Animal Behaviour 138 (2018) 59–67

Contents lists available at ScienceDirect

Animal Behaviour

journal homepage: www.elsevier.com/locate/anbehav

The adaptive significance of age-dependent changes in the tendency of individuals to explore

Thomas N. Sherratt [a,*], Julie Morand-Ferron [b]

[a] Department of Biology, Carleton University, Ottawa, ON, Canada
[b] Department of Biology, University of Ottawa, Ottawa, ON, Canada

A R T I C L E I N F O

Article history:
Received 8 August 2017
Initial acceptance 15 September 2017
Final acceptance 4 December 2017

MS. number: A17-00639R2

Keywords:
ageing
bandit model
Bayesian learning
dynamic programming
exploration–exploitation
neophilia
neophobia
novelty

Many organisms show a reduced tendency to investigate unfamiliar objects as they age. Although the phenomenon could arise for a range of reasons, it is possible that this age-related increase in conservatism is adaptive. In particular, we propose that novel objects encountered late in life will be perceived as being relatively rare, so the value of information from investigating their properties will be estimated to be low. In addition, agents that investigate novel objects late in their lives will have little time left to exploit this information should the objects turn out to be profitable. We formalize the above arguments by developing an exploration–exploitation ('bandit') model. In this model, agents must decide whether to explore or ignore a novel object that it has just encountered at a given stage in its life, despite uncertainty regarding the commonness of the object in the environment and the likelihood that the object is profitable/unprofitable. We assume that, as agents encounter (and possibly investigate) unfamiliar objects, they use Bayesian inference to update their beliefs about the objects' commonness and profitability. Dynamic programming is concurrently used to identify the conditions under which the agent should explore or ignore these objects. If the benefit/cost ratio of investigating novel objects is high, then all individuals will be selected to explore regardless of their age. Likewise, if the ratio is low, then all individuals should ignore novel objects. Under intermediate conditions, young individuals that encounter novel objects should investigate their properties, while older ones should ignore them. The optimal switch in strategy arises as a consequence of age-dependent variation in both the novel object's perceived abundance and the future value of information regarding the object's profitability. We highlight several additional testable predictions of the model and discuss alternative adaptive explanations.

Many animals show an aversion to investigating the properties of objects they have never encountered before, a behaviour known as neophobia (Greggor, Thornton, & Clayton, 2015). Intriguingly, there is evidence that the tendency of individuals to exhibit such neophobia increases with age, while the tendency to spontaneously explore novel objects, termed neophilia, decreases with age. For example, immature wild baboons (*Papio ursinus*) and geladas (*Theropithecus gelada*) were found to exhibit more neophilia towards novel objects (a doll, ball or can) than conspecific adults (Bergman & Kitchen, 2009). Likewise, juvenile chimango caracara, *Milvago chimango*, show a greater tendency to investigate novel plastic objects and less neophobia than adults (Biondi, Bo, & Vassallo, 2010, 2013, 2015). When presented with baited puzzle-boxes, wild juvenile spotted hyaena, *Crocuta crocuta*, were more exploratory, less neophobic and more persistent than adults (Benson-Amram & Holekamp, 2012; Benson-Amram, Weldele, & Holekamp, 2013). Similarly, juvenile great tits, *Parus major*, and blue tits, *Cyanistes caeruleus*, visited novel problem-solving devices more frequently than adult conspecifics (Morand-Ferron, Cole, Rawles, & Quinn, 2011, 2015). Laboratory rodents also often demonstrate age-related declines in exploratory tendency from maturity onwards (e.g. longitudinal designs: Joyal et al., 2000; cross-sectional designs: Lamberty & Gower, 1990), while extensive surveys indicate that the openness of humans to new experiences tends to decline (albeit slightly) from middle age (e.g. Costa et al., 1986; Donnellan & Lucas, 2008; Roberts, Walton, & Viechtbauer, 2006).

Importantly, not all studies have revealed a decline in the tendency to explore with age (e.g. see Herborn et al., 2010; Hopper et al., 2014; Kendal, Coe, & Laland, 2005), and sometimes the relationship between age and exploratory tendency is complex. For

* Correspondence: T. N. Sherratt, Department of Biology, Carleton University, Ottawa, ON K1S 5B6, Canada.
E-mail address: Tom.Sherratt@Carleton.ca (T. N. Sherratt).

https://doi.org/10.1016/j.anbehav.2018.01.025
0003-3472/

example, 2-month-old mound-building mice, *Mus spicilegus*, showed a higher willingness to investigate a novel object (a white ball) compared to 12-month-old adults, but mice at 24 months of age behaved in a similar way to the juveniles (Lafaille & Feron, 2014). Nevertheless, despite these exceptions, there appears to be a general pattern of reduced exploratory tendencies in a wide range of species as individuals get older (Greenberg & Mettke-Hofmann, 2001; Mata, Wilke, & Czienskowski, 2013; O'Hara et al., 2017). Of course, the capacity of individuals to learn and solve problems is conceptually distinct from the willingness to approach novel stimuli (Griffin & Guez, 2014), but the adage 'you can't teach an old dog new tricks' (or indeed an old parrot, see Loepelt, Shaw, & Burns, 2016) is likely in part a reflection of the relative impassiveness and conservatism of adults (Menzel, 1966).

