PROCEEDINGS B

rspb.royalsocietypublishing.org

Research

Cite this article: Arbilly M, Laland KN. 2017 The magnitude of innovation and its evolution in social animals. *Proc. R. Soc. B* **284**: 20162385.
http://dx.doi.org/10.1098/rspb.2016.2385

Received: 1 November 2016
Accepted: 16 January 2017

Subject Category:
Behaviour

Subject Areas:
behaviour, cognition, evolution

Keywords:
innovation, creativity, cognition, social learning

Author for correspondence:
Michal Arbilly
e-mail: michalarbilly@gmail.com

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.3677083.

THE ROYAL SOCIETY PUBLISHING

The magnitude of innovation and its evolution in social animals

Michal Arbilly[1] and Kevin N. Laland[2]

[1]Department of Biology, Emory University, Atlanta, GA 30322, USA
[2]School of Biology, University of St Andrews, Harold Mitchell Building, St Andrews, Fife KY16 9TH, UK

MA, 0000-0002-3227-882X; KNL, 0000-0002-2457-0900

Innovative behaviour in animals, ranging from invertebrates to humans, is increasingly recognized as an important topic for investigation by behavioural researchers. However, what constitutes an innovation remains controversial, and difficult to quantify. Drawing on a broad definition whereby any behaviour with a new component to it is an innovation, we propose a quantitative measure, which we call *the magnitude of innovation*, to describe the extent to which an innovative behaviour is novel. This allows us to distinguish between innovations that are a slight change to existing behaviours (low magnitude), and innovations that are substantially different (high magnitude). Using mathematical modelling and evolutionary computer simulations, we explored how aspects of social interaction, cognition and natural selection affect the frequency and magnitude of innovation. We show that high-magnitude innovations are likely to arise regularly even if the frequency of innovation is low, as long as this frequency is relatively constant, and that the selectivity of social learning and the existence of social rewards, such as prestige and royalties, are crucial for innovative behaviour to evolve. We suggest that consideration of the magnitude of innovation may prove a useful tool in the study of the evolution of cognition and of culture.

1. Introduction

Many animals are now known to invent new behaviours or to devise novel solutions to challenging problems [1,2], and such behaviour is now commonly referred to as 'innovation'. While innovation is widespread in animals, innovativeness varies across species, with high rates of innovation associated with both advanced cognition and increased brain size [3–6]. Innovation is also widely recognized as a crucial element in the evolution of complex culture, central to the ecological and demographic success of our species [7–10].

In spite of this attention, what constitutes an innovation remains a point of controversy [2]. Reader & Laland [1, p. 14] deployed a broad definition in describing innovation as 'the introduction of a new or modified learned behaviour not previously found in the population'. This definition makes clear that a primary characteristic of an innovation is its novelty, relative to the baseline of the population's existing behaviour. The logic behind deploying a broad definition, given the relatively poor understanding of animal innovation in a young science, is that it encourages data collection, potentially allowing for more refined definitions in the future, when a deeper understanding of the phenomenon has accrued [1]. Nonetheless, the use of this definition has stirred some debate [11,12], which is still ongoing [13–15]. Part of the criticism stresses that innovation should be a unique process, different from recognized processes related to other behavioural tasks [11].

Consider three putative cases of innovation: a bird foraging at a new food patch; another bird feeding on a novel food, and a third bird devising a novel foraging technique comprising carefully fashioning a stick with which to extract a worm out of a crevice in a tree. All three cases can be regarded as innovation, but we nonetheless recognize that there are also non-trivial differences between these cases that merit recognition. The three avian innovations do seem to be qualitatively different in their nature, perhaps making different cognitive demands on the animal, although pinning down such differences is a non-trivial task. Herein lies a problem that bedevils innovation research. Much disagreement over the definition of innovation stems from the difficulty of including the ordinary and extraordinary under the same umbrella term. Here we suggest that different forms of innovation can be distinguished from each other by a quantitative measure, which we label 'magnitude', and which specifies how different the innovation is from established behaviours that already exist in the population's repertoire. On this perspective, foraging in a new patch, which is rather similar to existing behaviour, is an innovation of low magnitude, while designing a task-specific tool is a high-magnitude innovation. While in practice it is likely to remain challenging to specify the magnitude of a particular real-world innovation in absolute terms, here we show that the concept of magnitude of innovation can nonetheless be of heuristic value to the scientific community, both by facilitating theoretical investigations of the evolution of innovation, and by allowing qualitative predictions to be made and tested.

An advantage of the magnitude of innovation concept is that it can easily be described quantitatively, as a deviation from the population's mean behaviour. Theoretically, this deviation can be specified over different and arbitrary dimensions, potentially solving the problem of attributing innovation to ecological conditions versus a higher-level cognitive process [11]. An innovation that is a response to a change in environmental conditions can be defined as a deviation over an ecological dimension, relative to the behaviour exhibited by the population in response to representative or modal conditions. In contrast, an innovation that involves, for example, extrapolating from an existing behaviour to a new behaviour or to the application of the behaviour in a new context, is occurring over a cognitive dimension, which, at least in principle, can be detailed. Of course, in practice a given innovation may encompass deviation over more than one dimension. Describing innovations by their magnitude has the advantage that it allows the researcher to approach one of the most interesting questions about innovations: why do some individuals, and some species, produce innovations that appear more advanced (or, at least, more novel) than others?

Many factors are thought to contribute to the occurrence of innovative behaviour, or the lack thereof, at the intra- and interspecies level. These range from a number of physiological, cognitive and emotion-related factors, including in particular energetic state [16], neophobia and exploratory behaviour [17–19], social rank [17,20], and motivation [21]. Of course, many of these factors will operate simultaneously, and will interact. Rank is an especially complicated issue. Many studies show a greater propensity to innovate among subordinate individuals [16,20,22–24], supporting the notion that necessity drives animals to innovate [1]. However, other studies suggest innovation to be more prevalent among dominants [17,25,26]. In addition, social rank itself might be affected by innovative behaviour, if innovation allows individuals to climb the dominance hierarchy [27] or, most obviously in the case of humans, if innovators receive some social reward for being copied. Such reward, in the form of prestige or royalties, can be a simple consequence of social learners' efforts to gain access to resources through the innovator [28,29].

Here, we use agent-based computer simulations to explore the evolution of innovation, characterized by its frequency and magnitude, and how it might be shaped by group-living. We investigate the effects of different social rewards and of the quality and accuracy of social learning, as well as factors not unique to social life, such as the strategic use of learned behaviours, and the strength of natural selection. We discuss the significance of considering the magnitude of innovation in the light of these results.

