

# The evolution of superstitious and superstition-like behaviour

Kevin R. Foster<sup>1,\*</sup> and Hanna Kokko<sup>2</sup>

<sup>1</sup>*Center for Systems Biology, Harvard University, 7 Divinity Avenue, Harvard, MA 02138, USA*

<sup>2</sup>*Laboratory of Ecological and Evolutionary Dynamics, Department of Biological and Environmental Sciences, University of Helsinki, 00014 Helsinki, Finland*

Superstitious behaviours, which arise through the incorrect assignment of cause and effect, receive considerable attention in psychology and popular culture. Perhaps owing to their seeming irrationality, however, they receive little attention in evolutionary biology. Here we develop a simple model to define the condition under which natural selection will favour assigning causality between two events. This leads to an intuitive inequality—akin to an amalgam of Hamilton’s rule and Pascal’s wager—that shows that natural selection can favour strategies that lead to frequent errors in assessment as long as the occasional correct response carries a large fitness benefit. It follows that incorrect responses are the most common when the probability that two events are really associated is low to moderate: very strong associations are rarely incorrect, while natural selection will rarely favour making very weak associations. Extending the model to include multiple events identifies conditions under which natural selection can favour associating events that are never causally related. Specifically, limitations on assigning causal probabilities to pairs of events can favour strategies that lump non-causal associations with causal ones. We conclude that behaviours which are, or appear, superstitious are an inevitable feature of adaptive behaviour in all organisms, including ourselves.

**Keywords:** optimality theory; behavioural ecology; animal behaviour

## 1. INTRODUCTION

Although the concept of superstition encompasses a wide range of beliefs and behaviours, most can be united by a single underlying property—the incorrect establishment of cause and effect: ‘a belief or practice resulting from ignorance, fear of the unknown, trust in magic or chance, or a false conception of causation’ (Merriam-Webster online dictionary). In a world increasingly dominated by science, superstitious and indeed religious thinking typically take a back seat in academic affairs. However, superstitions play a central role in many small-scale societies, and indeed remain prevalent in the popular culture of all societies. Why is this? Can science rationalize this seemingly most irrational aspect of human behaviour?

Superstitions receive considerable attention in several fields including popular psychology (Shermer 1998; Vyse 2000; Wheen 2004), philosophy (Scheibe & Sarbin 1965), abnormal psychology (Devenport 1979; Brugger *et al.* 1994; Shaner 1999; Nayha 2002) and medicine (Hira *et al.* 1998; Diamond 2001), which typically frame superstitions as irrational mistakes in cognition. A notable exception, however, is found in the introduction to the popular book of Shermer (1998). This argues that superstitions are the adaptive outcome of a general ‘belief engine’, which evolved to both reduce anxiety (proximate cause) and enable humans to make causal associations (ultimate cause) (Tinbergen 1963; West *et al.* 2007). Specifically, Shermer argued that in making causal associations,

humans are faced with the option to minimize one of two types of statistical error: type I errors whereby they believe a falsehood or type II errors whereby they reject a truth. And as long as the cost of type II errors is high enough, natural selection can favour strategies that frequently make type I errors and generate superstitions (see Beck & Forstmeier (2007) for a similar argument). Our goal here is to explore Shermer’s idea that superstitions are adaptive.

Previous biological accounts of superstition have focused upon the classic work of the behavioural psychologist Skinner who reported superstitious behaviour in pigeons (Skinner 1948; Morse & Skinner 1957). In one of his experiments on operant conditioning, Skinner presented the pigeons with food at random intervals and noted that they still displayed ritualized behaviours that he interpreted as superstitious, i.e. the pigeon was behaving as though its actions were causing the food to arrive. However, these behaviours were later reinterpreted as behaviours that improve foraging efficacy (analogous to salivation in Pavlov’s dogs), which suggests that the pigeons’ behaviour does not correspond to Skinner’s intended meaning of superstition (Staddon & Simmelhag 1971; Timberlake & Lucas 1985; Moore 2004). Nevertheless, Skinner’s early account is notable in two respects. First, it recognized the possibility of superstition occurring outside the human realm. Second, and linked to this, Skinner emphasized the behavioural aspect of superstition: ‘The bird behaves as if there were a causal relation between its behavior and the presentation of food, although such a relation is lacking.’ (Skinner 1948). That is, he focused on there being an incorrect response

\* Author for correspondence (kfoster@cgr.harvard.edu).

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2008.0981> or via <http://journals.royalsociety.org>.

**Box 1. Hamilton's rule and the outcome-based definition of altruism.**

In order to investigate the evolution of superstition, we focus upon the outcome of superstitious behaviours as opposed to the form of their psychological representation. This approach has important precedent in the study of behaviour that came some hundred years prior to Skinner: the founding discussions of the concept of altruism. The originator of the term, Augustus Comte, emphasized the psychological intent to do good (psychological process), but his contemporary Herbert Spencer often used a behavioural definition and applied it to the simplest of organisms (Dixon 2008). Spencer's perspective is the basis of the modern biologist's definition of altruism, where it has paid great dividends in explaining those apparently paradoxical behaviours that reduce lifetime personal reproduction but help others (Foster 2008). Importantly, studies based upon the biological definition of altruism have also made major (if incomplete) contributions to the study of such behaviours in humans (e.g. Darwin 1871; Sober & Wilson 1997; Fehr & Fischbacher 2003; Bowles 2006; Boyd 2006), even though human altruism functions through a complex underlying psychology.