Why would individuals show an age-related decline in exploratory tendency? Whether the changes are adaptive or not, younger individuals may not have learned that the world is a risky place and may exhibit a degree of naiveté not seen in adults. One could also argue that this age dependency is merely a reflection of some other factor associated with age. In hierarchical societies for instance, younger individuals may have a subordinate role that makes them more exploratory because they do not have the access to resources that more dominant individuals do (Reader & Laland, 2001). Here we consider a simple adaptive explanation for the decline in exploratory tendency that does not invoke any of these mechanisms. Specifically, we argue that information about new things will be of less future value to individuals as they age. Therefore, younger individuals should be more prepared than older individuals to accept risks associated with investigating novel objects.

Intuitively, one might expect that organisms are selected to engage in behaviours that help them gain information early in life (exploration) and then, having learned much of what is to be learned, switch to using this information late in life (exploitation) (see Mata et al., 2013). However, while exploration–exploitation models have been widely invoked to investigate the trade-off between gaining information and using it (Cohen, McClure, & Yu, 2007; Mata & von Helversen, 2015; Reader, 2015), remarkably few models have been developed to understand age-related investment in information acquisition. The model of Eliassen, Jorgensen, Mangel, and Giske (2007) has long served to highlight the general influence of age (and life span) in the exploration–exploitation trade-off. Unfortunately however, their model has limitations when applied to elucidating the adaptive significance of any age-dependent tendency to investigate unfamiliar objects. This is because Eliassen et al. (2007) framed their model in terms of patch foraging and contrasted the performance of learners with nonlearners, rather than asking directly when an agent of given age should explore or ignore a newly encountered object. Berger-Tal, Nathan, Meron, and Saltz (2014) developed a more general exploration–exploitation model that identified four key stages of knowledge acquisition and use over the agent's life span, culminating in exploitation. While elegant, their model was not developed specifically to understand age-dependent changes in the tendency of individuals to explore novel objects.

In this paper we present a tailor-made model that directly identifies the optimal (long-term payoff maximizing) strategy of an agent that encounters an unfamiliar object at a given stage in its life. On first and subsequent encounter(s) with an object, the agent must simply decide whether to take the risk and explore the unfamiliar object's properties (for a potential immediate cost or benefit, as well as information it could use in the future), or ignore it altogether. Assuming that the agents learn about the abundance and profitability of the novel object in a Bayesian manner, we can identify the optimal decision of the agent by combining the agent's current information and expected payoff with the expected future value of information. Formalizing the phenomenon in this way not only allows us to elucidate the agent's optimal exploration–exploitation strategy, but also to isolate the role of two correlated properties of an agent's age: namely the time it has left to live (which affects the future value of any information it gains from investigating) and the time it has lived before encountering the novel object (which affects the agent's estimate of the object's abundance in the environment, and hence the future value of information it gains from investigating).

The primary aim of developing our model was to identify quantitatively the conditions under which the tendency of an agent to explore an unfamiliar object would change with age. Our secondary aim was to use the model to generate testable predictions. Thus, if the observed changes in exploratory tendency can be readily understood from an informational perspective, then the model may also make some additional predictions that can be can be evaluated by experiment. Third, in presenting our model we hope to further highlight the value of exploration–exploitation models (otherwise known as 'bandit models', Press, 2009) to represent and resolve the ubiquitous trade-off between gaining new information and exploiting current information. Greenberg and Mettke-Hofmann (2001) call for a cost–benefit approach to understand adaptive decision making when organisms encounter unfamiliar objects, and this is precisely what we have attempted to present here. Our model is intended to be general, but to help frame it, one may think of the dilemma facing a forager that encounters a novel prey type (or object) for the first time in its life and has to decide whether to investigate it (i.e. 'handle'), or leave it alone.

