bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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.

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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. Classic studies on the cultural transmission of birdsong (2) and food handling techniques in macaques (3) find their roots in Darwin's naturalist musings on the culture of nectar robbing in bees (cited in (4)). Following Bonner's seminal examination of the topic (5), research interest in animal cultures and social learning soared near the end of the 20th century. This renewed interest was heavily fueled by long-term cross-site collaborations within primatology (6; 7; 8) and cetaceology (9; 10) that provided observational evidence from the wild that animal culture might be ubiquitous—particularly in large-brained, socially complex animals. As the diversity of taxa in which social learning is studied grows, increasing evidence is accumulating that traditions might be more widespread and ecologically meaningful than was previously appreciated (11; 12; 13; 14).

Increasing exploration of cultural transmission has also 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 (15). 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 to copy in a population of multiple potential demonstrators. Evaluating all available social information can be costly, particularly if the function of the behaviour is not understood. To cope with these difficulties, organisms use heuristics (15) and strategies (16; 17) 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 (18; 19; 20; 21).

Different strategies have different advantages and disadvantages. Two families of social learning strategies that have received both theoretical and empirical attention are conformity and payoff-bias (15; 22; 23). Conformist transmission, also known as positive frequency dependence, can be broadly adaptive, especially in the context of spatial environmental variation (15; 24; 25). However, unless it is combined with other, flexible strategies, conformity may reduce or prevent the spread of adaptive innovations or even cause population collapse (26). Payoff-biased social learning attends to behaviour that is associated with higher payoffs, such as a higher rate of food acquisition. In contrast to conformity, payoff-biased social learning is very effective at spreading novel adaptations. However, it can be outperformed by conformity or even unbiased social learning once adaptive behaviour is common (27; 28).

There is empirical evidence for both conformist and payoff-biased social learning in humans (23). In other animals, conformity has been studied more extensively than payoff-bias. For example, (29) and (30) find evidence consistent with positive frequency dependence in great tits and vervet monkeys,

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

respectively. 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 specific models of conformist and payoff-biased social learning. We investigated the innovation and transmission of extractive foraging techniques used to access the protected seeds of the *Sterculia apetala* fruit. This fruit occurs over much of the full range of *C. capucinus*, but only some subpopulations are experienced with it. By introducing the fruit in controlled settings, we were able to observe how new foraging traditions arose and spread through a wild group of monkeys. We use these data to investigate which social learning strategies best predict individual behaviour and how they influence the origins and maintenance of traditions. We also explore individual variation in learning and its relationship with age.

We accomplish this statistical analysis with a multilevel Bayesian dynamic learning model, of the form developed by (23). This model allows us to estimate for each individual in the sample a unique combination of social and individual learning strategies. The analysis takes advantage of dynamic social network data, which were available during each field experimental session. It also makes it possible to study the relationship between age, rank, or any other individual state and individual learning strategy. The multilevel Bayesian approach makes it possible to apply these models to field data that lack the precise balance required of other statistical approaches. 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. Over time, techniques that required less time to access the seed spread socially through the group and replaced other techniques. The statistical analysis suggests that payoff-biased social learning was responsible for this diffusion. We find no evidence of conformity, but do find weak evidence of anti-conformity, in which rarer 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.

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2. Study Design

2.1. **Field site and study system.** This study was conducted on a group of habituated white-faced capuchin monkeys, *Cebus capucinus*, dwelling in Reserva Biológica Lomas Barbudal (RBLB) and two private ranches adjacent to RBLB in northwest Costa Rica, during the months of December-February from December 2012 through February 2015. Capuchins provide an ideal study system for understanding social learning and traditions. They are extremely tolerant of foraging in proximity with conspecifics (31), independently evolved many brain correlates associated with intelligence (32; 33; 34; 35) and display the largest recorded repertoire of candidate behavioural traditions of any platyrrhine including diet preferences (36), social conventions (8), interspecific interactions (37) and extractive foraging techniques (38; 39; 40). Their reliance on social learning, frequency of innovation, and complexity of social interactions exemplifies what is predicted for long-lived primates with a slow life history strategy (41; 42).

RBLB has been the location of the Lomas Barbudal monkey project which, between May 1990 and the end of Feb 2015, collected 96,250 hours of observational behavioural data on 538 monkeys, and determined genetic parentage for most of these individuals (43). The habitat is neotropical dry forest, including much riparian forest, and is marked by a rainy wet season from May-November with 1,000-2,000 mm of rain (44), and a dry season marked by little to no rain and the deciduous shedding of the leaves of most tree species to cope with water limitation. Long-term research at this field site has made it possible to document many unique aspects of capuchin biology and behaviour that would not have been possible with short term observations (45), including many social conventions and other behavioural traditions.

2.1.1. *Extractive foraging techniques as behavioural traditions.* Capuchins use their intelligence and dexterous hands to exploit difficult to access resources via extractive foraging. This heavy reliance on extractive foraging is notable among monkeys (31), and has marked capuchins as an ideal comparative study system for understanding the evolution of extractive foraging in humans. In neotropical dry forests, capuchins increase their reliance on extractive foraging and expand diet breadth during the transition between the wet and dry season when resources are limited. Capuchins receive more close-range, directed attention from conspecifics when they are foraging on structurally protected fruits that require multiple steps to open and are large (46). Many of the techniques required to access structurally protected fruits have been proposed as potential behavioural traditions (38; 47).

Panamá fruits, *Sterculia apetala*, are an important dietary staple of capuchins at RBLB and comprise a large percentage (8%) of the diet of most capuchin groups in the early dry season (46). The fruits are *empanada* shaped and the fatty, protein rich seeds within are protected by a hardened outer husk, sticky exudates, and stinging hairs which potentially evolved

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

to aid in mammalian seed dispersal (48). Instead of waiting for the fruits to naturally 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, and is only exceeded by wasp nests (46). Panamá processing techniques are also observed to vary between groups at RBLB and other field sites in the area (Panger et al, 2002), suggesting that panamá processing techniques are socially-learned traditions. Interestingly, in his seminal work on *S. apetala* seed dispersal, (48) reported that wild-caught juvenile capuchins were unable to open the fruits and likely did not consume them prior to dehiscing.

Panamá foraging is noteworthy compared to other extractive foraging traditions. Processing techniques differ in efficiency, measured by the amount of time, on average, it takes to open a fruit with a particular technique. Techniques also differ in efficacy, both in their probability of being successful and due to the costs incurred by encountering irritating, stinging hairs. This contrasts with other extractive foraging traditions, particularly processing *Luehea* fruits, where the two main processing techniques show no difference in efficiency or efficacy (39). Near RBLB, panamá trees grow in riparian and/or evergreen habitat in primary forests where they are dominant canopy species, growing up to 35 meters. They produce crops from a few hundred to several thousand fruits, and often fruit in high densities some years followed by years of minimal or no fruiting.

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 1). Two trees were found in the territory during phenological surveys, but are at the periphery of the territory, have small crowns, and are in areas of the habitat shared with other groups of capuchins who regularly forage at them. When this study was designed, neither the authors nor other researchers at the field site had any recollection of ever 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 in their life. Efficiency at foraging for panamá markedly increased over the 3 years this experiment was conducted.

Several adults in the group (2 females and 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 adult males of unknown origin, it is unknown if they learned to process panamá fruits as

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

TABLE 1. Counts and percentages of total foraging time spent consuming panamá for RBLB study groups collected between July 2003 and December 2011. Each count is a activity point sample collected every 2.5 minutes from a 10 minute continuous focal follow. Rafiki's group was not observed foraging for panamá during focal follows since their late 2007 fission from Pelon group, however there are 14 panamá trees within their territory in which they have been observed foraging.

Group	**# times foraging**	**# times foraging panamá**	**% foraging time spent on panamá**
Abby's	34,263	78	0.23
Cupie's	4,431	3	0.07
Pelon	36,054	10	0.03
Flakes	17,200	0	0.00
Lost Boys	4,357	50	1.15
Musketeers	8,866	17	0.19
Newman's	3,951	48	1.21
Rafiki's	6,745	0	0.00
Rambo's	32,893	144	0.44
Splinter	2,944	18	0.61

juveniles, but this seems likely for at least one of them 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. **Fruit Collection.** The locations of 79 mature panamá trees near RBLB were recorded using GPS and compiled from previous phenological transects, the recollection of long term researchers, and surveys in ecological areas where they were most likely to exist. Panamá trees are easier to locate than many other species in the forest as they are often large, primarily grow in riparian areas or near annual springs, have conspicuous buttresses, and large palmate leaves that can easily be noticed in the canopy from a distance and densely litter the forest floor. Panamá fruits used in the study were collected from the forest floor and canopy at trees where fruits showed evidence of capuchin foraging. Fruits chosen were of early stage ripeness, had not dehisced, had mature-sized seeds, showed no evidence of physical tampering (i.e. biting or scratching) by animals, and were used within 48 hours of being collected. Capcuhins will naturally forage for panamá fruit both in tree canopies and on the forest floor.