2. The model

We evolve a population of social agents. In each generation, individuals acquire new behaviours (of varying payoff) through individual innovation and the copying of others, based on probabilities determined by their genes. They apply these behaviours, at some degree of selectivity, to receive the payoff associated with each behaviour. They may receive additional payoff if their innovation had been copied, or pay a cost for copying another's behaviour. The agents then reproduce proportionally to the payoff they have accumulated through their lifetime.

(a) The population

We modelled a population of $n = 100$ individuals, with each individual characterized by two focal genes, L and I. A learning gene, L, determined the average proportion of time or in other words, the probability, the individual allocated to individual learning (innovating) and social learning. There were 11 possible alleles in this gene: 0, 0.1, 0.2 . . . 1, where 0 coded for full-time social learning, 1 for full-time individual learning, and all other alleles for a combination of the two (e.g. a carrier of the 0.3 allele spent 30% of the time, on average, learning socially, complemented by an average of 70% learning individually). An innovation magnitude gene, I, determined how far from the population's norm an individual's innovations will be when learning individually. There were again 11 possible alleles in this gene: 0, 0.1, 0.2 . . . , 1, which represented standard deviations from the population's mean behaviour. The value of a behavioural innovation, produced based on the mean and standard deviation (as detailed below), is in fact the payoff an individual receives for learning/applying the behaviour, which later translates into the individual's fitness (see below). The population's mean value of behaviour does not change over a generation's lifetime, thus the magnitude of an innovation is measured only compared with behaviours in existence when the generation is born. A cumulative culture situation, where the mean changes as new innovations emerge, is outside the scope of this paper and will be considered in future analysis.

(b) Learning phase

All individuals in the population had a limited number of learning steps T (set to either $T = 10$ or $T = 100$), in which they acquired their behavioural repertoire. The two cases are designed to represent few and many opportunities for learning new behaviours within a lifespan. At each step, each individual learned either individually or socially; the probability of either learning strategy was dictated by the individual's L genotype. Individuals who learned individually generated an innovation: a new behaviour. The value of this innovative behaviour (i.e. its payoff) was drawn from a normal distribution whose mean was the population's mean behaviour and whose standard deviation was the innovator's allele in the I gene; for convenience, the population's mean behaviour value was set to 0. Then, individuals who learn socially in this learning step, copied the behaviours generated by innovators. All innovations were ranked according to their value, and which innovations would be copied depended on the selectivity of social learning in the population (which was kept constant per population). We controlled the selectivity of social learning using the variable D, defined as 1 – [the fraction of demonstrators copied]. When selectivity was high (high D) only innovations with the highest value were copied (e.g. when $D = 0.9$ only the top 10% of innovation was copied); as the selectivity of social learning became lower, copying became more random (and was completely random at $D = 0$). If there were no individual learners in a specific learning step, for computational reasons (distinguishing copiers from innovators with innovation magnitude of 0 in analysing the data), individuals were assigned with a behaviour of value (−1). Because if there was even one innovation at a specific time step, it was copied by all individuals in the population, and because reproductive success was always relative within a generation (see below), this should not affect the results.

We define three possible effects of copying on the payoff to innovators and copiers: social learning penalty, prestige, and royalties. The *social learning penalty* condition represents the possibility of innovators having an advantage in exploiting the behaviour they produced, or alternatively, the possibility that social learners have some difficulty copying the innovation faithfully, for example, owing to its complexity. The social learning penalty is represented by a deduction of some percentage φ from the value of the socially learned behaviour. To the copier, the reduced value is the true value of the behaviour: the reduced value is the payoff the copier receives for the behaviour when learning it and when applying it, and it is also the value by which the copier 'prioritizes' the behaviour compared with other behaviours in its repertoire.

The *prestige* condition accounts for an additional payoff to innovators whose innovation is copied. This addition is proportional, at a population-wide rate R, to the original value of the innovation (e.g. 1% of original value), and to the number of times it was copied. The prestige condition does not entail any direct cost to copiers ($\varphi = 0$). The *royalties* condition combines the social learning penalty condition with the prestige condition, and the bonus innovators receive is equal to the penalty paid by social learners ($R = \varphi > 0$). It should be noted that under this condition, the social learning penalty deduction is not paid 'directly' to innovators, because this deduction from the copiers perspective consistently affects the value of the behaviour, through both the learning and the application phase (see below). The bonus to the innovators, on the other hand, is calculated based on value of the behaviour and the number of times the behaviour has been copied. This bonus is added to the total payoff innovators obtained throughout the learning and application phases.

While there may be different ways to model prestige, our choice to tie it to the value of the innovation is based on our assumption that a greater magnitude of innovation would be tied with greater benefits at the social level: in terms of protection, access to resources and mating opportunities (see further analysis of this point in the discussion).

(c) Application phase

We assume that after acquiring the behaviours, individuals apply these behaviours and will tend to use them with a frequency directly proportional to the payoff they offer. To calculate the proportion of time allotted to each behaviour, and because payoffs can be negative as well as positive, we use an exponential transformation of the form:

$$p_x = \frac{e^{\sigma\beta_x}}{\sum_{i=1}^{j} e^{\sigma\beta_i}}, \tag{2.1}$$

where p_x is the proportion of time spent using behaviour x, β_x is the payoff of behaviour x, $i = 1 \ldots j$ are the behaviours the individual has acquired during its learning phase, and σ is the application sensitivity: the degree to which the population can distinguish between payoffs in choosing which behaviours to apply. High sensitivity (high σ) means agents will spend a higher proportion of their time applying the highest-paying behaviour and little to no time applying low value behaviours; low sensitivity means less difference in time investment between high payoff and low payoff behaviours.

We then calculate the payoff accumulated from applying the learned behaviours, W_A, by summing up the multiplications of each behaviour's payoff and the proportion of time spent applying it:

$$W_A = \sum_{i=1}^{j} p_i\beta_i. \tag{2.2}$$

(d) Selection and reproduction

To calculate the total payoff to individuals in the population, W_T, we summed the payoff obtained both during the learning phase, W_L (which is the sum of all payoffs of behaviours learned), and during the application phase, W_A, using a weight factor $0 < \alpha < 1$ which allowed us to control the relative amount of time invested in each phase:

$$W_T = \alpha W_L + (1 - \alpha) W_A. \tag{2.3}$$

We choose to include payoff received for behaviours regardless of whether they are applied (in the form of W_L), to account for costs of time and energy during the innovation and social learning processes.