METHODS

Consider a simple system in which agents have a fixed time horizon of T to make decisions. This period may reflect the life span of the agent or the duration of a foraging season, so long as the individuals remain in the same stage of development (rats, for instance, exhibit developmentally dependent responses to novel stimuli; see Ricceri, Colozza, & Calamandrei, 2000) and the environment is unchanging. At age x ($0 \le x \le T$) in the agent's life, it comes across an object it has never seen before. The agent does not know how common these objects are in its environment, but it does know that this is the first such object it has encountered in its x time units of existence. So, if the agent encounters the object late (or early) in its life, then it is reasonable to infer that the object is rare (or common). As Greenberg and Mettke-Hofmann (2001) note, the primary purpose of exploration of novel objects is the acquisition of information, such as whether the object can be approached without consequence and whether the object provides a resource that can be used now and in the future. We therefore assume that the agent does not know the true probability that the object will be profitable (net benefit b after handling) or unprofitable (net cost c after handling) although it will have prior beliefs about the probability of the novel object being profitable to handle based, for example, on generalizing its previous experiences (Gershman & Niv, 2015). These net benefits and costs take into account the fact that investigating a novel object may take up valuable time and energy that could instead be invested elsewhere (i.e. opportunity costs).

Despite the agent's lack of understanding as to how common the object is and how profitable it is to handle, it still has to decide whether to investigate its properties (i.e. interact with it) or ignore it (move away)! Thus, it has to make a decision based on incomplete information. To identify the optimal behaviour for an agent, we first need to express the uncertainty in a quantitative way, and then identify a means for quantifying the future value of any information the agent may seek. Note that we do not assume any age-dependent changes in the ability of individuals to obtain benefits or avoid costs when they approach these novel objects, since this phenomenon would obfuscate our information-only approach (see Discussion).

Uncertainty about the Commonness of an Unfamiliar Object

Let the agent live from time 0 to time x without seeing a given type of object. Clearly the longer an agent goes without seeing a given type of object, the rarer it should believe it is when it encounters it for the first time. But precisely how rare should that be? If the agent employs its information in the most rational way (i.e. it employs Bayesian inference), then its observations will be combined with its prior beliefs to shape its revised (posterior) beliefs (McNamara, Green, & Olsson, 2006).

Let us therefore assume that naïve agents at age 0 have a Beta(a_p,b_p) hypothesis distribution for the probability p of seeing a particular type of object per time step, with expectation π_p (= a_p/[$a_p + b_p$]). For Bayesian inference to work, we need priors, and these priors will inevitably be shaped by natural selection (Berger-Tal & Avgar, 2012; Gregory, 2006; Stamps & Frankenhuis, 2016). The Beta distribution we have assumed is sufficiently flexible to represent a wide range of prior beliefs. For example, it is possible that agents consider all values of p equally likely (a uniform prior, Beta(1,1)), or agents consider the likelihood of seeing such an object per time step is low (Beta(2,4)) or high (Beta(4,2)) (see Fig. 1a–c). More importantly, the Beta distribution is a conjugate prior of the binomial likelihood model (DeGroot, 1970, pp. 155–189; Kruschke, 2015, page 126). That is, if we seek to estimate the probability of an outcome arising when there are two alternatives and we use the Beta to represent the prior beliefs, then the posterior will also follow a Beta distribution, albeit with different parameters.

Given the conjugacy of the Beta with the binomial, an agent's hypothesis distribution for p will change from Beta(a_p,b_p) to Beta(a_p,$b_p + x$) as its age x increases and it has still not encountered the stimulus (for a proof, see Hobbs & Hooten, 2015, pp. 87–88; for examples of the same rule being used in earlier bandit models, see Green, 1980; Krebs, Kacelnik, & Taylor, 1978; Sherratt, 2011). More generally, if the agent lives to time x and sees y such objects of a given type in this time, then its hypothesis distribution for p will change from Beta(a_p,b_p) with expectation $a_p/(a_p + b_p)$ to Beta($a_p + y$, $b_p + x - y$) with expectation $(a_p + y)/(a_p + b_p + x)$. Extremely old individuals ($x \rightarrow$ infinity) will therefore estimate the likelihood of seeing the novel object per time step simply as y/x.

Uncertainty about the Profitability of the Novel Object

The second source of uncertainty that the agent has to contend with is the probability that the unfamiliar object is profitable (as opposed to unprofitable) to handle. For example, the novel object could provide a net benefit after taking the time to handle it, such as a meal, but it might alternatively pose a risk of injury. The newly encountered object is unfamiliar, so while the agent does not know its properties with any degree of certainty, it may have prior beliefs based on the object's appearance. Clearly, if the agent was confident that the novel object was always unprofitable to handle, then it would not approach it. Likewise, if it was confident that the novel object was always profitable, then it would approach it. If the agent was unsure, then it may help to find out more by investigating, so long as the estimated current and future payoffs from exploring are high enough.

While some types of unfamiliar objects will always be profitable or will always be unprofitable, there are many instances in nature in which the object (such as a prey type) is sometimes profitable to handle (e.g. a mimic) and other times unprofitable (e.g. a model). We therefore assume for generality that the newly encountered object can sometimes be profitable and sometimes unprofitable with unknown probability (in the Supplementary Material we extend our approach to consider multinomial outcomes, for example profitable, unprofitable and neutral to handle).