Our modern understanding of the evolution of altruism is often summarized using Hamilton's rule:  $br > c$  (Hamilton 1964). As for inequality (2.4), the logic of this equation rests upon there being a reliable cost to the social action  $c$  and a less reliable benefit  $b$ , which instead of weighting by  $p$  is weighted by  $r$ : genetic relatedness between an actor and recipient. There are various ways to interpret Hamilton's rule but the analogy with inequality (2.4) is the strongest for the 'direct-fitness' version, which captures all fitness effects of the action in terms of the effects on a focal individual (Frank 1998). In this case  $r$ , as  $p$ , captures the probability that the actor will receive the potential benefits from the action with the difference that Hamilton's rule is about social actions such that the benefits come from the actions of others rather than the action of focal individual themselves.

to a stimulus (behavioural outcome), rather than the conscious abstract representation of cause and effect (psychological relationship), with which human superstitions are often associated.

We follow the Skinnerian perspective here and adopt his outcome-based behavioural definition, rather than the one of psychological representation. This focuses the analysis upon the relevant evolutionary currency, the behaviour, that has fitness consequences, as has been done for the evolutionary study of altruism before us (box 1). Our approach then will not speak directly to the psychology of superstition, but instead aims to form some groundwork for understanding why innate tendencies towards superstitious behaviour might evolve in all organisms, including ourselves. In addition, although all that follows is fully compatible with the potential cultural influences, we deliberately do not model these here. This is not to deny the great importance of culture in shaping the exact nature of superstitious beliefs in humans (Laland & Brown 2002; Richerson & Boyd 2004), but rather to focus the analysis on one key question: under what conditions might a tendency for performing behaviours that incorrectly assign cause and effect be adaptive from an individual fitness point of view?

## 2. RESULTS AND DISCUSSION

### (a) Basic model

Our goal is to capture the conditions under which responding, as though there is a causal link between two events, will be favoured by natural selection and, in particular, to identify the conditions when a response is evolutionarily favoured but the causal link may be lacking. In the model, one event  $E_P$  precedes the other  $E_L$ . The prior event has no effect on fitness *per se* (e.g. a noise), but the latter does (e.g. a predator arrives), and the focal individual's reaction (e.g. evasive action) can modify the fitness consequences of this latter event. Will acting as though believing that the second event  $E_L$  will also occur increase or decrease the individual's reproductive fitness? Table 1 captures all possible fitness effects of the action and latter event in terms of  $\alpha$  variables that define additive effects upon fitness and  $\gamma$  variables that define non-additive interactions between the action and the latter event upon fitness. A fraction  $p$  of  $E_P$  events are followed

Table 1. Fitness effects of responding, or not responding, to a prior event that may precede a latter event.

	latter event occurs	event does not occur
action performed ( $s=1$ )	$W_b + \alpha_1 + \alpha_2 + \gamma_3$	$W_b + \alpha_2 + \gamma_2$
no action ( $s=0$ )	$W_b + \alpha_1 + \gamma_1$	$W_b$

by  $E_L$ , such that  $p = P(E_L|E_P)$  and an individual's fitness  $W$  can be expressed as

$$W = W_b + s[p(\gamma_3 - \gamma_1) + (1-p)\gamma_2 + \alpha_2] + p(\alpha_1 + \gamma_1). \quad (2.1)$$

Here,  $s$  denotes a tendency to perform as though  $E_L$  will occur. We assume that  $s$  can take any value between 0 and 1 and that it is affected by genetics. When denoting the relevant genotypic value by  $g$ , we have  $ds/dg > 0$ . Importantly, this does not exclude the possibility of other influences, including cultural factors that may make the relationship between the genotype  $g$  and the behavioural phenotype  $s$  complex. However, as long as there is a finite positive influence of genotype on the tendency to display phenotype  $s$ ,  $ds/dg$  will not affect the direction of selection, but only its magnitude. We are interested in the relationship between the genotype  $g$  and the fitness  $W$  (Price 1970). Therefore, from the chain rule, we obtain

$$\frac{dW}{dg} = \frac{dW}{ds} \cdot \frac{ds}{dg}. \quad (2.2)$$

This shows that the essential evolutionary question is whether  $dW/ds > 0$ , and we can, without loss of generality, focus on the simplest case of direct correspondence between phenotype and genotype,  $ds/dg = 1$ . Then  $dW/dg$  is simply given by

$$\frac{dW}{dg} = \frac{dW}{ds} = \alpha_2 + (1-p)\gamma_2 + p(\gamma_3 - \gamma_1). \quad (2.3a)$$