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2.3. **Foraging Data Collection.** 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 behaviours. Panamá fruits were placed on a 25 cm diameter wooden platform. Platforms provided a 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 out for two primary reasons. First, when individuals are naturally foraging for panamá, there are multiple fruits in a tree from which they choose. Second, we wanted to see whom they chose to bias their attention toward when given a choice of multiple potential demonstrators. While many learning experiments often 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 animals learn in the wild. How they use information (i.e. via conformity or payoff-bias), must be consider jointly with whom they choose to observe (i.e. kin or cohort-mates).

Fruits were set out under the cover of a poncho when the monkeys were not looking in our vicinity. Ponchos obscured the monkey's view of us manipulating the fruits and platform, and since ponchos were worn regularly when not experimenting, the 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. For processing behaviours, we verbally described how they were processing each fruit, and whether or not they were stung by hairs.

In addition to processing techniques of foraging animals, we recorded the behaviours and interactions with potential audience members. The location of potential audience members relative to the focal individual was recorded as one of: in contact, <1 , <5, <10, or <20 adult monkey lengths. One adult capuchin monkey length is approximately 0.75 m. Monkeys were assumed to be observing processing if they stopped motion and focused their gaze on the processing individual for more than 3 seconds. If they briefly glanced at the foraging individual, or focused their stare on them while the demonstrator was ingesting seeds but did not observe processing, they were assumed to be not acquiring information about food processing. After each experimental bout, fruits were recovered and photographed to note technique-specific markings and validate the observations of researchers. Only 4 panamá processing events were not viewed by the observer due to an animal darting away

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

quickly with a fruit. Unfortunately, an estimated 114 processing bouts are missing from this data set collected on days 39,40,42,43,45,48, and 49 that were lost by a research assistant. While we have photographs of most of these processed fruits, we did not analyze them or impute processing times. Videos of several of the processing techniques and corresponding pictures of processed fruits can be found in the supplemental video.

3. Results: Innovation and diffusion of techniques

Processing techniques were sorted into 7 categories based on the predominant technique used to process the fruit (Table 2). Techniques varied in both the time required to reach the seed, when successful, and the proportion of attempts resulting in success. Mean (median) duration ranged from 50 (29) seconds to 330 (210) seconds. Proportion of attempts resulting in success ranged from 0.38 to 0.89.

The frequencies of different techniques changed over time both in the group and in most individuals (Figure 2). The most successful technique, canine seam, went from being non-existent in the group to becoming the most common technique employed, particularly for older individuals. However, it never reached fixation in the population as juveniles born before 2009 almost never performed the technique. Another technique, bite and pop, had a slightly lower estimated mean time to open (when opened) but had a much lower overall success rate compared to canine seam (Table 2). This technique only worked for strong, older individuals on riper fruits. The two techniques which had no success, pound and scrub, were rarely observed and only observed in juveniles. The canine seam technique first appeared on experimental day 15, when a socially peripheral immigrant adult male (NP) displayed it with his second attempt at processing a fruit, which he opened in 6 seconds. He continued to open fruits in this manner, with limited observation until it diffused to other group members. Two knowledgeable adults, an adult female (ME) and the alpha male (QJ), switched to the more efficient canine seam technique while all others tried it at least once (Figures 2, S3).

4. Results: Learning strategies

Our goal is to directly compare the explanatory power of alternative models of social learning. A satisfactory approach should both model the temporal dynamics of behaviour as well as accommodate the substantial imbalance in sampling among individuals. Therefore, we analyzed these data using a series of hierarchical experience weighted attraction (EWA) models (49; 50). EWA models are a flexible family of statistical models that directly use mathematical models of social learning to estimate the learning strategy or strategies used by individuals, and link individual learning rules to population-level dynamics. They were originally developed in behavioural economics to estimate preferences for learning strategies in economic games,

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

TABLE 2. Summary statistics for the 7 panamá processing techniques observed in this study. Mean and median duration presented in seconds.

Technique	**Description**	**Mean**	**Median**	**% Open**	**N**
Back attack	peel fibers off back from fruit with seam facing away from mouth, bite to pop open at seam	169.0	119.0	51.1	176
Bite and pop	bite opposite corners of each fruit forcefully, bite to pop open at seam	49.7	29	37.8	283
Canine seam	hold fruit perpendicular to mouth, insert upper and lower canines into seam to split open	70.5	42	88.5	511
Chew hole	chew hole or rip fibers off fruit at corner, back, or side, seam not chewed	330.5	211.5	65.5	247
Pound	pound fruit on hard substrate	n/a	n/a	0	15
Scrub	scrub fruit on hard substrate	n/a	n/a	0	5
Seam Strip	hold fruit parallel to mouth, strip fibers off along the seam, bite to pop open at seam	130.6	211.5	65.0	200
All techs		131.5	95	65.6	1437

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

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

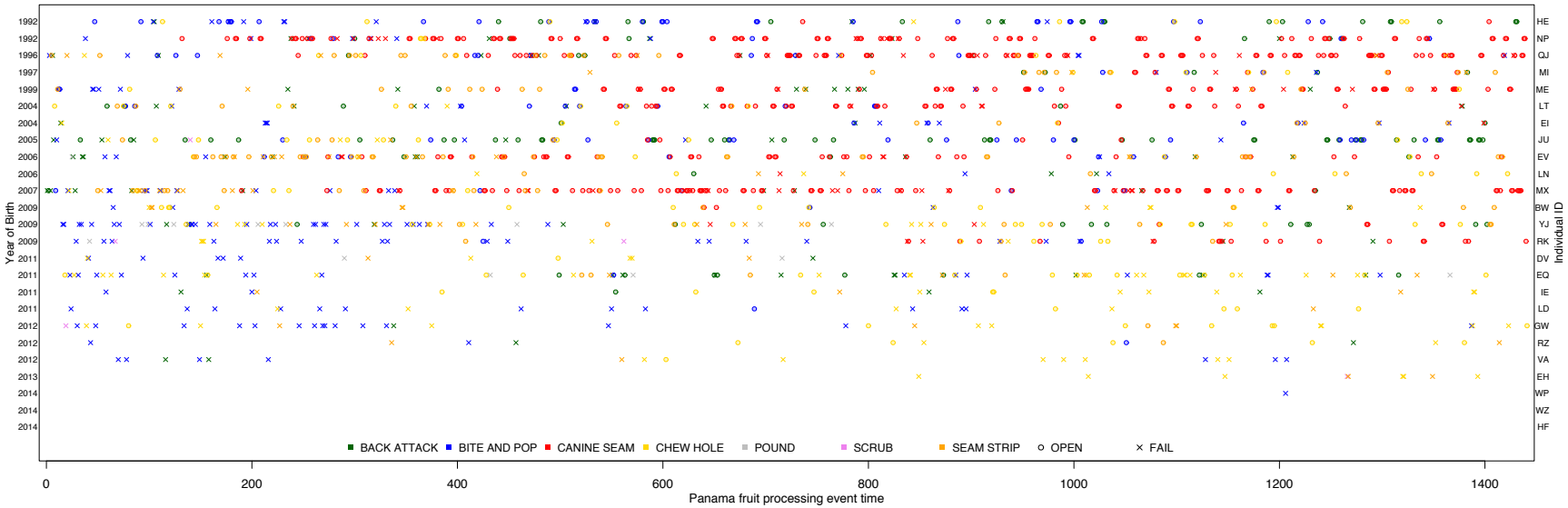


FIGURE 2. Plot of techniques used by all individuals during the experiment from raw data. Rows are unique individuals, and are arranged from top to bottom by decreasing age. X-axis is event time with each point representing a unique fruit in chronological experimental order. Each color represents a unique processing technique, shaded circles are successful uses of the corresponding technique, while an "x" indicates a failed attempt. For older individuals, the most successful technique in red (canine seam) diffused to older members of the population over the course of the experiment. Younger individuals toward the bottom of the chart did not use the canine seam technique, had more failed attempts at opening the fruit, and fewer processing attempts overall when compared to older individuals.