Given the binomial probabilistic outcome (each time it is handled, the object is either profitable or unprofitable), we can again exploit conjugacy by assuming the agent starts with a Beta(a_q,b_q) prior for the probability q that the novel object is profitable to handle (e.g. it contains a mealworm). If the agent decides to handle the novel object on n occasions, and on r occasions the novel object is profitable (each time with benefit b) while on u occasions the novel object is unprofitable (each time with cost c) such that $r + u = n$, then its posterior belief distribution for the likelihood of a reward should change to Beta($a_q + r$, $b_q + n - r$) with expectation $(a_q + r)/(a_q + b_q + n)$. So the expectation of q for an agent with plenty of opportunity to sample ($n \rightarrow$ infinity) will be r/n (hence independent of its priors, however strong these initial beliefs may be).

There is always information to be gained from handling a relatively unfamiliar object, but is it worth taking the risk that the

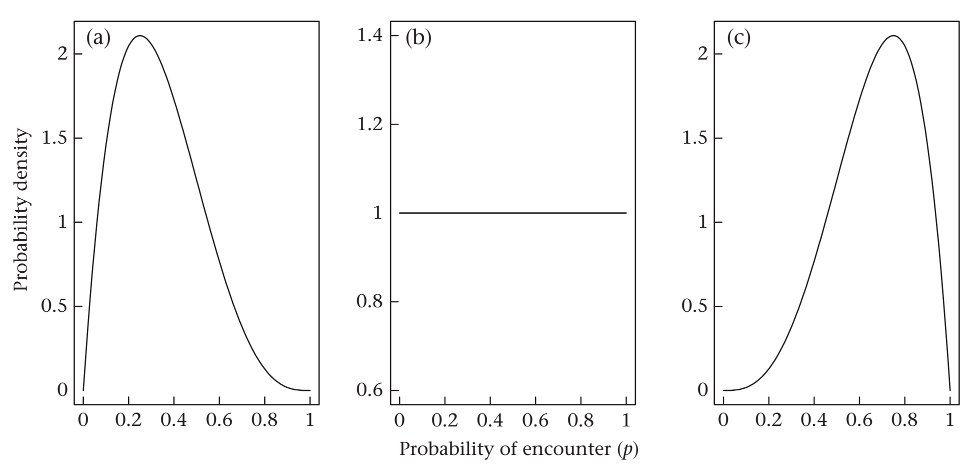

Figure 1. (a–c) Initial hypothesis distributions (priors) as to the probability (p) of an agent age 0 encountering a novel object of a particular type per time step: (a) pessimistic (Beta(2,4)); (b) agnostic (Beta(1,1)); (c) optimistic (Beta(4,2)). Similar initial hypothesis distributions can be used to reflect the probability (q) that the novel object is profitable (as opposed to unprofitable) to handle.

object is unprofitable? To answer this we must combine uncertainty over how common the unfamiliar object is (which will affect the anticipated likelihood of seeing it in the future, hence the future value of any information it gains) and uncertainty as to how profitable the object is (which will affect its immediate rewards, but also its long-term rewards if it continues to use its available information in the optimal way). This requires the agent look to the future and identify the decision (explore or ignore) given how old it is (x), how many objects of the given type it has encountered before (y) ($y \leq x$), how many times an object of this type was handled in the past (n) ($n \leq y$) and how many times these unfamiliar objects proved to be profitable (r) ($r \leq n$).

The standard method for evaluating the optimal state-dependent (state variables x, y, r and n) decision in a multistage decision process where the decisions taken at each stage affect the states under which future decisions have to be made is dynamic programming (Clark & Mangel, 2000; Houston, Clark, McNamara, & Mangel, 1988; Jones, 1978; Mangel & Clark, 1988). In the next section we identify the dynamic programming equation (DPE) that must be solved to identify the optimal decision for any agent with initial priors Beta(a_p,b_p) and Beta(a_q,b_q) and given informational state variables (x, y, r, n). In the results that follow the presentation of the DPE, we characterize the key properties of this optimal sampling solution.

