W, Laland KN. Detecting social learning using networks: a users guide. American Journal of Primatology. 2011 Aug;73(8):834–844.

[63] Allen J, Weinrich M, Hoppitt W, Rendell L. Network-Based Diffusion Analysis Reveals Cultural Transmission of Lobtail Feeding in Humpback Whales. Science. 2013 Apr;340(6131):485–488.

[64] Russon AE. Developmental perspectives on great ape traditions. In: Fragaszy DM, Perry S, editors. The biology of traditions: models and evidence. Cambridge; UK: Cambridge University Press; 2003. p. 329–364.

[65] Kline MA, Boyd R, Henrich J. Teaching and the Life History of Cultural Transmission in Fijian Villages. Human Nature. 2013 Oct;24(4):351–374.