And a response to the prior event will be favoured when

$$p(\gamma_3 - \gamma_1) > -\alpha_2 - (1-p)\gamma_2, \quad (2.3b)$$

where  $\alpha_2$  can be interpreted as an energetic cost  $c$  of performing the action, such that  $c = -\alpha_2$ , and the sum  $-\gamma_1 + \gamma_3$  captures the net benefit to performing the action when it is appropriate, which can be manifest either through

a benefit to acting ( $\gamma_3$ ) or a cost to not acting ( $-\gamma_1$ ). The final interaction term  $\gamma_2$  is important if performing the action in the absence of the latter event carries a specific cost. Take, for example, the tendency to eat plants to cure diseases in a small-scale society. Call the prior event  $E_P$  being ill, the latter  $E_L$  that conditions will worsen and the action eating a particular plant. The cost of obtaining the plant is  $-\alpha_2$ , the improvement in health from eating the plant is captured by  $\gamma_3$ , and if conditions were going to improve and the plant causes poisoning, this would mean  $\gamma_2 < 0$ . However, focusing on what is probably the typical case where the action provides a benefit  $b$  specific to the occurrence of the latter event, i.e.  $-\gamma_1 + \gamma_3 = b > 0$ ,  $\gamma_2 = 0$ , responding to the prior event is favoured when

$$pb > c. \quad (2.4)$$

This simple inequality provides an intuitive solution to the question of when it will be beneficial to assign causality between two events. Centrally, natural selection can favour associated events whose relationship is uncertain ( $0 < p < 1$ ) whenever there is a high fitness benefit to assigning causality correctly, e.g. not being eaten by a predator or being cured of a dangerous disease.

For extremely low values of  $p$ , inequality (2.4) can also be viewed as an evolutionary fitness version of Pascal's wager (Jordan 1994). This is the argument that one should be Christian even if the existence of God is highly improbable, because the pay-off from being correct is considered very great (figure 1). However, inappropriate actions based on extreme probabilities may be rare in the biological setting. Although assigning causality to highly improbable events in our model will indeed generate a large proportion of actions that do not individually carry a benefit ( $1-p$ , figure 1), it also requires disproportionately large fitness benefits for the action to be favoured overall by natural selection. This latter effect—the probability that a response is favoured by natural selection—works against the first by making the evolution of traits associated with a very low  $p$  unlikely.

We can include it in our evolutionary model by considering a prior-latter event relationship where again  $P(E_L|E_P) = p$ , and including the probability that the benefit-to-cost ratio ( $b/c = z$ ) is high enough to favour the strategy, from inequality (2.4), i.e. when  $z > 1/p$ . Therefore, for a given distribution of  $z$ , we can calculate the probability that the benefit-to-cost ratio is above  $1/p$  and then weight this by the probability that the resulting actions are a correct assignment of causality ( $1-p$ ). Taking a Gaussian distribution in  $z$  with a mean of 1 for illustrative purposes, we obtain

$$F = (1-p) \operatorname{Erfc} \left( \frac{\frac{1}{p} - 1}{\sqrt{2}\sigma} \right), \quad (2.5)$$

where  $\operatorname{Erfc}$  is the complementary error function, and  $\sigma$  is the standard deviation of the Gaussian. Figure 2 shows this function plotted for various frequency distributions of the benefit-to-cost ratio of actions. This predicts that, as long as low benefit-to-cost ratios are more likely than high, selection will lead to most incorrect assignments for associations based upon low to intermediate probabilities of causality. Very strong associations are rarely incorrect, while natural selection will rarely favour making very weak associations (as is required for Pascal's wager).

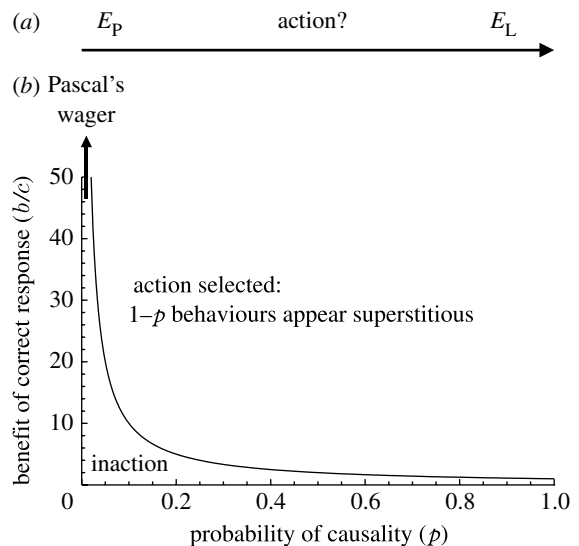


Figure 1. Natural selection for an action that leads to behaviours associated with an incorrect causal association. (a) One event  $E_P$  precedes the other  $E_L$ . The prior event has no effect on fitness *per se* (e.g. a noise), but the latter does (e.g. a predator arrives), and the focal individual's response (e.g. evasive action) can modify the fitness consequences of this latter event. (b) The evolution of the response. In the parameter space above the line, a focal individual will be selected to act as though the two events are causally associated even though  $1-p$  of the time the response is not needed, e.g. running away when a sound is heard in the absence of impending danger.