Combining the Uncertainty to Arrive at a Decision: the DPE

Let us define $S(x, y, r, n)$ as the maximum future payoff to an agent at the start of age x given state variables (x, y, r, n) assuming that it continues to adopt the optimal decision rules as it learns more about how common the novel objects are, and potentially (if it decides to handle them) how frequently they are profitable. Collectively, x, y, r and n allow the agent to estimate the expected probability of encountering the novel object per time step ($\pi_p = (a_p + y)/(a_p + b_p + x)$) and the expected probability that the novel object is profitable to handle ($\pi_q = (a_q + r)/(a_q + b_q + n)$) based on its priors and what it has learned from its environment. These expectations will change over time as information is gathered. Should it encounter the initially novel object in the time step it is about to take, then it has to decide whether to explore or ignore. Let $S_E(x, y, r, n)$ and $S_I(x, y, r, n)$ be the expected maximum future payoff from exploring (E) or ignoring (I) a novel object it has just encountered based on its current informational variables x, y, r and n at the start of the time step. Naturally, the agent should choose to explore if $S_E(x, y, r, n) > S_I(x, y, r, n)$ and to ignore if the reverse inequality holds (in the unlikely event of a tie, we can assume, entirely arbitrarily, that it chooses to ignore the novel object).

How can we calculate its likely future rewards S? Given the way S is defined, we can begin by recognizing that:

$$S(x, y, r, n) = (1 - \pi_p)\, S(x + 1, y, r, n) + \pi_p\{\max[S_I(x, y, r, n), S_E(x, y, r, n)]\} \quad (1)$$

Thus, either the agent does not encounter the unfamiliar object in the current time step (in which case, it just gets older with no change in its other state variables), or it does encounter the unfamiliar object and it chooses the payoff-maximizing decision (Explore or Ignore). The former event arises with estimated probability $(1 - \pi_p)$ and the latter arises with estimated probability π_p. Ignoring an object it has just encountered would leave the agent one time unit older, and aware that it has just seen the unfamiliar object but otherwise uninformed as to its likelihood of being profitable:

$$S_I(x, y, r, n) = S(x + 1, y + 1, r, n) \quad (2)$$

By contrast, exploring the unfamiliar object would leave the agent one time unit older and better informed for the future, while subject to immediate gains or losses, which arise with estimated probabilities π_q and $(1 - \pi_q)$ respectively:

$$S_E(x, y, r, n) = \pi_q\, (S(x + 1, y + 1, r + 1, n + 1) + b) + (1 - \pi_q)(S(x + 1, y + 1, r, n + 1) - c) \quad (3)$$

Substituting equations (3) and (2) into equation (1) yields the DPE, which is all we need to know to identify the agent's optimal decision under any conditions. The DPE can be solved numerically via backward induction, assuming that the agents have a fixed time horizon of T. Under these conditions we know that $S(x = T, y, r, n) = 0$ for all combinations of y, r and n because there is no future to consider. We can then work backwards, starting with $x = T - 1$ and evaluate the optimal decision under all possible combinations of y, r and n since it involves just a single time step. Following this we can move to $x = T - 2$, using the DPE to identify the maximum expected future payoff based on the immediate rewards and payoffs from time $T - 1$ and so on.

RESULTS

The response variable of interest here is the tendency of an agent to engage in exploratory behaviour at a given stage in its life (age x) when it encounters an object for the first time. For this reason, the model predictions are entirely independent of the actual probability that the object is profitable to handle. One might wonder why one needs to consider the future at all, if we only wish to understand the agent's initial reluctance to sample. The answer is that it is an integral part of the decision-making process for the agent – it has to consider the likelihood of seeing this type of unfamiliar object again as well as the payoff if it makes the optimal decisions now and in the future. Thus, exploration not only yields immediate costs or benefits, but also information that the agent can use for the rest of its life. If there was no future to think about, we would have a familiar one-off gambit (analogous to Pascal's wager), based simply on the expected benefits and costs multiplied by their likelihoods (Foster & Kokko, 2009).

Let us begin by assuming the agent starts out with a uniform prior for q (so that the expected probability of the object being profitable to handle per interaction is 0.5). Whenever the expectation for q is 0.5 ($= \pi_q$) and the anticipated benefit to cost (b/c) ratio of handling the unfamiliar object exceeds 1, then the agent should always explore the properties of a novel object simply because there is an estimated immediate benefit to interacting with the object (i.e. $b\ \pi_q > (1 - \pi_q)\ c$), above and beyond the additional information it will gain. Nevertheless even when ($b/c < 1$) such that $b\ \pi_q < (1 - \pi_q)\ c$, then it may still pay agents (of appropriate age) to investigate the object, because they could be wrong about the true value of q (π_q is simply an estimate) and the improved estimate of q they gain from sampling may therefore have some future value. The estimated future value of information is higher when the agent is young for two inter-related reasons (partitioned below). First, it will tend to consider the novel object relatively common (since it has encountered it early in life), so the object may be something it will frequently encounter in the future. Second, with a long time left to live, it will have more opportunity to use any information it gains. Under these conditions younger individuals require a lower b/c ratio threshold to justify exploration and this is precisely what we see (Fig. 2a).