Individuals then reproduced, producing a number of offspring proportional to their total payoff relative to the payoff of all other individuals in the population. Because the total payoff could be negative, we again use an exponential transformation of the form:

$$r_y = \frac{e^{\lambda W_{T,y}}}{\sum_{k=1}^{n} e^{\lambda W_{T,k}}}, \tag{2.4}$$

where r_y is the probability of reproduction for individual y, and λ is the strength of selection. When λ is large, selection is strong: the individuals who obtained a higher total payoff had much higher chances to reproduce than individuals who obtained a lower payoff. As λ decreases, selection becomes weaker: the chances of reproduction of individuals with high payoffs and of those with low payoffs become similar. Among the offspring, we assumed a mutation rate of $\mu = 10^{-1}\ n$ in both genes. Mutation was random and the new variant was drawn from each gene's predefined allele pool.

3. Results

(a) Effect of the quality of social learning

Across conditions, we found both the rate and magnitude of innovation to be higher as the selectivity of social learning,

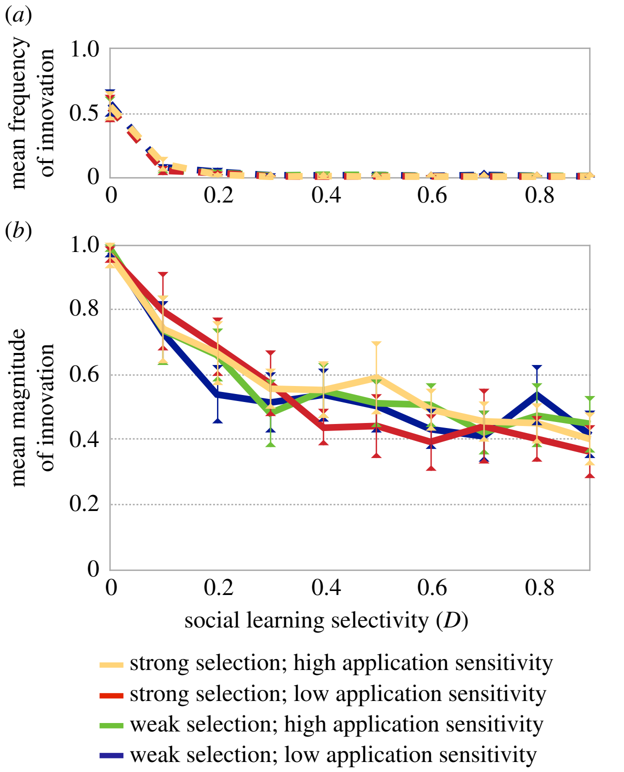

Figure 1. The effect of application sensitivity and selection strength on (*a*) innovation frequency, and (*b*) innovation magnitude. Strong selection: $\lambda = 3.3$; weak selection: $\lambda = 0.6$; high application sensitivity: $\sigma = 3.3$; low application sensitivity: $\sigma = 0.6$. Number of learning steps $T = 10$ (similar results were obtained for $T = 100$). Means and standard errors calculated for generations 5001–10 000, across 10 repeats of each simulation. Magnitude of innovation was calculated among individuals whose innovativeness probability was greater than zero ($L > 0$). Mutation rate $\mu = 10^{-1}$, learning phase weight $\alpha = 0.1$, social learning penalty $\varphi = 0$, prestige rate $R = 0$.

D, dropped. Because of this consistency and owing to the complementary (but not simple) relationship of innovation and social learning, we chose to examine the effects of other variables using *D* as a scale (figures 1–4). High selectivity of social learning (high *D*) means that social learners copy only the best (high value) behaviours. Because innovating involves a risk of producing a maladaptive (below average) behaviour, innovating in this condition involves only costs and no benefits: an innovator may by chance produce a high-value new behaviour, but this behaviour will be copied by many others. However, if an innovator produces a low-value behaviour, it cannot avoid its low value, while social learners can copy a better behaviour from another innovator. Therefore, social learners fare better than innovators, and innovating becomes maladaptive in itself. When the selectivity of social learning is low, social learners risk copying a maladaptive innovation, and at the same time, because more innovations are being copied and not only the best ones, innovators have the chance of their high-value innovation not being copied by others, or only by few, thus increasing the possible relative benefit. The combination of these two factors seems to increase both the frequency and magnitude of innovation when *D* is low.

(b) Strength of selection and application sensitivity

Strength of selection and application sensitivity each affect the frequency and magnitude of innovation only in interaction with other factors, and do not seem to make a difference in the basic model. An interaction between the two does not seem to have an effect either (figure 1). Both play a role in interaction with social learning penalty, prestige and royalties (see below): once innovation frequency is increased owing to other factors, they may become significant.

(c) Number of learning steps

The number of learning steps is the total number of opportunities, either to innovate or learn socially, that a population had during a generation's lifetime. Innovativeness evolved at higher frequencies when the number of learning steps was small than when the number of learning steps was large ($T = 10$ compared with $T = 100$; demonstrated in figure 2, qualitatively similar results were obtained under all conditions). This difference is due to the stochastic nature of a short learning window: given very little opportunities to innovate, individuals may, by chance, produce a high-value behaviour or a low-value behaviour, putting them at very high or very low chances to reproduce, respectively. However, given many chances to innovate, individuals will produce both high and low value innovations. These may balance each other out, and unless social learning is completely random ($D = 0$), individuals who are strictly social learners will manage to avoid learning the lowest value behaviours and on average accumulate a higher payoff than individuals who innovate. While the frequency of innovativeness and magnitude of innovation varied under the different conditions detailed below, the difference between large *T* and small *T* remained notable.

(d) Social learning penalty

A penalty to socially learned behaviours was found to increase the frequency and magnitude of innovation, but any such increases declined with the selectivity of social learning (figure 2). The increases in innovation frequency and magnitude were greater when selection was stronger compared with weaker (compare pink lines with red lines, or light blue lines with orange lines in figure 2). Application sensitivity interacts with social learning penalty in a very similar manner to strength of selection (and therefore not shown). Furthermore, a difference in strength of selection (or application sensitivity) had a larger effect the larger the social learning penalty was (i.e. difference between red and pink is greater than difference between light blue and orange in figure 2). Because the social learning penalty decreases the value of copied behaviours, pushing the value distribution of behaviours to the negative side, the advantage to innovation under these conditions is clear, especially when social learning selectivity is low and social learners copy not only the best behaviours (low *D*). Higher application sensitivity, that is, higher sensitivity to behaviours' value, increases the chances of applying high-value behaviours; while this is true under any condition, a greater penalty to social learning increases the difference between the distribution of behavioural values of copiers and innovators, therefore a greater penalty results in a greater effect when sensitivity is high. The same is true for strength of selection.