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

but have also been used in studies of human (22; 23) and non-human social learning (51). Here we present the first fully Bayesian hierarchical analysis of social learning using EWA models, as well as the first analysis of social learning in wild non-human animals using EWA models.

4.1. **Social learning strategies.** Our main focus is the contrast between two well-studied types of social learning, conformity and payoff-bias. However, there are several other strategies of social learning that we also investigate in the a common modeling framework. We quickly describe the background of these strategies. We then describe how the modeling framework incorporates them.

Content biases: payoff-biased learning. Copying the most successful observable behaviour, or behaviours with the highest payoff, is an intuitively useful social learning strategy (27; 28). In a foraging context, selectively copying fitness-maximizing behaviours can increase the efficiency of diet and resource acquisition. Guppies will choose food patches with higher return rates (52), and field studies suggest that wild tufted capuchin monkeys bias attention toward the most efficient tool-users (53). However, surprisingly little work has examined whether animals use pay-off biased social learning. Part of this may be a consequence of early methods used to identify traditions such as the group contrasts method, where behaviours that differed between groups that could not be explained by genetics or ecology were proposed as candidate traditions (6). This likely overlooked many adaptive behaviours that differ in payoff; despite being socially learned the best behaviour is less likely to differ between groups than more "arbitrary" traditions.

Model-biased learning. In situations where the content of a behaviour cannot be evaluated, individuals might bias their attention toward particular demonstrators in a population. These model biases (54) can be efficient shortcuts to acquiring behaviours. Simply copying successful or high-ranking demonstrators might be correlated with some aspect of success, and does not require the cognition or time required to evaluate the content of behavioural choices. Prestige-biased learning is a common example of a model bias in humans (55). While many animals may lack the concept of prestige, they may have similar analogues. Captive chimpanzees have been found to be more likely to copy dominant, high ranking individuals (56). This rank bias in chimps was replicated by (51), who also found evidence consistent with copying successful individuals. An experiment using a single 2-action puzzle-box displayed to a group of free-ranging vervet monkeys (57) suggests that vervets adopt the behaviour of the high ranking members of their own sex. This bias, at least in philopatric females, might maintain group specific traditions. However, it is difficult to determine if this was rank-bias, similarity-bias (i.e. copy the same sex), or both, as high ranking individuals monopolized the puzzle boxes.

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Copying the behaviours of one's parents, or vertical transmission, is another widespread learning strategy. If a parent can survive and successfully reproduce, its offspring's mere existence serves as a reliable cue that her parents must have gotten something right by surviving long enough to produce offspring (58). In wild white-faced capuchins *Luehea candida* processing techniques for individuals were predicted in part by the technique their mother used and in part by the non-maternal technique they saw performed most often. This kin-biased effect was also stronger in females than males (39). Kin-biased learning has been found in many carnivores including black bears (59) and members of the family *Herpestidae* (60; 61). However, it can be difficult to tease apart whether this is due to cognition or is a consequence of family-unit social systems.

Copying similar individuals can be an adaptive heuristic, especially if behaviours differentially suit members of a population based of a trait. In populations where organisms differ in ecology, strength, size, strength, or cognition due to age or sex, 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 (62), while evidence for sex-biased learning has been found in several primate species (39; 57).

4.2. **Frequency-dependent learning.** Frequency-dependent learning is copying behaviours relative to their observed frequency in a population. It includes unbiased learning, positive frequency dependent learning, and negative frequency dependent learning. Unbiased learning, also known as random copying or linear imitation, is copying what is observed proportional to its frequency in a population. It is arguably the simplest form of learning and is likely widespread (63). Negative frequency-dependence, or anti-conformity, is preferentially copying rare behaviours in a group. It can also be thought of as a form of neophilia. Positive frequency dependence, more commonly known as conformity or majority-biased learning, is preferentially copying the most common behaviour observed in a group. To illustrate, if the most common trait in a group is observed n percent of the time, the chance of copying it is greater than n percent. Conformity is of particular interest because it can lead to fixation of a cultural trait in a group, maintain the stability of that trait, and is considered an important aspect of human evolution and behaviour (15; 24). For these reasons evidence for conformity has been sought out in many non-human animals. Captive experiments in guppies (64), sticklebacks (65), rats (66), chimpanzees (67) (but see (68)), and robust capuchins (69) have found evidence of social learning consistent with conformity. Wild experiments in vervets (70) and great tits (29) have also found evidence supporting conformity-biased learning.

4.3. **Model design.** An EWA model consist of two parts: a set of expressions that specify how individuals process and accumulate experience and a second set of expressions that turn accumulated experience into probability

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

distributions over choices. Accumulated experience is represented by *attraction* scores, A_{ijt}, which are unique to each behaviour i, individual j, and time t. A simple, common, and successful formulation is reinforcement learning, in which an individual payoff π_{ij} updates the attraction of individual j to behaviour i:

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

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

To turn these attraction scores into behavioural choice, some function that defines a probability for each possible choice is needed. In the simplest case, only individual attraction scores influence choice, and a standard multinomial logistic, or *soft-max* choice rule, is employed:

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

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. We label the above expression I_{ij} for ease of later reference and compactness of notation. In this study the individuals were assigned a payoff of zero, $\pi_{ij} = 0$, if they failed to open a panamá fruit. If they were successful, payoff was realized as the inverse-log amount of time it took to open the fruit, $\pi_{ij} = \log(t_{\text{open}})^{-1}$. 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, via a weight parameter. 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:

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

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_{ijt} by use of a multinomial probability expression with a log-linear component B_{ijt} that is an additive combination of cue frequencies. Specifically, we allow the probability of each option i, as a function only of social cues, be:

$$S_{ijt} = \frac{N^f_{ijt} \exp B_{ijt}}{\sum_m N^f_{mjt} \exp B_{mjt}} \tag{4}$$

This is easiest to understand in pieces. Ignoring the $\exp B_{ijt}$ terms for the moment, the N_{ijt} variables are the observed frequencies of each technique i

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

TABLE 3. Summary of payoffs and social cues used in social learning strategies & heuristics. Social payoffs and cues are specific to the observer depending on which foragers they observe, and correspond with equations 4-7

Social cue	Definition of κ
Payoff	Fail: 0 , Succeed: $\log(t_{\text{open}})^{-1}$
Rank	1 for α rank male and female, 0 others
Matrilineal kinship	1 for matrilineal kin, 0 for others
Age similarity	$(1 + \lvert\text{age}_{\text{forager}} - \text{age}_{\text{observer}}\rvert)^{-1}$
Age bias	$1 - \text{age}_{\text{forager}}$

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_{ijt} term:

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

which is just the sum of the products of the influence parameters β_k and the cue values $\kappa_{k,ijt}$. We consider five cues: payoff to behaviour, rank of demonstrator, matrilineal kinship of demonstrator to observer, age similarity, and age preference for older demonstrators. These cues are summarized in Table 3.

The final components needed to define the model 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 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 S2).

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. The full set of models is presented in the supplemental. Models

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

TABLE 4. Posterior medians and standard deviations for parameters from the global model. Estimates of $\sigma_{individual}$ are the standard deviations of varying effects for that parameter across all individuals.

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

were fit using the Hamiltonian Monte Carlo engine Stan, version 2.14.1 (71), in R version 3.3.2 (72). We compared models using the information criterion WAIC (73). WAIC is a generalized form of information criterion, of which the more familiar AIC is a special case. WAIC values were calculated and models were compared using the `compare` function from the `rethinking` package in R (74). To help ensure that our statistical model functions as intended, we simulated the hypothesized data generating process and recovered the true data-generating values from our simulated data. We chose conservative, weakly informative priors for our estimated parameters. This made our models skeptical of large effects and helped ensure chain convergence. Model code, data simulation code, graphing code, and raw data for this study are available at https://github.com/bjbarrett/panama1.

4.4. **Results of EWA models.** Of the evaluated models, there was overwhelming support for some mix of individual and social learning over individual learning alone (Table S1). This is unsurprising. The highest ranked model according to WAIC values 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 (Table 4). Visualizations of the posterior distributions estimated for the global model can be found in Figures S1 and S2.