The prior beliefs of the agents shape their decisions in predictable ways. For example, if the agents initially feel pessimistic in that they believe they are relatively unlikely to see the novel object ($a_p = 2$, $b_p = 4$) and the novel object is likely to be unprofitable

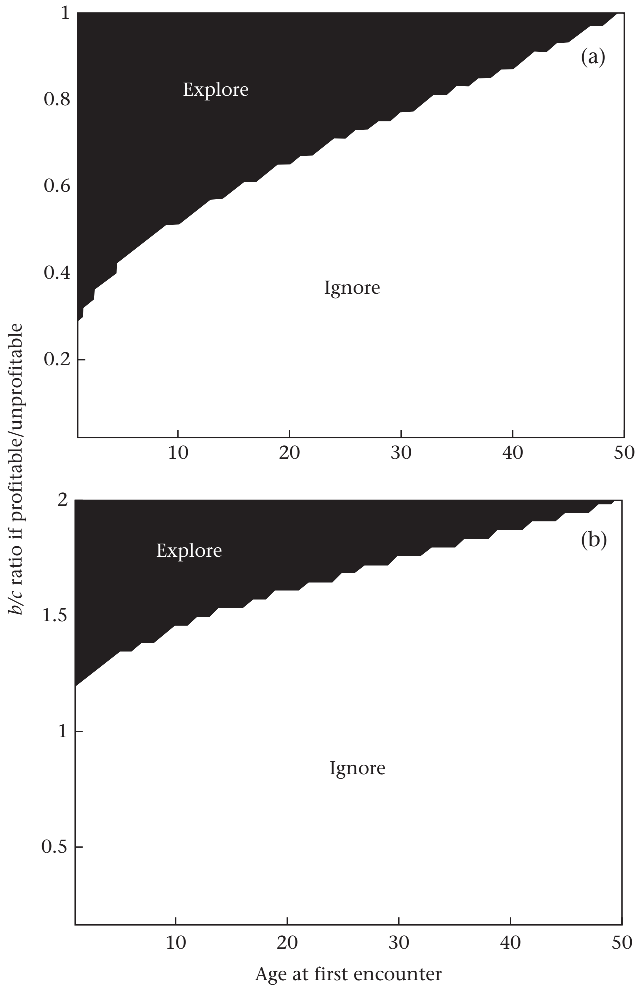

Figure 2. The optimal decision (Explore or Ignore) of an agent that encounters an object at a given age for the first time. Older individuals should be more conservative and ignore the novel object because the future value of the information is judged to be low on the basis that (1) the object is perceived to be rare and (2) the individual has a relatively short time to live. Only when the benefit-to-cost ratio (b/c) is high, should older agents be prepared to sample. (a) Agents have agnostic priors for encountering a novel object ($a_P = b_p = 1$) and for the probability of the object being profitable if handled ($a_q = b_q = 1$). (b) Agents begin with more pessimistic priors for the probability of encountering a novel object ($a_p = 2$, $b_p = 4$) and for the probability of the object being profitable if handled ($a_q = 2$, $b_q = 4$). Note rescaled graph: any agent whatever its age should decide to explore an object it has encountered for the first time if $b/c > (1 - \pi_q)/\pi_q$, which explains the upper b/c 'asymptote' of 1 when $\pi_q = ½$ (a), and of 2 when $\pi_q = 1/3$ (b). Additional model parameter: $T = 50$.

($a_q = 2$, $b_q = 4$), then all individuals (whatever their age) are expected to be more conservative. However, despite the fact that a higher relative benefit (b/c) is required for even younger individuals to risk sampling (see Fig. 2b, note rescaled Y axis), there is still a clear effect of age in that older individuals that encounter a novel object late in life require a higher b/c to justify sampling. Supplementary Fig. S1a, b shows similar effects when the agent is only pessimistic about its rate of encounter with the unfamiliar object, or only pessimistic about the likelihood of it being profitable. As might be expected, pessimism as to the likelihood of the unfamiliar object being profitable does much more to reduce the tendency of agents to explore an unfamiliar object than pessimism regarding the likelihood of encountering it. This is because the agent is constantly updating its beliefs about its probability of encountering the object. By contrast, it will only update its beliefs about the probability of the object being profitable after it handles it, so its initial beliefs will hold until its first interaction with the object.

How can we distinguish the effect of age (time x) from the effects of the time to live (time $T - x$)? While challenging to disentangle experimentally, this is entirely possible from a modelling perspective. Thus, if prior experience (specifically, a lack of previous encounters with the object) alone mattered, then the critical age ($= x^*$) at which agents switch from exploring to ignoring novel objects should be independent of the agent's time horizon T so long as $T > x^*$. Likewise, if only the future matters, then the residual time span at the optimal switch point x^* ($T - x^* = \delta^*$) should be independent of T. If both the past experience and the future value of information matters, then we would expect neither x^* nor δ^* to be constant as T is varied (see Fig. 3).