Solutions (2.4) and (2.5) then show that natural selection will readily favour strategies that generate a high frequency of individual behaviours that will appear superstitious, i.e. an action that implies a causal relationship that is lacking. And the general tendency for organisms to associate improbable events, which will often appear superstitious, may be thus explained, e.g. eating a particular plant to cure a particular disease may work only on very rare occasions. However, the logic of such examples ultimately rests on the two events sometimes, albeit rarely, being causally linked. If we take a strict definition of superstition, therefore, and demand the association of truly unrelated events, then these cases do not appear to be superstitious but rather, as Pascal himself emphasized, a good wager.

#### (b) *Extended model: multiple events*

Can natural selection ever favour the association of truly unrelated events, which would satisfy the most stringent definitions of superstition? Our model of a single cause-effect relationship suggests not but what about a mixture of prior and latter events, where some are causally related but others are not? We explore this possibility with a model where two prior events precede either one or two latter events (figure 3) and there is the possibility that the actor incorrectly assigns the probabilities of causality ( $p$  and  $q$ ) to the respective events. We focus on the first and simpler case here (figure 3a), but the conclusions from both cases are similar (below; electronic supplementary material). For clarity, we base the analysis on a specific scenario: two different types of disturbance that occur in an environment where a predator may appear, e.g. rustling in grass and rustling in trees, one of which is more typically caused by

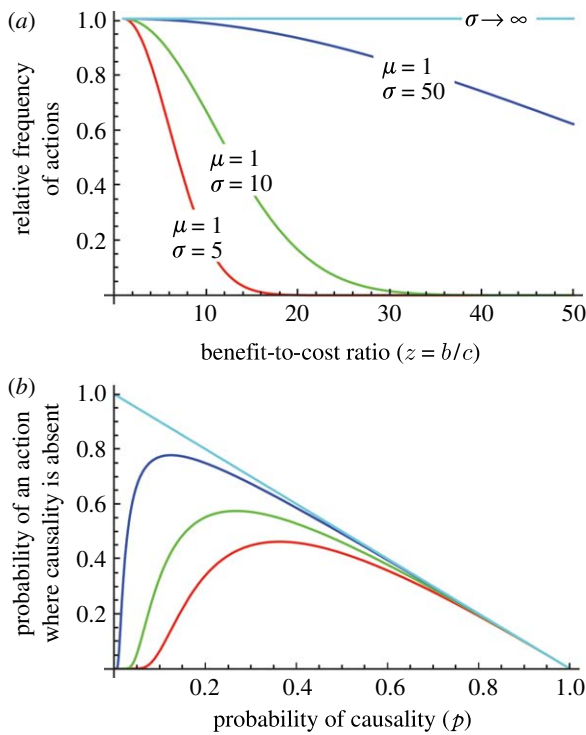


Figure 2. (a) Gaussian relative frequency distributions of actions as a function of their benefit-to-cost ratio ( $z$ ). Distributions have mean  $\mu$  and standard deviation  $\sigma$ . (b) Probability of an action evolving that predicts a latter event that does not occur, as a function of  $p$  (the probability that the prior and latter events are causally associated). The different plots are for the different probability density distributions of the benefit-to-cost ratio shown in (a). The line for  $\sigma \rightarrow \infty$  corresponds to all benefits from actions being equally likely, which reduces the relative frequency of incorrect responses to simply  $(1-p)$  (figure 1).

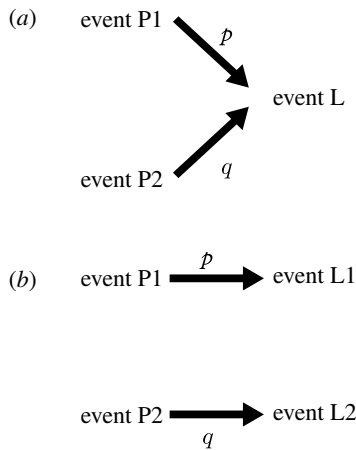


Figure 3. Models of multiple prior and latter events. (a) Single latter event and (b) two latter events.

wind, the other one by a moving predator. As before, our question is when will selection favour assuming that the prior and latter events are causally associated, even though the relationship between them is uncertain?

During a time unit, grass movement occurs with probability  $f$  and tree movement with probability  $g$ , and the probabilities the predator follows the events are  $p$  and  $q$ , respectively. There is additionally a conditional probability that the predator appears without any prior event; this equals  $r$  (table 2). If tree movement is not causally linked to the presence of the predator, we have

$q=r$ . The focal individual can make an anti-predator response that removes all risk from the predator (e.g. running down a burrow) but this comes at a cost  $c$  (equivalent to cost  $c = -\alpha_2$  in the first model). Meanwhile, not acting while the predator is around implies a risk of death,  $b$  (or equivalently the benefit from escaping when the predator is there), where  $b > c$ . Finally, we include the possibility that the actor may incorrectly assign the probabilities of causality ( $p$  and  $q$ ) to the respective events. This might occur for a number of reasons, including poor information on the real values of  $p$  and  $q$ . Here we assume that the incorrect assignment results simply from constraints on recognition of the two prior events, where movement in the grass is heard as movement in the trees with conditional probability  $P$  (hear prior event 2|prior event 1 occurs) =  $a_{21}$ , while the reverse occurs with probability  $a_{12}$ . There are then four possible strategies: (i) ignore all prior events, (ii) respond to grass, (iii) respond to trees, and (iv) respond to both.