Individual marginal posterior distributions of each parameter are displayed in Table 4. We quickly summarize these distributions below. However, keep in mind that while the marginal posterior distribution of each learning parameter provides some information, the model is too complicated to always interpret these parameters directly. For example, the weight of

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

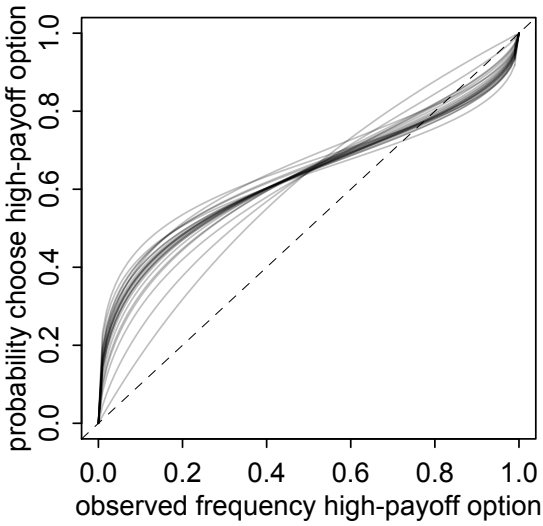
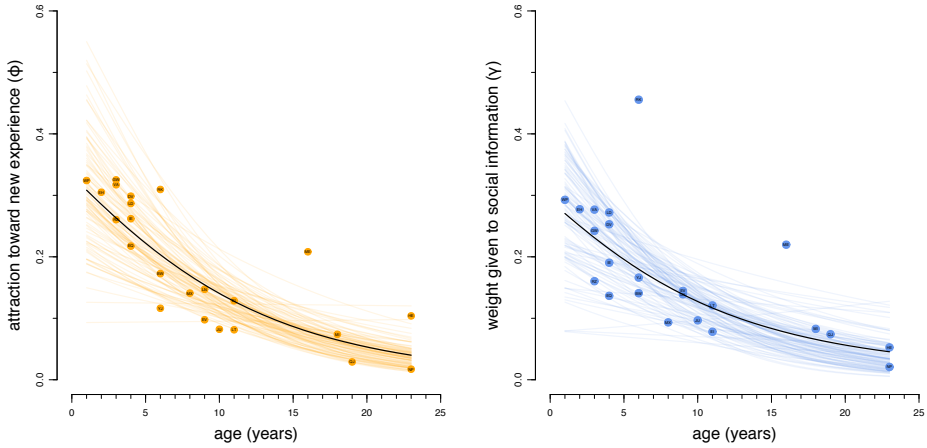


FIGURE 3. Posterior predictions of probabilities of choosing a socially observed option with payoff $\log(t_{\text{open}})^{-1} = 0.5$, relative to an observed option that was not successfully opened.

social information γ applies only at each choice. It is not therefore a simple partitioning of the importance of social versus individual information. 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 (Supp Fig S3).

Influence of conformity and payoff-bias (f and β_{pay}): The raw marginal conformist exponent is below 1 on average, indicating mild anti-conformity, if anything. The marginal payoff-bias coefficient is strongly positive, indicating attraction to high-payoff actions. However, these parameters always interact to influence social learning. In order to appreciate their joint action, we visualize the individual social learning functions 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 monkey. The diagonal dashed line represents unbiased social learning, in which each option is chosen in proportion to its observed frequency among tutors. 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

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



(A) ϕ: attraction to new experience (B) γ: influence of social information

FIGURE 4. Relationships between age and (A) attraction to new experience (ϕ) and (B) influence of social information (γ). Black line represents posterior mean relationship, solid points are posterior mean estimates of varying effects for each individual. Lighter lines are 100 randomly drawn posterior predictions. Darker shades indicated higher densities of posterior probability.

corner. There is no compelling evidence of positive frequency dependence under this model.

Weight of past experience (ϕ): On average, capuchins more heavily favor previous experiences over new ones ($\phi = 0.15; (0.11, 0.20)$ 89% posterior credible interval (PCI)) , Table 4 and Figure S1a). 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 (Figure 4a). Older capuchins have lower estimates of ϕ than younger capuchins ($\mu_{age} = -0.11; (-0.16, -0.06)$ 89% PCI). This suggests that older individuals are more canalized; they are less likely to change their behaviour and more heavily favor past experience, whereas younger individuals are more likely to switch behaviours.

Weight of social information (γ): This parameter estimates the relative weight of social learning compared to individual learning. Overall, individuals weigh social learning less than individual learning ($\gamma = 0.14; (0.08, 0.19)$ 89% PCI; Figure S1b, Table 4). Varying effects for individuals ranged from 0.07-0.39 ($\sigma_{individual} = 0.66$). Weight assigned to social learning was also negatively predicted by age ($\mu_{age} = -0.10; (-0.18, -0.03)$ 89% PCI ; Fig. 4b). This suggests that younger individuals rely more on social learning than older individuals.

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Contribution of age bias (β_{age}): Age bias contributed notably to social learning in our global model($\beta_{age} = 0.69$; (-0.79, 2.14) 89% PCI; Figure. S1h , Table 4), suggesting that capuchins were more likely to copy older individuals in their social group, Varying intercepts of β_{age} for individuals ranged from 0.64 to 0.85 ($\sigma_{individual} = 0.25$).

Contribution of age cohort bias (β_{coho}): Cohort bias contributed moderately to social learning in our global model ($\beta_{coho} = 0.48$; (-0.99, 2.00) 89% PCI; Figure S1g , Table 4), suggesting that capuchins also show some preference for copying individuals similar to them in age. Varying intercepts of β_{age} for individuals ranged from 0.43 to 0.64 ($\sigma_{individual} = 0.25$).

Contribution of kin bias (β_{kin}): Kin bias has a small contribution to social learning ($\beta_{kin} = 0.19$; (-1.22, 1.75) 89% PCI; Figure S1e , Table 4), suggesting that on average capuchins show a slight preference for copying their mother's techniques. Varying intercepts of β_{kin} for individuals ranged from 0.10 to 0.27.

Contribution of rank bias (β_{rank}): Rank bias has an estimated near zero effect on social learning ($\beta_{rank} = -0.11$; (-1.53, 1.33) 89% PCI; Figure S1f , Table 4).

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 find however strong evidence of payoff-biased learning. We also found evidence that other social cues, such as age, influence social learning, and that individual age was strongly related to the strength of individual reinforcement learning. 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 key advantages 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. It validates claims that food processing techniques are socially learned traditions (38; 40; 39). It has been notoriously challenging to find conclusive experimental evidence for social learning of object manipulation tasks in capuchins, despite much effort (75; 32). One of the important implications of this research is that evidence for social learning might be found across a broader range of taxa if more ecologically valid behaviours, or behaviours that differ in efficiency, are studied in the wild. This study also displays that animals, and primates in particular, can acquire new, more efficient behaviours quickly, in only a matter of a few days or weeks. This rapid pace

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of social transmission suggests that learning can act to rapidly facilitate behavioural responses to environmental change.

5.1.1. *Capuchins relied on pay-off biased and negative frequency dependent learning.* Our findings suggest that payoff-biased learning and negative frequency dependence guided diffusion of panamá processing techniques in this group (Table 4). These learning strategies are consistent with our observed finding that the rarest and most efficient panamá processing technique in the population, canine seam, eventually became the most common at the population level. This was the case for most, but not all, naive and knowledgeable adult and subadult individuals born after 2009 (Figure 2 and S3).Juveniles born before 2009 did not use the canine seam technique (Figure 2). We believe this is simply because their mouths were not large and strong enough to use this technique on panamá fruits.

Payoff-biased learning had the largest effect on the probability of choosing a behaviour, while negative frequency dependence discouraged 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, which can be interpreted as a type of neophilia. However since a rare behaviour was also the most efficient, studies to see if a less efficient, rare technique can invade a population are necessary to see the conditions under which these two primary learning strategies can be teased apart.

5.1.2. *Animals jointly use multiple social learning strategies.* While we found the strongest support for payoff-biased and negative-frequency dependent learning, our models suggest that animals use *multiple* social learning strategies simultaneously or that social biases and content biases might be equifinal. Age-biased learning and age-cohort biased learning also had support in the global model (Table 4). This could be due to older individuals' increased likelihood of being efficient panamá processors compared to juveniles. However, the preference for one individual (JU) to use the back attack and bite and pop techniques, despite often observing the more efficient canine seam technique, is consistent with age-bias, as he spent much of his time within observation range of his father, HE, who also preferred the back attack and bite and pop techniques.