Here we show that both the agent's past experience and the future value of information influence its optimal decision. Fig. 4 shows the critical age (x^*) above which individual agents should simply ignore a novel object (and below which they should explore) for a range of time horizons (T) under four different b/c ratios. Intriguingly, the critical age at which adults switch from being prepared to explore, to ignoring a newly encountered unfamiliar object increases with the total life span (T) of the individual. The longer an agent has left to live at a given age x, the more opportunity it will have to use its information in the future, so it will be more likely to explore. So, there is no single threshold critical age (x^*) beyond which agents should cease exploring, and the decision to explore or ignore an unfamiliar object is not simply a reflection of the agent's previous lack of encounter with the unfamiliar object.

While the optimal switch point x^* increases approximately linearly with T, the gradient of this relationship is less than 1. With an intercept of 0 and a gradient less than 1, the amount of time an agent has left to live (δ^*) when it switches from exploring to ignoring novel objects also increases with its life span (see Supplementary Fig. S2). In this way, the optimal rule for switching from exploration to ignoring an unfamiliar object is not simply based on the time left for the agent to use its information either, because it depends in part on the estimate of the frequency at which this information will be useful. We can see this most clearly when we set the agent's belief that it will see a novel object as a constant (0.5 say) entirely independent of its age and experiences

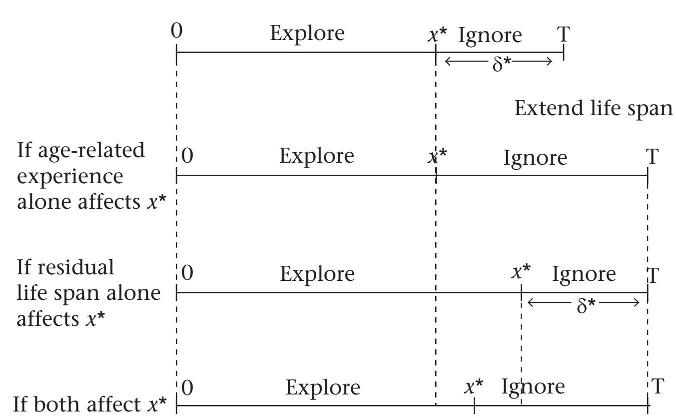

Figure 3. A schematic diagram showing the optimal critical age (x^*) beyond which an agent should choose to ignore an object it has encountered for the first time. If the position of x^* was entirely dependent on the agent's experience it has accumulated up to that time, then x^* should not change if the agent's time horizon T is increased. Likewise if the position of x^* is dependent simply on reaching a given critical time left to live (δ^*), then the agent should switch to ignoring a novel object when its residual life span reaches δ^*, whatever its longevity. If the optimal position of x^* is influenced by previous experience and by likely future experiences, then neither x^* nor δ^* will be constant as T is varied.

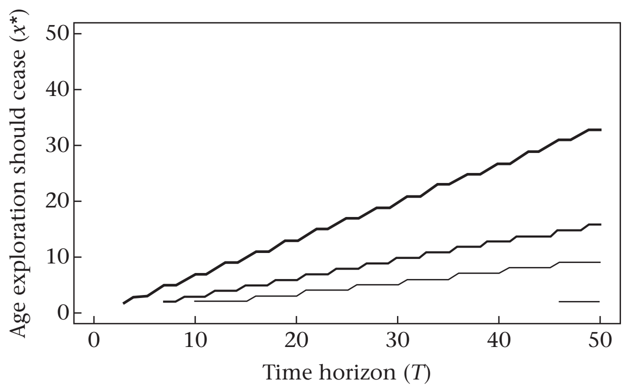

Figure 4. The critical age (x^*) beyond which an agent should choose to ignore an object it has just encountered for the first time, for agents that live for a particular time horizon T. Separate lines show different b/c ratios ($b/c = 0.8$, 0.6, 0.5, 0.3 portrayed in lines of diminishing thickness). When b/c is low and the time horizon is short, then agents will not be selected to explore in the first place, so no critical age is given. The critical age for ceasing exploration increases with the b/c ratio. Given that the gradient between x^* and T is less than 1, both x^* and residual life span at this switch point ($\delta^* = T - x^*$) increase as the life span of the agent increases.

(so the agent does not use Bayesian learning to update its beliefs regarding the distribution of p). Only then does the optimal switch rule result in a fixed residual life span (δ^*) for all values of T and a given ratio of b/c. The residual life span on switching to ignoring novel objects is fixed in this case because estimates of the future likelihood of encountering the object per time step are now independent of the age at which the agent first encountered it. Instead, the only relevance of age in influencing the optimal decision derives from the time left before the agent reaches T.