Innovation magnitude increased to its maximum as soon as the frequency of innovation became higher that 0.1 (figure 2).

(e) Social prestige

Social prestige for innovators generally increases both the frequency of innovativeness and the magnitude of innovation

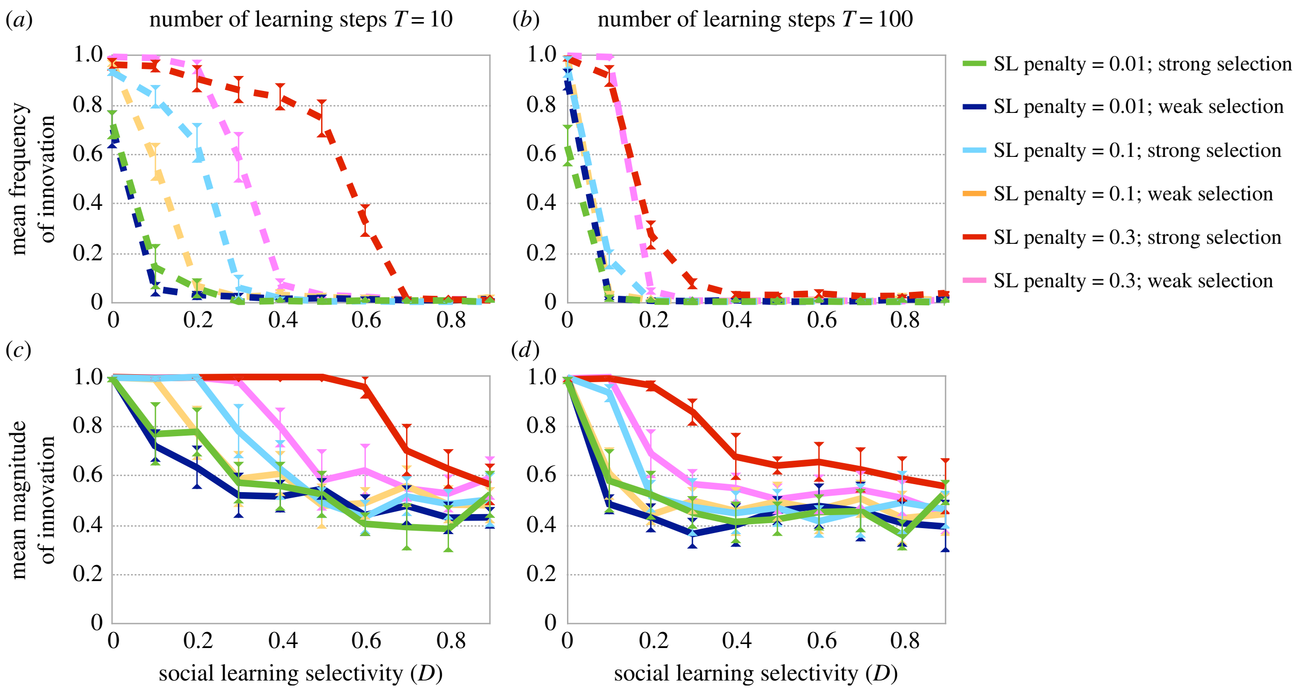

Figure 2. Effect of a social learning penalty φ on innovativeness. (*a*) Frequency of innovation, number of learning steps $T = 10$; (*b*) frequency of innovation, number of learning steps $T = 100$; (*c*) magnitude of innovation, number of learning steps $T = 10$; (*d*) magnitude of innovation, number of learning steps $T = 100$. Strong selection: $\lambda = 3.3$; weak selection: $\lambda = 0.6$. Means and standard errors calculated for generations 5001–10 000, across 10 repeats of each simulation. Magnitude of innovation was calculated among individuals whose innovativeness probability was greater than zero ($L > 0$). Mutation rate $\mu = 10^{-1}$, learning phase weight $\alpha = 0.1$, prestige rate $R = 0$, application sensitivity $\sigma = 1$. Similar results were obtained when manipulating application sensitivity.