This extended model is analysed with the methodology of the first model (equations (2.1)–(2.4)). In brief, we use the set of conditional probabilities shown in table 2 to derive a fitness equation for each of the four possible strategies that allow their fitness consequences to be compared (table S1 in the electronic supplementary material shows the equivalent calculations for a case with two latter events corresponding to figure 3b). The model contains many more parameters than the first but its key behaviour is captured by focusing on a simple case whereby the predator always disturbs the grass ( $r=0$ ) and grass movement precedes the predator's arrival with probability  $p$  but, importantly, the other prior event (movement in the trees) is never causally associated with the predator ( $q=r=0$ ). In addition, for illustration, we assume that the probability that movement in the grass is heard as in the trees is equal to the reverse ( $a_{21} = a_{12} = a$ ) and the two prior events occur rarely enough that the probability that both occur is negligible ( $fg \approx 0$ ), although table 2 shows how to calculate the case where sounds can co-occur (see also figure 4).

With these assumptions, we can evaluate the fitness consequences of various possible strategies. Most simply, a corollary of the tree movement never being associated with the predator is that it will nearly always be better to respond to the grass. In particular, the model predicts that selection will favour a response to grass rather than trees when

$$(2a - 1)[c(f - g) - bfp] > 0, \quad (2.6)$$

which is always true for  $a < 0.5$  (hearing the grass is a better predictor of movement in the grass occurring than hearing the trees) and  $bp > c$  (responding to the grass carries a net benefit when it is detected perfectly). Responding to the grass is favoured over no response when

$$\frac{f(1-a)}{f(1-a) + ga} pb > c. \quad (2.7)$$

Responding to both the grass and the trees rather than just grass gives

$$\frac{fa}{fa + g(1-a)} pb > c, \quad (2.8)$$

and responding to both events rather than no response is favoured when

$$\frac{f}{f+g} pb > c, \quad (2.9)$$



Table 2. Calculations of conditional survival probabilities as a function of behavioural strategy.

actual event	assigned event	frequency <sup>a</sup>	survival probability			
			no response	respond P1	respond P2	respond both
none	none	$(1-f)(1-g)$	$1-rb$	$1-rb$	$1-rb$	$1-rb$
P1	P1	$f(1-g)(1-a_{21})$	$1-pb$	$1-c$	$1-pb$	$1-c$
P2	P1	$(1-f)ga_{12}$	$1-qb$	$1-c$	$1-qb$	$1-c$
P2	P2	$(1-f)g(1-a_{12})$	$1-qb$	$1-qb$	$1-c$	$1-c$
P1	P2	$f(1-g)a_{21}$	$1-pb$	$1-pb$	$1-c$	$1-c$
both	P1	$fg(1-a_{21})a_{12}$	$(1-pb)(1-qb)$	$1-c$	$(1-pb)(1-qb)$	$1-c$
both	P2	$fga_{21}(1-a_{12})$	$(1-pb)(1-qb)$	$(1-pb)(1-qb)$	$1-c$	$1-c$
both	both	$fg(1-a_{21})(1-a_{12})+fga_{21}a_{12}$	$(1-pb)(1-qb)$	$1-c$	$1-c$	$1-c$

<sup>a</sup>The conditional survival probability for one strategy is calculated by multiplying frequency by the survival probability for that strategy and then summing these products for each row.