Age cohort-biased learning also received support, suggesting individuals were more likely to copy the techniques of their age-mates. While this might be due to a psychological bias, it is also possible that this finding is a result of the physiological limitations of young capuchins to performing the most efficient technique previously mentioned. Individuals born in 2009 reached adult age (6 years) at the end of the study, and were consequently large enough to successfully employ canine seam in the second study year, but not the first. In the context of this experiment it is difficult to tease apart the physiological limitations and psychological biases. We do however, believe

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

that mathematically modeling age-cohort biased learning is a worthwhile endeavor, especially under conditions where proclivities to learn socially or innovate vary by age, or where behavioural phenotypes provide different fitness benefits to different demographic groups.

5.2. **Age predicts individual variation in social and individual learning.** There has been a call for researchers of cultural evolution and animal social learning to explore individual variation in social learning, as it might have meaningful evolutionary and social implications (18). We found that younger individuals more heavily relied on social learning than older individuals (Fig. 4b) and that older individuals were less likely to observe conspecifics (Supplemental Fig. S4). This age structure in proclivity to learn socially might suggest that instead of evolution favoring a single learning strategy for a particular species, flexible learning strategies that change over development might be favored by natural selection, particularly for long lived organisms in variable social environments. Individual variation might also be important in wildlife conservation and reintroduction efforts (17). If animals need to observe knowledgeable conspecifics to acquire behaviours to survive in the wild, limited funds for reintroduction can be targeted towards those individuals more likely to learn socially.

We also observed that older individuals had lower values of ϕ, suggesting that they were less likely to update information and had a greater attraction to previous experiences (Fig. 4a). 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 and had higher quality and quantities of personal information. Older individuals foraged more frequently and were more successful, thus giving them more opportunities to evaluate higher quality personal information (Figure 2).

Age effects unrelated to learning strategies may also indirectly affect whether and how individuals learn. In social contexts where resources are monopolized by dominant and older individuals, age structured differences may emerge in which younger individuals are more likely to acquire social information, because they are excluded from performing the behaviour and acquiring personal experience. Older individuals who are more likely to perform a behaviour on a monopolizable resource, may also be limited in the amount of social information they observe. This may create a unique opportunity where individuals of intermediate age learn more quickly and efficiently, because they are simultaneously acquiring personal and social information at a joint rate higher than the tails of the demographic pyramid. However, it is important to note that during our experiments, processing capuchins of all ages would often stop and focus their attention on another processing individual. Currently we lack quantitative theory about the implications of individual and age-related variation of social learning in affecting cultural transmission and evolutionary dynamics, and these empirical

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

findings are an important step toward developing theories we can evaluate in future studies.

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. This allows us to simultaneously consider contrasting hypotheses within the same modeling framework and investigate the relative influence of, for example, conformity and payoff-biased learning. 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. As a consequence, social and individual learning interact meaningfully in our statistical approach, rather than being antagonistic or social learning serving a merely parasitic role. 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 (76). For example, (29) found that the sigmoidal curve typical of conformity biased learning best described their data, aggregated at the group level. This approach is productive, but may waste information available in the full time series and contained in individual heterogeneity. It is also possible that other social learning strategies could result in a similar sigmoidal curve (77), although this is a risk to some extent with all model-based analysis. 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) (78). NBDA has provided insight into the presence of social learning, especially in wild animal populations (79; 62; 19). Like our approach, NBDA uses a description of social interactions, typically a matrix of dyadic associations, to model a time series of behaviour. 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 uses a full dynamic time series for available social information, rather than a static social network. There is no reason in principle that ordinary NBDA models could not make similar use of these data. Another contrast is with our emphasis on modeling the entire behavioural sequence. Again, in principle, NBDA can achieve the same goal. But one common NBDA approach is time of acquisition diffusion analysis, or TADA. TADA looks at only the first observed instance of an focal individual performing a behaviour. It then evaluates whether that focal's behaviour is predicted by the first instance of social information they observed or is

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

rather due to innovation (80). In contrast, we attempt to model the entire diffusion time series, from initial innovations to adoption and retention. We attempt this difficult task, because while social learning may not be responsible for the initial instance of an individual performing a behaviour, it may affect the probability of retaining it in the future. As a thought exercise, imagine buying a pair of turquoise trousers before the first day of the school year without ever observing peers wear them before. You then notice a particularly hip group of classmates (or colleagues) wearing the same turquoise trousers. If the probability of you wearing turquoise trousers in the future increases because of this observation, then the initial decision to wear turquoise trousers was made individually. However, the decision to wear them on future occurrences is influenced by social information. In this case, TADA would not find evidence of social learning, while our framework could.

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 (71). 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. Technological improvements like these are essential for allowing us to more satisfactorily confront theory with data.

5.4. **Learning, exploration, and canalization.** Our model recovers the observation that all adult individuals at one point tried the canine seam technique, even if they have settled on another primary technique. Many adult and sub-adult individuals born after 2009 (Fig. 2 and S3; individuals: HE, JU, MI, EI, BW, LN), tried the canine seam technique at intermediate timesteps and switched to another processing technique that was successful for them. Individuals who had considerable success with the canine seam technique, made it their primary technique, but also sporadically tried other behaviours (individuals: NP, ME, QJ, LT, RK, MX, EV, YJ). This result is consistent with the possibility that social learning is guiding the strategy space they are willing to explore, but individual reinforcement and experience may dictate the technique(s) they settle on. This is consistent with our observation that older, naive individuals were less likely to observe others after settling on a predominant processing technique (Figure 2 and S4). The interplay between social learning in exploration and individual learning in canalization is an area that warrants further exploration.

5.5. **Implications for the origins and maintenance of traditions.** This model suggests that payoff-biased learning can cause the spread of a tradition. However, as (61) point out, social learning may increase within-group homogeneity, while individual learning may act to decrease it. Our

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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.

5.6. **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 (older individuals appear to have changed their behaviour after observing younger, more skilled individuals), it is a possibility in general. Such equifinality 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 apply social learning to applied contexts such as wildlife conservation (81; 82), and applying cultural evolutionary models to humans in areas such as political policy and sustainability. 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. This will enable us to better understand the cognitive, social, and ecological underpinnings of cultural transmission across all taxa.

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

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

behaviour, such as avoidance of a potentially dangerous demonstrator. For example, developing male orangutans have reduced opportunities to observe female demonstrators since females actively avoid males nearing reproductive maturity (21; 83). 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. 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 (84), 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 the study, collected field data, carried out statistical analysis, and drafted the manuscript; RM participated in statistical analysis and helped draft manuscript; SP established the field site, assisted with field logistics, collected data, and helped draft the manuscript. All authors gave final approval for publication.

7. Acknowledgments

We would like to thank M. Grote, M. Crofoot,and the UC Davis CE/HBE lab group who provided useful feedback on early versions of this manuscript. Thanks to A. Cobden, B. Davis, E. Seabright, M. White, M. Ziegler, C. Angyal, for assistance with data collection and transcription. Thanks to W. Lammers, J.C Ordoñez Jiménez, I. Godoy, and K. Kajokaite, for assistance with field logistics. A. and E. van Halen provided moral support to RM and BB while drafting this manuscript. We thank the Costa Rican park service (MINAET and SINAC, currently), Hacienda Pelon de la Bajura, Hacienda Brin D'Amor, and the residents of San Ramon de Bagaces for permission to work on their land. This research was performed in compliance with the laws of Costa Rica, and the protocol was approved by the UCLA animal care committee (ARC #1996-122 and 2005-084 plus various renewals and followed UC Davis animal care and use protocols (IAUCUC permit no. 17297).

8. Research Funds

Funding to BB was provided by grants awarded by the American Society of Primatologists and an ARCS Foundation, Northern California Chapter award. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. 1650042 awarded to BB. This project is based on work supported by the funding provided to SP by the Max Planck Institute for Evolutionary Anthropology, the National Science Foundation (grants No. SBR-0613226 and BCS-0848360),

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

two grants from the L.S.B. Leakey Foundation, two grants from the National Geographic Society, and multiple COR grants from the University of California-Los Angeles. 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. The Biology of Traditions: Models and Evidence. Cambridge University Press; 2003.

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

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

[4] Romanes GJ. Mental evolution in animals. D. Appleton; 1884.