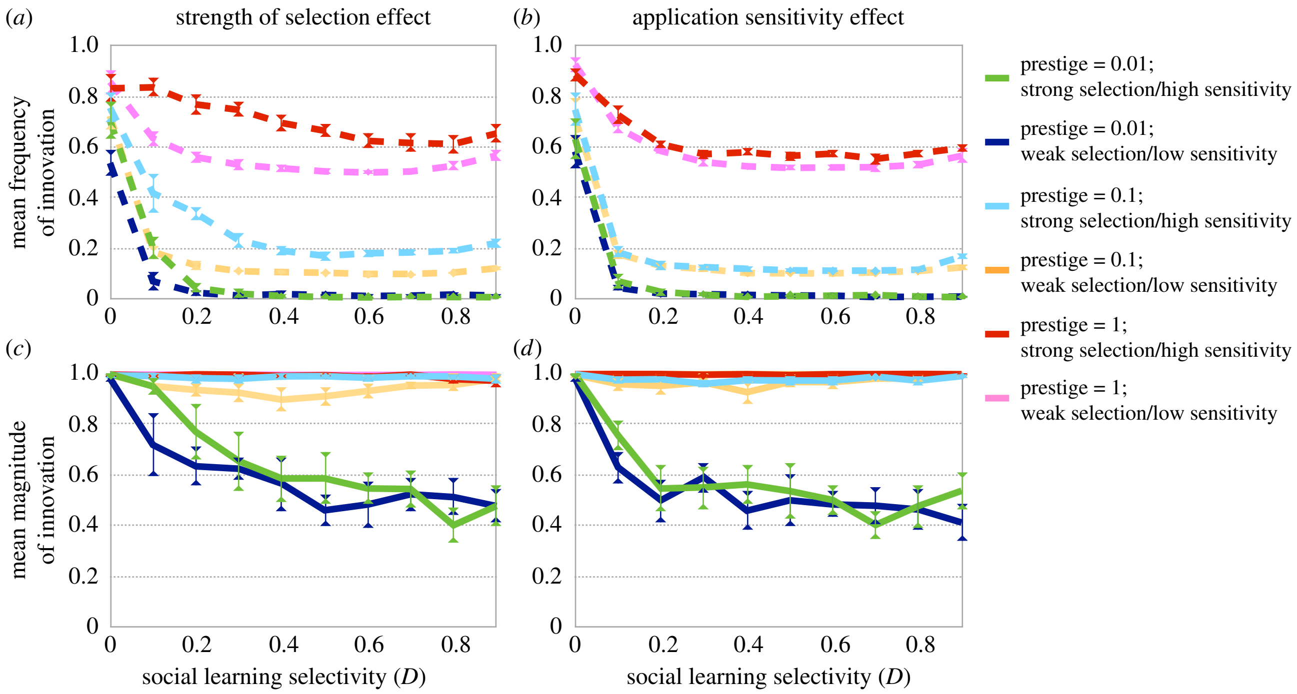

Figure 3. Effect of prestige on innovativeness. (*a*) Mean frequency of innovation under weak and strong selection; (*b*) mean frequency of innovation under low and high application sensitivity; (*c*) mean magnitude of innovation under weak and strong selection; (*d*) mean magnitude of innovation under low and high application sensitivity. Strong selection: $\lambda = 3.3$; weak selection: $\lambda = 0.6$. High application sensitivity: $\sigma = 3.3$; low application sensitivity: $\sigma = 0.6$. Means and standard errors calculated for generations 5001–10 000, across 10 repeats of each simulation. Magnitude of innovation was calculated among individuals whose innovativeness probability was greater than zero ($L > 0$). Mutation rate $\mu = 10^{-1}$, learning phase weight $\alpha = 0.1$, social learning penalty $\varphi = 0$, application sensitivity $\sigma = 1$ in (*a*,*b*), selection strength $\lambda = 1$ in (*c*,*d*).

(figure 3). As in the case of the social learning penalty, any mean innovativeness frequency that was 0.1 or higher, was associated with the highest possible innovation magnitude. Not surprisingly, social prestige did not interact with application sensitivity but only with the strength of selection—stronger selection resulting in higher frequency and magnitude of innovation (compare difference between pink and red lines in figure 3*a*, and in figure 3*b*). This is because

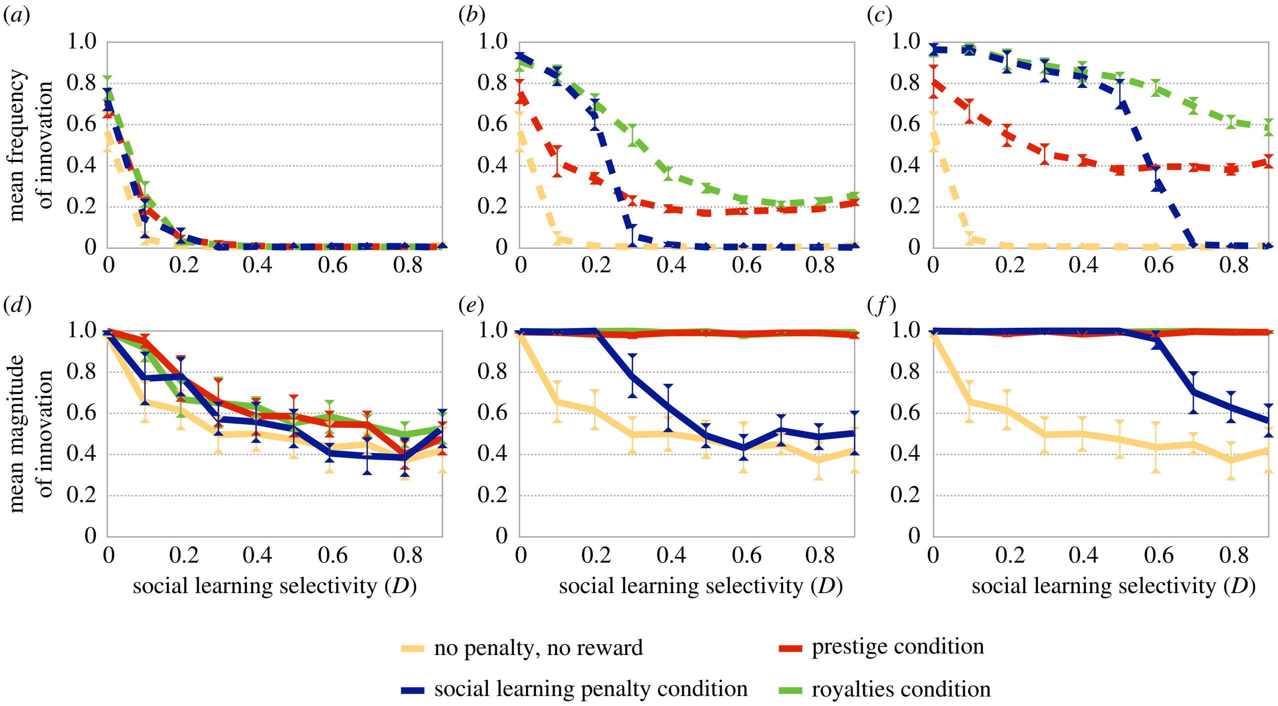

Figure 4. Comparison of the effects of social learning penalty (SLP) rate (φ), prestige rate (R) and royalties ($\varphi + R$) on innovation frequency (dashed lines) and magnitude (solid lines). (*a,d*) SLP condition: $\varphi = 0.01$, $R = 0$; prestige condition: $\varphi = 0$, $R = 0.01$; royalties condition: $\varphi = R = 0.01$. (*b,e*) SLP condition: $\varphi = 0.1$, $R = 0$; prestige condition: $\varphi = 0$, $R = 0.1$; royalties condition: $\varphi = R = 0.1$. (*c,f*) SLP condition: $\varphi = 0.3$, $R = 0$; prestige condition: $\varphi = 0$, $R = 0.3$; royalties condition: $\varphi = R = 0.3$. Means and standard errors were calculated for generations 5001–10 000, across 10 repeats of each simulation. Magnitude of innovation was calculated among individuals whose innovativeness probability was greater than zero ($L > 0$). Mutation rate $\mu = 10^{-1}$, learning phase weight $\alpha = 0.1$, application sensitivity $\sigma = 1$, selection strength $\lambda = 3.3$.

prestige adds to the total payoff of innovators but does not affect the value of the copied behaviour itself for either the innovator or the copier. As in the case of a social learning penalty, a higher prestige rate resulted in greater impact of the strength of selection.

(f) Social learning penalty, prestige and royalties

A comparison of the effects of a social learning penalty and social prestige shows that while a social learning penalty resulted in higher frequency and magnitude than prestige when the selectivity of social learning was low, prestige led to higher frequency and magnitude than a social learning penalty when social learning was good (compare blue and red lines in figure 4). Combining the effect of a social learning penalty and prestige to create the royalties condition showed a general trend of increase in the frequency and magnitude of innovation. Interestingly, while royalties generally had a greater effect than a social learning penalty or prestige, in parts of the parameter range the effect of royalties was no different than the effect of a social learning penalty alone (figure 4). This phenomenon was observed for a low selectivity of social learning (low D), and is likely due to the fact that prestige, either by itself or as part of the royalties condition, relies on the presence of social learners—if there are not enough of them in the population, there is little payoff to reap from being copied.