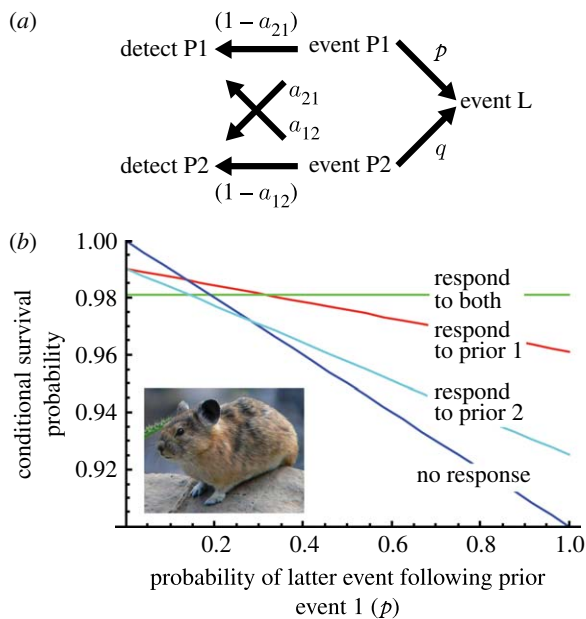


Figure 4. (a) The evolution of superstitious behaviour when multiple events precede a latter one. Should a prey species respond to neither, one or both prior events (e.g. moving grass or trees) as though they signal the arrival of a predator? (b) Effect of different strategies upon survival when prey species cannot perfectly discriminate between the prior events ( $a=0.3$ ). Responding to both events is often selected even though prior event 2 has no true association with the predator ( $q=0$ ). Other parameters:  $c=0.1$  (cost of anti-predator response),  $b=1$  (cost of predator being present),  $f=g=0.1$  (probability that each sound occurs) and  $r=0$  (probability that predator arrives when there is no sound). Inset photo of Pika used with kind permission from J. Bailey. See <http://bailey.aros.net/nature/>.

where  $(f/(f+g))pb > (fa/(fa+g(1-a)))pb$  when  $a < 0.5$ . So unless there are extreme errors in identifying the prior events (hearing prior event 1 is a better predictor of prior event 2 than prior event 1, and vice versa), when inequality (2.8) holds, so will inequality (2.9).

Inequalities (2.7) and (2.8) then allow us to assess the effects of multiple prior events on the evolution of superstitious or superstition-like behaviours. Centrally, they show that when there is some error in the discrimination between events and their associated effects, it will often to be beneficial to respond to both events, even though one, in this case movement in the trees, has no true association with a benefit. Figure 4 illustrates the

behaviour of the model: as the probability of a causal association between movement in the grass and the predator rises ( $p$ ), so does the likelihood of responding to non-causal prior events (movement in trees). This occurs because increased  $p$  means increased benefit to responding to the causal stimulus, which more easily compensates for the cost of mistakenly responding to the non-causal stimulus.

(c) **A hierarchical view of superstition**

Further intuition about responses to multiple events can be obtained by considering the two extreme cases where there are no errors in assigning causal probabilities, and complete error. With no errors ( $a=0$ ), inequality (2.8) is never satisfied and it is always better to respond to just the causal event, and from inequality (2.7) we recover our general inequality that a response is favoured when  $pb > c$ . More interesting, however, is the case when discrimination errors are so common that the two prior events are indistinguishable ( $a=0.5$ ). Then both inequalities 7 and 8 reduce to inequality (2.9), which captures the fact that the focal individual now has only two options: respond to both stimuli or do not respond at all. This scenario is useful because it distinguishes two levels of uncertainty that can drive superstitious or superstition-like behaviours. The lower level captures the fact that one prior event is truly causally associated with the latter event, but there may nevertheless be occasions when the latter event does not follow. As a result, there is the probability  $p$  that the latter follows the prior, where  $p \neq 1$ . In the absence of any second prior event ( $g=0$ ) then, we again recover equation (2.1).

The higher level effect occurs when there exists a second prior event that is not causally associated (movement in the trees,  $g > 0$ ). Now there can be cases where the focal individual responds to a prior event that has no causal association with the latter event, which is an association that might be considered formally superstitious. The benefit of responding depends upon the Bayesian probability that a prior event is the causal event rather than the non-causal one ( $f/(f+g)$ ), and when the first event is perfectly causally associated with the latter event ( $p=1$ ), it is this Bayesian ratio coupled to the benefit-to-cost ratio that drives any response. One can, of course, combine the two levels of uncertainty into a single probability,

$$P = \frac{f}{f+g}p, \tag{2.10}$$

where substituting  $P$  into inequality (2.9) returns again to the form of inequality (2.4) and the associated logic (figure 2).

This simple substitution reveals the hierarchical structure of the arguments presented here. When a causal prior event is not perfectly predictive of a latter event, it will often be possible, with more information, to subdivide the prior event into occasions that are sometimes causal and some that are never causal. And with more information, one might go further and subdivide the former set, and so on. However, whenever an actor cannot fully dissect out the prior events that carry perfect causality, there will be a level at which they are forced to respond to an aggregate of causal and non-causal events, or not respond at all, i.e. responding based upon  $P$  (higher level probability) rather than  $p$  and  $q$  (lower level probabilities). To the extent that the associations used by individuals can be viewed as an aggregate of causal and non-causal relationships, therefore, there is a case for the existence of superstitions that arise through natural selection.

The limits on an individual's ability to estimate and distinguish among causal probabilities will, in turn, depend on the available mechanisms of assessment. While estimates of causality may arise through learning, an evolutionary account of superstitious behaviours highlights that learning is not required for their generation. Our model of multiple prior events fits a vertebrate prey species that learns predator responses (figure 4) but it can also be applied to innate responses, where 'learning' will instead operate over evolutionary time scales. For example, there is evidence of an innate avoidance exhibited by some animals of harmless yellow and black insects because others of this coloration are dangerous (Werner & Elke 1985). There may also be a genetic component to the finding that predators only avoid non-poisonous snakes that mimic a poisonous species in areas where the poisonous species is common (Pfennig *et al.* 2001). Such data also support the predictions of our model. In our terminology, an increase in the frequency of the causal relationship 'eat snake and be poisoned' leads to an incorrect and superstitious association being formed with the non-causal association. In other words, the ratio  $f/(f+g)$  increases with increasing  $f$  and favours avoiding the harmless snakes. Moving even further away from the potential influence of learning, our model could equally describe a bacterium swimming towards a substance that carries no metabolic benefits, e.g. *Escherichia coli* cells will swim towards physiologically inert methylaspartate presumably owing to an adaptation to favour true aspartate (Sourjik & Berg 2002).