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

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

[7] 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).

[8] 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.

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

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

[11] Box HO, Gibson KR, editors. Mammalian Social Learning: Comparative and Ecological Perspectives. Cambridge; New York: Cambridge University Press; 1999.

[12] Janik VM, Slater PJ. Traditions in mammalian and avian vocal communication. In: Fragaszy DM, Perry S, editors. The biology of traditions: models and evidence; 2003. .

[13] Brown C, Laland KN. Social learning in fishes: a review. Fish and Fisheries. 2003 Sep;4(3):280–288.

[14] Leadbeater E, Chittka L. Social Learning in Insects — From Miniature Brains to Consensus Building. Current Biology. 2007 Aug;17(16).

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

[16] Laland KN. Social learning strategies. Animal Learning & Behavior. 2004 Feb;32(1).

[17] 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).

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[18] 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.

[19] 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.

[20] Brainard MS, Doupe AJ. What songbirds teach us about learning. Nature. 2002;417(6886):351–358.

[21] 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. .

[22] 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.

[23] 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.

[24] 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.

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

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

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

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

[29] 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.

[30] 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.

[31] Oppenheimer JR. Behavior and Ecology of the White-faced Monkey: Cebus Capucinus, on Barro Colorado Island, C.Z. University of Illinois.; 1968.

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

[33] Jerison H. Evolution of the Brain and Intelligence. 1973;.

[34] Roth G, Dicke U. Evolution of the brain and intelligence. Trends in cognitive sciences. 2005;9(5):250–257.

[35] Rilling JK, Insel TR. The primate neocortex in comparative perspective using magnetic resonance imaging. Journal of Human Evolution. 1999 Aug;37(2).

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[36] Chapman CA, Fedigan LM. Dietary differences between neighboring Cebus capucinus groups: local traditions, food availability or responses to food profitability? Folia Primatologica. 1990;54(3-4).

[37] 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).

[38] 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).

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

[40] 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.

[41] Reader SM, Laland KN. Social intelligence, innovation, and enhanced brain size in primates. Proceedings of the National Academy of Sciences. 2002;99(7).

[42] Reader SM, Hager Y, Laland KN. The evolution of primate general and cultural intelligence. Philosophical Transactions of the Royal Society B: Biological Sciences. 2011;366(1567).

[43] 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.

[44] 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).

[45] Perry S, Panger M, Rose LM, Baker M, Gros-Louis J, Jack K, et al. Traditions in wild white-faced capuchin monkeys. In: Fragaszy DM, Perry SE, editors. The biology of traditions: Models and evidence; 2003. .

[46] Perry S, Ordoñez Jimenez JC. The efiects of food size, rarity, and processing complexity on white-faced capuchins' visual attention to foraging conspecifics. In: Feeding Ecology in Apes and Other Primates. vol. 48. Cambridge University Press; 2006. p. 203.

[47] 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.

[48] 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).

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

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

[51] 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.

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

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[53] 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.

[54] Henrich J, McElreath R. The evolution of cultural evolution. Evolutionary Anthropology: Issues, News, and Reviews. 2003;12(3):123–135.

[55] 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.

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

[57] 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.

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

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

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

[61] 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.

[62] Aplin LM, Sheldon BC, Morand-Ferron J. Milk bottles revisited: social learning and individual variation in the blue tit, Cyanistes caeruleus. Animal Behaviour. 2013 Jun;85(6):1225–1232.

[63] Chou LS, Marsh RE, Richerson PJ. Constraints on social transmission of food selection by roof rats, Rattus rattus. Acta Zoologica Taiwanica. 2000;11(2):95–109.

[64] 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).

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

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

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

[68] Vale G, Davis S, van de Waal E, Schapiro S, Lambeth S, Whiten A. Lack of conformity to new local dietary preferences in migrating captive chimpanzees. Animal Behaviour. 2017 2;124:135–144.

[69] 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.

[70] van de Waal E, Claidière N, Whiten A. Social learning and spread of alternative means of opening an artificial fruit in four groups of vervet monkeys. Animal Behaviour. 2013 Jan;85(1):71–76.

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

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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

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

[74] McElreath R. rethinking: Statistical Rethinking book package; 2016. V. 1.59.

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

[76] 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`.

[77] Acerbi A, van Leeuwen EJC, Haun DBM, Tennie C. Conformity cannot be identified based on population-level signatures. Scientific Reports. 2016 Oct;6:36068.

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

[79] 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.

[80] Hoppitt W, Boogert NJ, Laland KN. Detecting social transmission in networks. Journal of Theoretical Biology. 2010 Apr;263(4):544–555.

[81] Custance D, Whiten A, Fredman T. Social learning of an artificial fruit task in capuchin monkeys (Cebus apella). Journal of Comparative Psychology. 1999;113(1).

[82] Whitehead H, Rendell L, Osborne RW, Würsig B. Culture and conservation of non-humans with reference to whales and dolphins: review and new directions. Biological Conservation. 2004 Dec;120(3):427–437.

[83] Russon AE. Developmental perspectives on great ape traditions. The biology of traditions: models and evidence Cambridge University Press, Cambridge. 2003;.

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

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplemental Materials

Additional Acknowledgements: For long term data collected at Lomas, we are grateful to H. Gilkenson, W. Lammers, C. Dillis, M. Corrales, and R. Popa, for helping manage the field site. We thank The following field assistants assisted with data collection, each contributing a year or more of data to the Lomas Barbudal Monkey Project data set: L. Beaudrot, M. Bergstrom, R. Berl, A. Bjorkman, L. Blankenship, T. Borcuch, J. Broesch, J. Butler, F. Campos, C. Carlson, S. Caro, M. Corrales, N. Donati, C. Dillis, R. Dower, K. Feilen, K. Fisher, A. Fuentes J., M. Fuentes A., C. Gault, H. Gilkenson, I. Gottlieb, L. Hack, S. Herbert, C. Hirsch, C. Holman, S. Hyde, L. Johnson, W. Lammers, S. Lee, S. Leinwand, K. Kajokaite, M. Kay, E. Kennedy, D. Kerhoas-Essens, E. Johnson, S. Kessler, S. MacCarter, J. Manson, W. Meno, C. Mitchell, Y. Namba, A. Neyer, C. O'Connell, J.C.Ordoñez J., R. Popa, K. Ratliff , H. Ruffler, S. Sanford, M. Saul, I. Schamberg, C. Schmitt, A. Scott, J. Verge, A. Walker-Bolton, and E. Williams.

TABLE S1. Widely applicable information criteria (WAIC) estimates, WAIC model weights (wWAIC), posterior median (PME), posterior standard deviation (SD), and standard deviation of varying effects for individuals (σ) for all evaluated models.

Model	Global		Freq. Dep.		Payoff Bias		Age Bias		Cohort Bias		Rank Bias		Kin Bias		Individual		Null
WAIC	4005.12		4010.77		4021.88		4031.79		4053.04		4054.69		4060.97		4071.97		5526.40
wWAIC	0.94		0.06		<0.01		<0.01		<0.01		<0.01		<0.01		<0.01		<0.01
	PME	SD	PME	SD	PME	SD	PME	SD	PME	SD	PME	SD	PME	SD	PME	SD	
λ	20.97	*1.11*	21.03	*1.08*	23.22	*1.41*	22.58	*1.36*	22.42	*1.27*	22.44	*1.31*	22.33	*1.32*	19.51	*0.96*	
ϕ	0.15	*0.03*	0.16	*0.03*	0.15	*0.03*	0.15	*0.03*	0.15	*0.03*	0.15	*0.03*	0.15	*0.03*	0.13	*0.03*	
γ	0.14	*0.03*	0.14	*0.03*	0.15	*0.05*	0.10	*0.04*	0.09	*0.04*	0.09	*0.03*	0.08	*0.03*			
f_c	0.38	*0.28*	0.53	*0.12*													
β_{pay}	1.02	*0.84*			3.49	*0.89*											
β_{kin}	0.19	*0.93*											0.35	*0.99*			
β_{rank}	-0.11	*0.91*									1.12	*0.96*					
β_{coho}	0.48	*0.93*							1.07	*1.03*							
β_{age}	0.69	*0.92*					2.11	0.83									
μ_{age} (ϕ)	-0.11	*0.03*	-0.11	*0.03*	-0.09	*0.03*	-0.10	*0.03*	-0.09	*0.03*	-0.09	*0.03*	-0.09	*0.03*	-0.08	*0.03*	
μ_{age} (γ)	-0.10	*0.05*	-0.09	*0.05*	-0.01	*0.05*	0.02	*0.06*	0.02	*0.06*	0.03	*0.05*	0.03	*0.05*			
σ_ϕ	0.66	*0.03*	0.67	*0.03*	0.66	*0.03*	0.66	*0.03*	0.67	*0.03*	0.67	*0.03*	0.67	*0.03*	0.67	*0.04*	
σ_γ	0.69	*0.06*	0.68	*0.06*	0.72	*0.07*	0.74	*0.07*	0.69	*0.09*	0.67	*0.09*	0.62	*0.08*			
σ_{f_c}	1.29	*1.19*	1.37	*1.18*													
$\sigma_{\beta_{pay}}$	0.28	*0.42*			0.71	*0.84*											
$\sigma_{\beta_{kin}}$	0.25	*0.40*											0.22	*0.35*			
$\sigma_{\beta_{rank}}$	0.26	*0.42*									0.29	*0.51*					
$\sigma_{\beta_{coho}}$	0.26	*0.40*							0.52	*0.83*							
$\sigma_{\beta_{age}}$	0.25	*0.38*					0.84	*0.70*									