The royalties condition was also not a simple combination of the effect of a social learning penalty and prestige when the value of these two parameters was high. When social learning selectivity was high (high D), where a social learning penalty in itself had no effect on frequency but prestige in itself had a moderate effect (figure 4*c*), royalties resulted in a notably higher frequency of innovation than prestige in itself. It appears that although a social learning penalty in itself was not enough to disfavour copying and increase the frequency of innovation, in the presence of prestige in the royalties condition it did have a significant effect.

As in the cases of social learning penalty and prestige, in the presence of royalties any mean innovativeness frequency that was 0.1 or higher, was associated with the highest possible innovation magnitude.

4. Discussion

From a strictly logical perspective, in the company of social learners, whatever advantage a behavioural innovation offers its producer might quickly dissipate as others adopt the behaviour and enjoy its benefits, free of cost. This argument has been proven repeatedly in studies of social learning and the diffusion of innovations [9]. It potentially holds for all innovations, but is likely to be even more pronounced when innovations greatly differ from existing behaviours, and might explain why these sorts of innovations are relatively rare. The results of our model show a positive, but not linear, relationship between the frequency of innovation and its magnitude—while lower frequency of innovation was associated with lower magnitude, and higher frequency with higher magnitude, the frequency of innovation had to be extremely low for the magnitude to be below its highest possible value. In fact, a frequency of innovation that was greater than 0.1 was usually associated with maximal magnitude. This may seem to conflict with the suggestion that innovations that greatly differ from the norm should be rare owing to their cost in the presence of copiers. However, owing to the normal distribution we assume for the magnitude of innovation, a high magnitude does not mean that all innovations produced are extreme, but rather a greater spectrum of innovations, compared with a low magnitude; the majority of innovations are always close to the population's norm, and the magnitude only determines the width of the normal distribution around this norm. Altogether, these results suggest that as long as the population is innovating at a consistent rate, extreme innovations (i.e. innovations of large magnitude) are expected to arise regularly, at this rate or another, depending on the general frequency of innovation.

Most notably, and perhaps surprisingly, innovation did not evolve when the selectivity of social learning was high, that is, when individuals were able to identify the highest-paying behaviour at each learning step and to copy it exclusively. While some accounts find a correlation between the tendency to innovate and social learning abilities between [3] and within species [30–32], these usually refer to the occurrence of social learning and not to its selectivity (but see [33]). An example of a form of social learning that may be functionally equivalent to low selectivity is the tendency of many animals to copy high-ranking individuals. This rule may be a good rule of thumb under restricted circumstances (i.e. if dominants are older, experienced and/or successful individuals), but there is at least some evidence from chimpanzees that dominants themselves do not necessarily use the best behaviour in their repertoire, yet they are being copied by subordinates nonetheless [34]; this finding has been suggested to explain the low occurrence of traditions observed in chimpanzee populations compared with the rate of innovation in this species. Innovations in chimpanzees are often produced by juveniles/subordinates, and are attributed to their limited access to resources [35,36]. Our model suggests that poor discrimination in copying might also be responsible, at least to some extent, for the high rate of innovation. The same is true for difficulties in copying innovations, represented in the model by the social learning penalty. Taken together, these findings point to the challenge of establishing a tradition: it requires innovation alongside selective, faithful copying. Such conditions may arise relatively rarely in most social learning species.

As humans are both capable of high-fidelity social learning and impressive innovation, it is perhaps natural to expect these traits to go together, and facilitate each other. In fact, highly accurate and selective social learning may hinder innovation, by reducing the advantages that it brings to the innovator. That may explain why humans have evolved prestige, and devised institutions that confer royalties to the innovator: these may greatly enhance the likelihood of innovation being beneficial, and adaptive traditions being propagated. While prestige is usually regarded as a marker of an individual from whom it is good to copy [37,38], prestige can also result from producing a high-value innovation. It is easy to think of examples from our modern human society, where innovators of technology, science, art and sports receive great praise and enjoy a high social status. In the rest of the animal kingdom, prestige is not so easy to identify; indeed, it may not occur [29]. Zahavi [39] suggested social benefits to altruistic behaviour. For example, helpers at the nest may be advertising their quality and motivation through their seemingly altruistic behaviour, thus potentially gaining in social standing or reputation, in a manner that may later help them find a good mate and successfully pass on their genes [39]. Still, there is as of yet little compelling data to support this hypothesis. The idea that problem-solving, or intelligence, might be under sexual selection [40] is at least functionally consistent with the notion of prestige, and specifically with social prestige as a consequence of behavioural innovation. A few studies found a positive correlation between problem-solving and mating success [41–43], but it is difficult to show that improved mating success is directly owing to preference for problem-solving mates rather than these problem-solving individuals faring better and being in good condition, which makes them attractive mates and successful breeders. While there is some evidence for improved social status following innovative behaviour in macaques [44] and chimpanzees [45], much remains to be studied to fully gauge the extent of this phenomena in animals other than humans. Henrich & Gil-White [29] suggest that prestige is an adaptation that originally evolved through giving deference to successful models in order to gain access to the resources they may provide, and eventually became a way for social learners to evaluate the models [29]. We view prestige in our model as deference given for resource access, which is translated into greater reproductive success. It would be interesting, however, to examine in future analyses how this basic form of prestige might evolve into a cue of demonstrator quality.

The concept of royalties—where a copier pays the innovator for exploiting the innovation—is of essence to modern human life. The earliest documentation of patents and copyrights does not date very far in our history [46], but royalties do not necessarily require formal institutions to exist; trading in information for goods was likely a feature of human societies early on, and can be tied to our species' tendency to cooperate with non-kin [47]. Other animals do engage in simple forms of trade, perhaps the best-studied being the trade in grooming in exchange for dominant tolerance [48], infant handling [49] or mating [50]. However, in non-humans, there is little evidence for any trade in information that innovators would stand to gain from. While a behaviour that is easy to copy by observation does not provide the opportunity for the innovator to request a payment in return, more difficult to acquire behaviours probably require teaching or complex language to transmit, which are rare and absent, respectively, in other animals [51,52]. In humans, it is possible to envisage how a system of royalties might develop from the system of prestige, as competition over access to a demonstrator increases. An increase in payoff to innovators owing to prestige might be regarded as at the expense of copiers, if reproduction is based on relative success, in which case prestige might be considered a form of royalty.