Appreciating the biological basis for such simple responses is more than a curiosity. It emphasizes that organisms will frequently display behavioural responses that appear poorly adapted to a particular situation. The existence of superstitious behaviours that are part of an adaptive strategy provides an alternative explanation for behaviours that might otherwise be seen as maladaptive 'mistakes' (Gadagkar 1993) or a poor match between a species and a changing environment ('ecological traps'; Schlaepfer *et al.* 2002). This said, we do not intend to suggest that all behaviours are adaptive and many ecological traps are probably real. Indeed, an evolutionary lag following a changed environment provides another route to superstitious behaviours, whereby an organism associates two events that once were, but are no longer,

causally related, e.g. a predator goes extinct but the prey still hides at night.

In humans, assessments of causality and the associated responses reach their most complex level owing to the potential for reasoning and cultural transmission. Along these lines, Beck & Forstmeier (2007) have recently argued that prior experience (an individual's 'world view') will weigh heavily in whether a current relationship is deemed true or false. This potential interaction between reasoning and culture is apparent in the observation that the dawn of human agriculture coincided with a tendency to use green beads in jewellery. This may indicate that when crops became important, individuals began to reason that green icons brought good fortune (Bar-Yosef Mayer & Porat 2008). In addition, alternative medicine has a strong culturally learned component that appears to frequently group ineffective medicines with effective ones (Astin *et al.* 1998).

The enormous potential for cultural evolution to affect the links between genetics and behaviour (Laland & Brown 2002; Richerson & Boyd 2004; Lehmann *et al.* 2008) means that our reductionist model must be cautiously applied to humans. Nevertheless, our work suggests that the acquisition of new information through learning, copying and hearsay is all underlain by the innate and adaptive propensities to act on uncertainties. In particular, the inability of individuals—human or otherwise—to assign causal probabilities to all sets of events that occur around them will often force them to lump causal associations with non-causal ones. From here, the evolutionary rationale for superstition is clear: natural selection will favour strategies that make many incorrect causal associations in order to establish those that are essential for survival and reproduction.

Authors express thanks to Andrés López-Sepulcre, Katja Bargum, Daniel Rankin, Katharina Ribbeck, Carey Nadell, Andrew Murray, Allan Drummond, Martin Willensdorfer, and two anonymous referees for their helpful comments and discussions. K.R.F. is supported by National Institute of General Medical Sciences Center of Excellence Grant 5P50 GM 068763-01.

## REFERENCES

- Astin, J. A., Marie, A., Pelletier, K. R., Hansen, E. & Haskell, W. L. 1998 A review of the incorporation of complementary and alternative medicine by mainstream physicians. *Arch. Int. Med.* **158**, 2303–2310. (doi:10.1001/archinte.158.21.2303)
- Bar-Yosef Mayer, D. E. & Porat, N. 2008 Green stone beads at the dawn of agriculture. *Proc. Natl Acad. Sci. USA* **104**, 9964–9969. (doi:10.1073/pnas.0709931105)
- Beck, J. & Forstmeier, W. 2007 Superstition and belief as inevitable by-products of an adaptive learning strategy. *Hum. Nat.* **18**, 35–46. (doi:10.1007/BF02820845)
- Bowles, S. 2006 Group competition, reproductive leveling, and the evolution of human altruism. *Science* **314**, 1569–1572. (doi:10.1126/science.1134829)
- Boyd, R. 2006 The puzzle of human sociality. *Science* **314**, 1555–1556. (doi:10.1126/science.1136841)
- Brugger, P., Dowdy, M. A. & Graves, R. E. 1994 From superstitious behavior to delusional thinking: the role of the hippocampus in misattributions of causality. *Med. Hypotheses* **43**, 397–402. (doi:10.1016/0306-9877(94)90015-9)
- Darwin, C. 1871 *The descent of man and selection in relation to sex*. London, UK: John Murray.