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

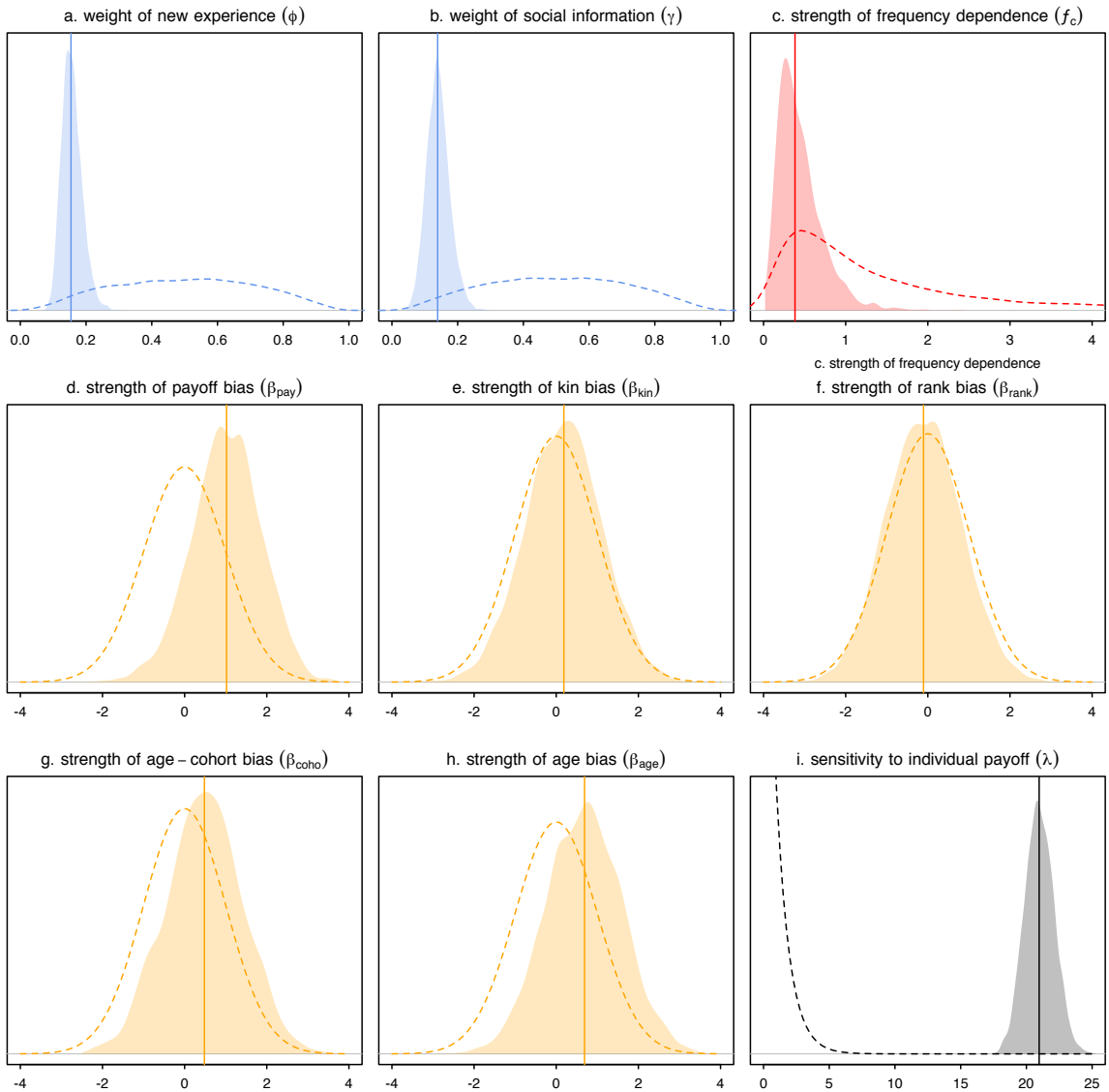


FIGURE S1. Posterior distributions of the main effects of the estimated parameters for the global model. Vertical solid lines lie at posterior mean. Dashed curves are the prior probabilities used to estimate each parameter.

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

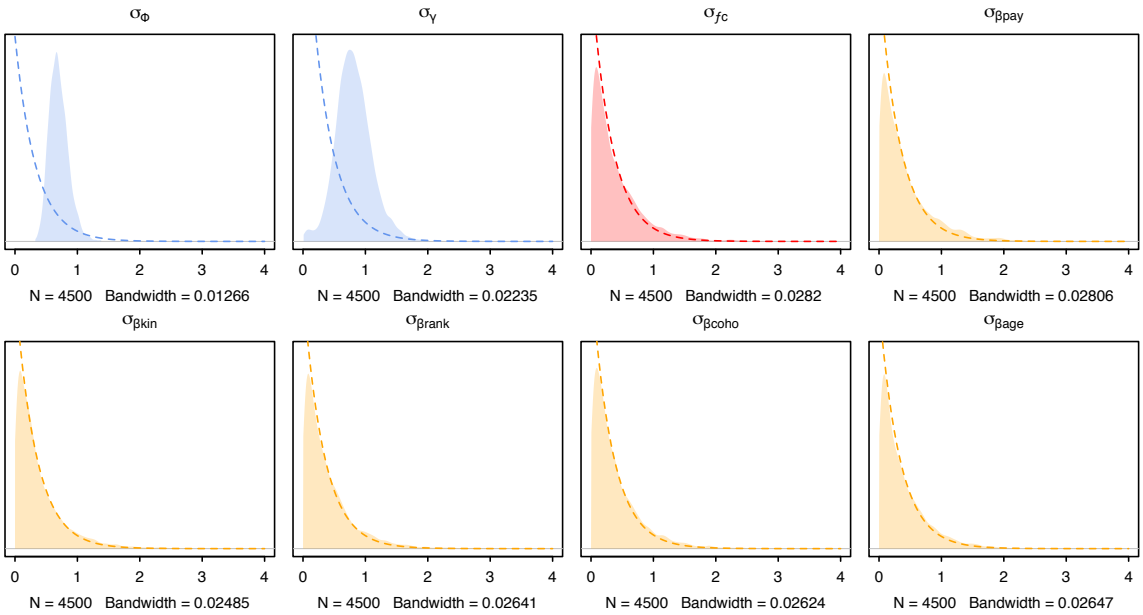


FIGURE S2. Posterior distributions of estimates of σ, the standard deviation of varying effects across all individuals. Vertical solid lines lie at posterior mean. Dashed curves are the prior probabilities used to estimate each parameter.