Our model suggests that royalties and prestige result in identical magnitude of innovation, but their effect on the frequency of innovation is different: royalties typically lead to higher (much higher under some conditions) frequencies of innovation. While both forms of reward promote extreme innovation, populations with a royalties regime will see many more of them. This difference is crucial for a population's ability to adjust to changes in its environment—a high rate of innovation facilitate a quick response—but it is also significant in the context of cultural evolution, and the pace at which it progresses [15].

Finally, it should be noted that, while we believe it represents the prevalent case, our assumption of normal distribution, where high-magnitude innovations are relatively rare and associated with higher variation in payoff, may not apply to all situations. For example, in cases where environmental variation is very high, low magnitude innovations may sometimes entail a greater variation in outcome than those of high magnitude. Such an environment is also not likely to promote social learning [53], and therefore the evolution of innovation magnitude would be under a somewhat different set of selective pressures, offering further avenues for investigation.

In sum, we submit that characterizing innovation in terms of both frequency and magnitude offers practical advantages. These include rendering the concept accessible to quantitative analysis and circumventing some challenging grey areas that arise in discussing the phenomena.

Supporting materials. The Matlab code for the simulations used in this paper has been uploaded as part of the electronic supplementary material.

Authors' contributions. K.N.L. and M.A. designed the model and co-wrote the paper; M.A. wrote the simulations and analysed the results.

Competing interests. We have no competing interests.

Funding. The study was funded by The John Templeton Foundation.

Acknowledgements. We thank Laurel Fogarty, Daniel Weissman, and two anonymous reviewers for their helpful comments. M.A. wishes to thank Lucia Jacobs for hosting her at UC Berkeley while conducting part of this analysis.

References

1. Reader SM, Laland KN. 2003 *Animal innovation*. Oxford, UK: Oxford University Press.
2. Reader SM, Morand-Ferron J, Flynn E. 2016 Animal and human innovation: novel problems and novel solutions. *Phil. Trans. R. Soc. B* **371**, 20150182. (doi:10.1098/rstb.2015.0182)
3. Reader SM, Laland KN. 2002 Social intelligence, innovation, and enhanced brain size in primates. *Proc. Natl Acad. Sci. USA* **99**, 4436–4441. (doi:10.1073/pnas.062041299)
4. Lefebvre L, Reader SM, Sol D. 2004 Brains, innovations and evolution in birds and primates. *Brain Behav. Evol.* **63**, 233–246. (doi:10.1159/000076784)
5. Benson-Amram S, Dantzer B, Stricker G, Swanson EM, Holekamp KE. 2015 Brain size predicts problem-solving abilities in mammalian carnivores. *Proc. Natl Acad. Sci. USA* **113**, 2532–2537. (doi:10.1073/pnas.1505913113)
6. Navarrete A, Reader SM, Whalen A, Laland KN. 2016 The co-evolution of innovation and technical intelligence in primates. *Phil. Trans. R. Soc. B* **371**, 20150186. (doi:10.1098/rstb.2015.0186)
7. Cavalli-Sforza LL, Feldman MW. 1981 *Cultural transmission and evolution: a quantitative approach*. Princeton, NJ: Princeton University Press.
8. Basalla G. 1988 *The evolution of technology*. Cambridge, UK: Cambridge University Press.
9. Rogers EM. 2003 *Diffusion of innovations*, 5th edn. New York, NY: Free Press.
10. Henrich J. 2015 Culture and social behavior. *Curr. Opin. Behav. Sci.* **3**, 84–89. (doi:10.1016/j.cobeha.2015.02.001)
11. Clayton NS. 2004 Is necessity the mother of innovation? *Trends Cogn. Sci.* **8**, 98–99. (doi:10.1016/j.tics.2004.01.001)
12. Ramsey G, Bastian ML, van Schaik C. 2007 Animal innovation defined and operationalized. *Behav. Brain Sci.* **30**, 393–407. (doi:10.1017/S0140525X07002373)
13. Griffin AS, Guez D. 2014 Innovation and problem solving: a review of common mechanisms. *Behav. Process.* **109**, 121–134. (doi:10.1016/j.beproc.2014.08.027)
14. Fogarty L, Creanza N, Feldman MW. 2015 Cultural evolutionary perspectives on creativity and human innovation. *Trends Ecol. Evol.* **30**, 736–754. (doi:10.1016/j.tree.2015.10.004)
15. Kolodny O, Creanza N, Feldman MW. 2015 Evolution in leaps: the punctuated accumulation and loss of cultural innovations. *Proc. Natl Acad. Sci.* USA **112**, E6762–E6769. (doi:10.1073/pnas.1520492112)
16. Laland K, Reader S. 1999 Foraging innovation in the guppy. *Anim. Behav.* **57**, 331–340. (doi:10.1006/anbe.1998.0967)
17. Boogert NJ, Reader SM, Laland KN. 2006 The relation between social rank, neophobia and individual learning in starlings. *Anim. Behav.* **72**, 1229–1239. (doi:10.1016/j.anbehav.2006.02.021)
18. Tebbich S, Stankewitz S, Teschke I. 2012 The relationship between foraging, learning abilities and neophobia in two species of Darwin's finches. *Ethology* **118**, 135–146. (doi:10.1111/j.1439-0310.2011.02001.x)
19. Benson-Amram S, Holekamp KE. 2012 Innovative problem solving by wild spotted hyenas. *Proc. R. Soc. B* **279**, 4087–4095. (doi:10.1098/rspb.2012.1450)
20. Thornton A, Samson J. 2012 Innovative problem solving in wild meerkats. *Anim. Behav.* **83**, 1459–1468. (doi:10.1016/j.anbehav.2012.03.018)
21. Kacelnik A. 2009 Tools for thought or thoughts for tools? *Proc. Natl Acad. Sci. USA* **106**, 10 071–10 072. (doi:10.1073/pnas.0904735106)
22. Kendal RL, Coe RL, Laland KN. 2005 Age differences in neophilia, exploration, and innovation in family groups of callitrichid monkeys. *Am. J. Primatol.* **66**, 167–188. (doi:10.1002/ajp.20136)
23. Katzir G. 1982 Relationship between social structure and response to novelty in captive jackdaws, *Corvus monedula* L. I. Response to novel space. *Behaviour* **81**, 231–263. (doi:10.1163/156853982X00157)
24. Cole EF, Quinn JL. 2012 Personality and problem-solving performance explain competitive ability in the wild. *Proc. R. Soc. B* **279**, 1168–1175. (doi:10.1098/rspb.2011.1539)
25. Barnard CJ, Luo N. 2002 Acquisition of dominance status affects maze learning in mice. *Behav. Process.* **60**, 53–59. (doi:10.1016/S0376-6357(02)00121-3)
26. Spencer KA, Buchanan KL, Goldsmith AR, Catchpole CK. 2004 Developmental stress, social rank and song complexity in the European starling (*Sturnus vulgaris*). *Proc. R. Soc. Lond. B* **271**, S121–S123. (doi:10.1098/rsbl.2003.0122)
27. Kummer H, Goodall J. 1985 Conditions of innovative behavior in primates. *Phil. Trans. R. Soc. Lond. B* **308**, 203–214.
28. Kummer H. 1978 On the value of social relationships to nonhuman primates: a heuristic scheme. *Soc. Sci. Inf.* **17**, 687–705. (doi:10.1177/053901847801700418)
29. Henrich J, Gil-White FJ. 2001 The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evol. Hum. Behav.* **22**, 165–196. (doi:10.1016/S1090-5138(00)00071-4)
30. Aplin LM, Sheldon BC, Morand-Ferron J. 2013 Milk bottles revisited: social learning and individual variation in the blue tit, *Cyanistes caeruleus*. *Anim. Behav.* **85**, 1225–1232. (doi:10.1016/j.anbehav.2013.03.009)
31. Reader S, Laland K. 2000 Diffusion of foraging innovations in the guppy. *Anim. Behav.* **60**, 175–180. (doi:10.1006/anbe.2000.1450)
32. Bouchard J, Goodyer W, Lefebvre L. 2007 Social learning and innovation are positively correlated in pigeons (*Columba livia*). *Anim. Cogn.* **10**, 259–266. (doi:10.1007/s10071-006-0064-1)
33. Burkart JM, Strasser A, Foglia M. 2009 Trade-offs between social learning and individual innovativeness in common marmosets, *Callithrix jacchus*. *Anim. Behav.* **77**, 1291–1301. (doi:10.1016/j.anbehav.2009.02.006)
34. Kendal R, Hopper LM, Whiten A, Brosnan SF, Lambeth SP, Schapiro SJ, Hoppitt W. 2015 Chimpanzees copy dominant and knowledgeable individuals: implications for cultural diversity. *Evol. Hum. Behav.* **36**, 65–72. (doi:10.1016/j.evolhumbehav.2014.09.002)
35. Biro D, Sousa C, Matsuzawa T. 2006 Ontogeny and cultural propagation of tool use by wild chimpanzees at Bossou, Guinea: case studies in nut cracking and leaf folding. In *Cognitive development in chimpanzees* (eds T Matsuzawa, M Tomonaga, M Tanaka), pp. 476–508. Tokyo, Japan: Springer.
36. Reader SM, Laland KN. 2001 Primate innovation: sex, age and social rank differences. *Int. J. Primatol.* **22**, 787–805. (doi:10.1023/A:1012069500899)
37. Laland KN. 2004 Social learning strategies. *Learn. Behav.* **32**, 4–14. (doi:10.3758/BF03196002)
38. Boyd R, Richerson PJ. 1985 *Culture and the evolutionary process*. Chicago, IL: University of Chicago Press.
39. Zahavi A. 1995 Altruism as a handicap—the limitations of kin selection and reciprocity. *J. Avian Biol.* **26**, 1–3. (doi:10.2307/3677205)