- Devenport, L. D. 1979 Superstitious bar pressing in hippocampal and septal rats. *Science* **205**, 721–723. (doi:10.1126/science.462183)
- Diamond, J. 2001 *Snake oil and other preoccupations*. London, UK: Vintage.
- Dixon, T. 2008 *The invention of altruism: making moral meanings in Victorian Britain*. Oxford, UK: Oxford University Press for the British Academy.
- Fehr, E. & Fischbacher, U. 2003 The nature of human altruism. *Nature* **425**, 785–791. (doi:10.1038/nature02043)
- Foster, K. R. 2008 Behavioral ecology: altruism. In *Encyclopedia of ecology*, pp. 154–159. Amsterdam, The Netherlands: Elsevier.
- Frank, S. A. 1998 *The foundations of social evolution. Monographs in Behavior and Ecology*. Princeton, NJ: Princeton University Press.
- Gadagkar, R. 1993 Can animals be spiteful? *Trends Ecol. Evol.* **8**, 232–234. (doi:10.1016/0169-5347(93)90196-V)
- Hamilton, W. D. 1964 The genetical evolution of social behaviour. I & II. *J. Theor. Biol.* **7**, 1–52. (doi:10.1016/0022-5193(64)90038-4)
- Hira, K., Fukui, T., Endoh, A., Rahman, M. & Maekawa, M. 1998 Influence of superstition on the date of hospital discharge and medical cost in Japan: retrospective and descriptive study. *Br. Med. J.* **317**, 1680–1683.
- Jordan, J. (ed.) 1994 *Pascal's wager: pragmatic arguments and belief in God*. Lanham, MD: Rowman & Littlefield.
- Laland, K. N. & Brown, G. R. 2002 *Sense & nonsense: evolutionary perspectives on human behaviour*. Oxford, UK: Oxford University Press.
- Lehmann, L., Foster, K. R., Borenstein, E. & Feldman, M. W. 2008 Social and individual learning of helping in humans and other species. *Trends Ecol. Evol.* **23**, 664–671. (doi:10.1016/j.tree.2008.07.012)
- Moore, B. R. 2004 The evolution of learning. *Biol. Rev. Camb. Philos. Soc.* **79**, 301–335. (doi:10.1017/S1464793103006225)
- Morse, W. H. & Skinner, B. F. 1957 A second type of superstition in the pigeon. *Am. J. Psychol.* **70**, 308–311. (doi:10.2307/1419345)
- Nayha, S. 2002 Traffic deaths and superstition on Friday the 13th. *Am. J. Psychiatry* **159**, 2110–2111. (doi:10.1176/appi.ajp.159.12.2110)
- Pfennig, D. W., Harcombe, W. R. & Pfennig, K. S. 2001 Frequency-dependent Batesian mimicry. *Nature* **410**, 323. (doi:10.1038/35066628)
- Price, G. R. 1970 Selection and covariance. *Nature* **227**, 520–521. (doi:10.1038/227520a0)
- Richerson, P. J. & Boyd, R. 2004 *Not by genes alone: how culture transformed human evolution*. Chicago, IL: University of Chicago Press.
- Scheibe, K. E. & Sarbin, T. R. 1965 Towards a theoretical conceptualisation of superstition. *Br. J. Philos. Sci.* **16**, 143–158. (doi:10.1093/bjps/XVI.62.143)
- Schlaepfer, M. A., Runge, M. C. & Sherman, P. W. 2002 Ecological and evolutionary traps. *Trends Ecol. Evol.* **17**, 474–480. (doi:10.1016/S0169-5347(02)02580-6)
- Shaner, A. 1999 Delusions, superstitious conditioning and chaotic dopamine neurodynamics. *Med. Hypotheses* **52**, 119–123. (doi:10.1054/mehy.1997.0656)
- Shermer, M. 1998 *Why people believe weird things: pseudoscience, superstition, and other confusions of our time (paperback)*. New York, NY: W. H. Freeman & Co.
- Skinner, B. F. 1948 Superstition in the pigeon. *J. Exp. Psychol.* **38**, 168–172. (doi:10.1037/h0055873)
- Sober, E. & Wilson, D. S. 1997 *Unto others: the evolution of altruism*. Cambridge, MA: Harvard University Press.
- Sourjik, V. & Berg, H. C. 2002 Receptor sensitivity in bacterial chemotaxis. *Proc. Natl Acad. Sci. USA* **99**, 123–127. (doi:10.1073/pnas.011589998)
- Staddon, J. E. R. & Simmelhag, V. L. 1971 The “Superstition” experiment: a reexamination of its implications for the principles of adaptive behavior. *Psychol. Rev.* **78**, 3–43. (doi:10.1037/h0030305)
- Timberlake, W. & Lucas, G. A. 1985 The basis of superstitious behavior: chance contingency, stimulus substitution, or appetitive behavior? *J. Exp. Anal. Behav.* **44**, 279–299. (doi:10.1901/jeab.1985.44-279)
- Tinbergen, N. 1963 On aims and methods of ethology. *Z. Tierpsychol.* **20**, 410–433.
- Vyse, S. A. 2000 *Believing in magic. The psychology of superstition*. Oxford, UK: Oxford University Press.
- Werner, S. & Elke, H. 1985 On the function of warning coloration: a black and yellow pattern inhibits prey-attack by naive domestic chicks. *Behav. Ecol. Sociobiol.* **16**, 249.
- West, S. A., Griffin, A. S. & Gardner, A. 2007 Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *J. Evol. Biol.* **20**, 415–432. (doi:10.1111/j.1420-9101.2006.01258.x)
- Wheen, F. 2004 *How mumbo-jumbo conquered the world*. New York, NY: Harper Perennial.