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

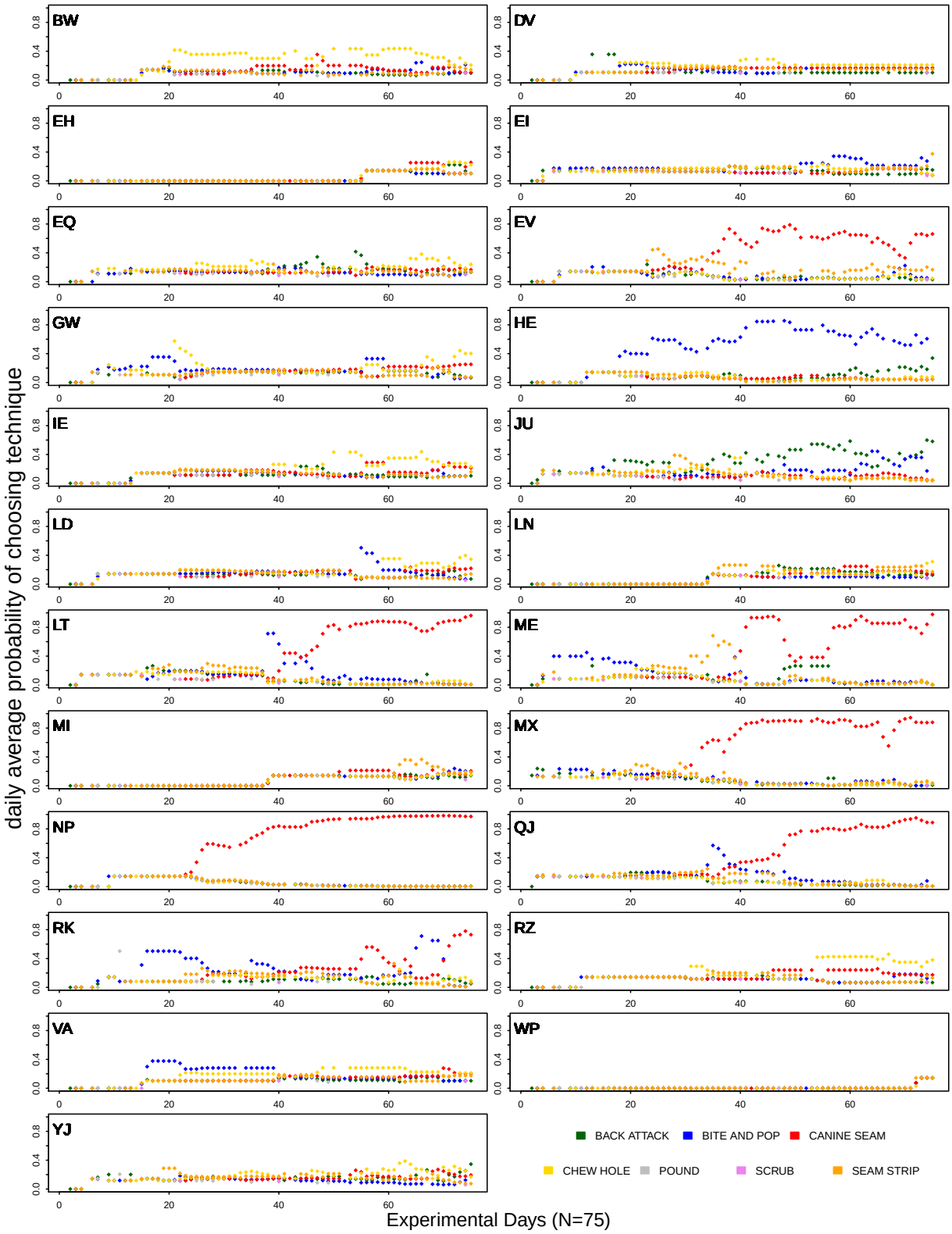


FIGURE S3. Daily average mean probability of choosing processing technique for each individual. ID codes are in upper left hand corner.

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

TABLE S2. Posterior median estimates (PME) and standard deviations (SD) for global models evaluating social from temporal windows of the previous 7, 21, and 28 days. Predictions are comparable to one another and to those presented in the main paper with a width of the previous 14 days.

Social Info Window Width	7 days		21 days		28 days	
Parameter	PME	SD	PME	SD	PME	SD
λ	20.98	*1.10*	20.81	*1.10*	20.95	*1.13*
ϕ	0.16	*0.03*	0.15	*0.03*	0.15	*0.03*
γ	0.14	*0.03*	0.16	*0.04*	0.14	*0.04*
f_c	0.39	*0.28*	0.35	*0.21*	0.39	*0.28*
β_{pay}	1.03	*0.86*	1.16	*0.73*	1.04	*0.87*
β_{kin}	0.17	*0.96*	0.05	*0.93*	0.19	*0.95*
β_{rank}	-0.10	*0.93*	0.08	*0.90*	-0.10	*0.92*
β_{coho}	0.47	*0.95*	0.32	*0.91*	0.48	*0.93*
β_{age}	0.68	*0.88*	0.87	*0.87*	0.67	*0.90*
μ_{age} (ϕ)	-0.11	*0.03*	-0.11	*0.03*	-0.11	*0.03*
μ_{age} (γ)	-0.10	*0.05*	-0.12	*0.05*	-0.10	*0.05*
σ_ϕ	0.67	*0.03*	0.67	*0.03*	0.66	*0.03*
σ_γ	0.69	*0.06*	0.70	*0.06*	0.69	*0.06*
σ_{f_c}	1.28	*1.17*	1.29	*0.77*	1.28	*0.96*
$\sigma_{\beta_{pay}}$	0.28	*0.43*	0.38	*0.54*	0.28	*0.42*
$\sigma_{\beta_{kin}}$	0.24	*0.39*	0.26	*0.45*	0.25	*0.37*
$\sigma_{\beta_{rank}}$	0.26	*0.42*	0.26	*0.41*	0.26	*0.43*
$\sigma_{\beta_{coho}}$	0.25	*0.40*	0.25	*0.40*	0.26	*0.38*
$\sigma_{\beta_{age}}$	0.27	*0.38*	0.24	*0.36*	0.25	*0.38*

bioRxiv preprint doi: https://doi.org/10.1101/110221; this version posted February 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

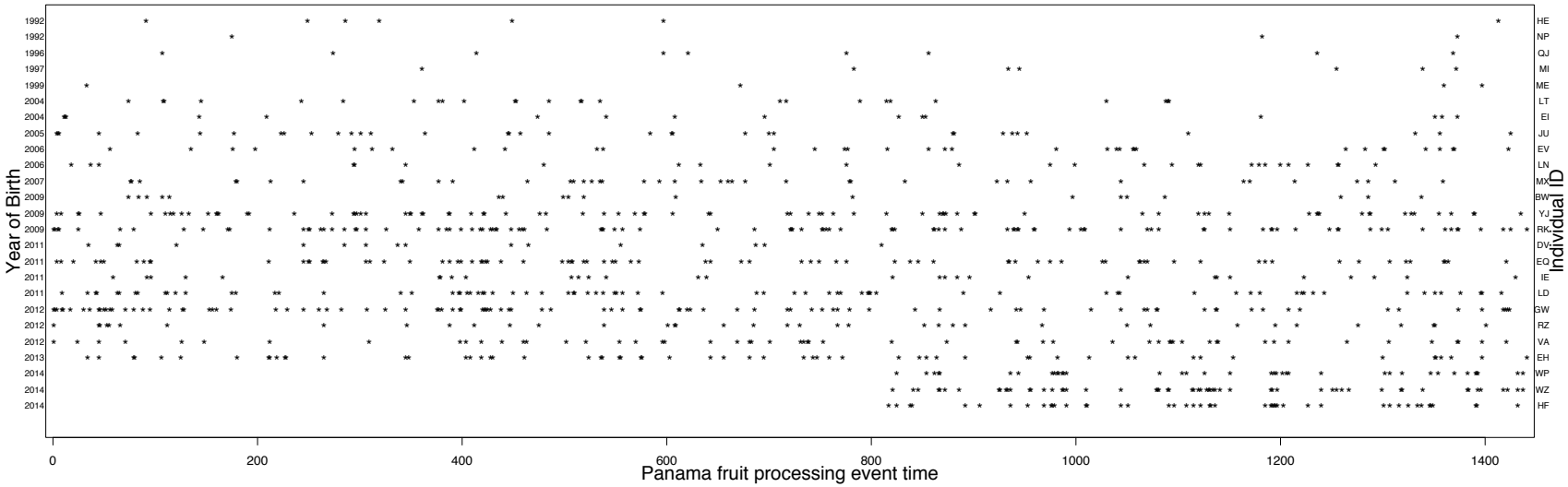


FIGURE S4. Plot of the use incidences of focused, sustained observation of another panamá foraging capuchin. Older individuals are at the top of the graph. The x-axis is the event time of each fruit processed by a single individual. Darker areas indicate more observation of conspecifics. Each star indicated an observation event by the individual on the y-axis.