40. Boogert NJ, Fawcett TW, Lefebvre L. 2011 Mate choice for cognitive traits: a review of the evidence in nonhuman vertebrates. *Behav. Ecol*. **22**, 447–459. (doi:10.1093/beheco/arq173)
41. Cole EF, Morand-Ferron J, Hinks AE, Quinn JL. 2012 Cognitive ability influences reproductive life history variation in the wild. *Curr. Biol*. **22**, 1808–1812. (doi:10.1016/j.cub.2012.07.051)
42. Cauchard L, Boogert NJ, Lefebvre L, Dubois F, Doligez B. 2013 Problem-solving performance is correlated with reproductive success in a wild bird population. *Anim. Behav*. **85**, 19–26. (doi:10.1016/j.anbehav.2012.10.005)
43. Keagy J, Savard JF, Borgia G. 2009 Male satin bowerbird problem-solving ability predicts mating success. *Anim. Behav*. **78**, 809–817. (doi:10.1016/j.anbehav.2009.07.011)
44. Stammbach E. 1988 Group responses to specially skilled individuals in a *Macaca fascicularis* group. *Behaviour* **107**, 241–266. (doi:10.1163/156853988X00368)
45. Goodall J. 1986 *The chimpanzee of Gombe: patterns of behavior*. Cambridge, MA: Belknap Press.
46. Frumkin M. 1945 The origin of patents. *J. Pat. Off. Soc*. **27**, 143–149.
47. Pinker S. 2010 Colloquium paper: the cognitive niche: coevolution of intelligence, sociality, and language. *Proc. Natl Acad. Sci. USA* **107**, 8993–8999. (doi:10.1073/pnas.0914630107)
48. Borgeaud C, Bshary R. 2015 Wild vervet monkeys trade tolerance and specific coalitionary support for grooming in experimentally induced conflicts. *Curr. Biol*. **25**, 3011–3016. (doi:10.1016/j.cub.2015.10.016)
49. Fruteau C, van de Waal E, van Damme E, Noë R. 2011 Infant access and handling in sooty mangabeys and vervet monkeys. *Anim. Behav*. **81**, 153–161. (doi:10.1016/j.anbehav.2010.09.028)
50. Gumert MD. 2007 Payment for sex in a macaque mating market. *Anim. Behav*. **74**, 1655–1667. (doi:10.1016/j.anbehav.2007.03.009)
51. Morgan TJH *et al*. 2015 Experimental evidence for the co-evolution of hominin tool-making teaching and language. *Nat. Commun*. **6**, 6029. (doi:10.1038/ncomms7029)
52. Dean LG, Kendal RL, Schapiro SJ, Thierry B, Laland KN. 2012 Identification of the social and cognitive processes underlying human cumulative culture. *Science* (*80-*.). **335**, 1114–1118. (doi:10.1126/science.1213969)
53. Aoki K, Wakano JY, Feldman MW. 2005 The emergence of social learning in a temporally changing environment: a theoretical model. *Curr. Anthropol*. **46**, 334–340. (doi:10.1086/428791)