DISCUSSION

Here we have asked how an agent should react to a novel object, when the likelihood that the object is profitable to the agent is unknown (although it will have its beliefs) and the frequency of the object in the environment is also unknown (although the agent will have its beliefs, based in part on the age at which it first encountered it). Formalizing the conundrum as an exploration–exploitation problem, we predict that younger individuals should be more liable to investigate an unfamiliar object than older individuals for two inter-related reasons: (1) the earlier the object is encountered in the agent's life, the more common it is perceived to be and (2) the earlier the object is encountered in the agent's life, the longer the agent has to exploit this knowledge in the future. These explanations are not mutually exclusive, and indeed we show that both may play a role in influencing the relationship between exploratory tendency and age. We also show that cost–benefit considerations can sometimes result in all agents choosing to explore a novel object independent of their age (if the object is considered to be sufficiently profitable), or no agents choosing to explore the object (if the object is considered sufficiently unprofitable). In this way, the nature of the response of organisms to novel objects is likely to be shaped by a range of ecological factors (such as whether individuals live on predator-free islands, etc.) that each affect the nature of the risks involved (Mettke-Hofmann, Winkler, & Leisler, 2002; O'Hara et al., 2017).

Although we present our model in terms of foraging, it should be applicable to a broad range of situations, including predator inspection and the interactions of individuals with strangers. Indeed, studies of social interactions in humans indicate that there is a gradual decline in the tendency of individuals to establish relationships with unfamiliar people as they age. Carstensen's socioemotional selectivity theory (Carstensen, 1995) attempts to explain the phenomenon from an adaptive perspective, proposing that as humans age they have less to gain from any newly formed relationships in the future. In addition, an individual may be less inclined to get to know an acquaintance if they feel that they are unlikely to see them again. Our model formalizes these two basic arguments, where the age-related decline in the tendency to try new things is shaped by both time left to exploit knowledge and the agent's estimate of the frequency of future contact.

While the tendency of individuals to approach novel objects will vary among species and with age within species (see Crane & Ferrari, 2017, for a review of studies on predator neophobia), there are of course numerous factors that can shape an individual's beliefs and subsequent decisions. These factors include the individual's personal experiences that could act to obscure any age-related trend. For example, an agent that encountered an object similar to one that it has just encountered may consider the profitability of the newly encountered object rather differently from conspecifics that did not have this experience. Likewise, an agent that observes others getting food from the unfamiliar object will naturally be more inclined to approach the object itself. Both generalization and social learning are readily built into exploration–exploitation models by allowing the prior beliefs to be shaped by factors other than the agent's direct experience of the object in question (see for example Abbott & Sherratt, 2011), but as with all models, simplifications are required to maintain transparency and tractability.

Our model simply focuses on the tendency of payoff-maximizing young and old individuals to approach and investigate novel objects that may be profitable or unprofitable. We use this model to argue that the widely observed age-related decrease in exploratory tendency is likely to be adaptive. This change in behaviour could manifest itself as a gradual increase in the agent's fear of novel objects (neophobia), although it is worth emphasizing that in our model it is simply the value of information to the agent that alters with its age, not the perceived risk of the novel object on approach. One way to identify the underlying mechanism is to determine the physiological responses of individuals of different age towards novel objects. If the primary reason for an increase in conservatism with age is the devaluation of information, then there should not be an age-dependent change in physiological response; however, if the perceived risks of harm are higher in older individuals, then novel objects should elicit a heightened fear response in these individuals. The same sort of age-dependent changes in exploratory tendency could arise from a gradually decreased attraction to novel objects (neophilia). Operationally, neophilia might be characterized by a positive attraction towards a novel item of unknown net profitability (positive or negative) compared to a more familiar object of known (and positive) net profitability. Our model is currently presented in terms of a two-arm bandit with one arm (= approaching the novel object) providing an uncertain net reward (positive or negative) and the other arm (= ignoring the novel object) providing a known reward of 0. However, it would be straightforward to cast it as a two-arm bandit in which the arm with the known reward provides a guaranteed payoff that is higher than 0 (Jones, 1978). This would simply make all individuals more conservative, but despite moving the b/c thresholds, the underlying age effect would remain. Two-arm bandit models involving (initially) unknown rewards in both arms have previously been invoked to understand problem solving of birds faced with binary choices that yield different rates of reward (Krebs et al., 1978; Logan, 2016), although age was not considered.

Like many discrete-time dynamic programming models (Clark & Mangel, 2000; Mangel & Clark, 1988), we have assumed that agents make decisions within a fixed time horizon T and we have argued that it is often optimal for them to switch behavioural strategies during this time period. In many cases, while an individual may have some awareness of its age, it will rarely know T and hence its residual operational life span, with certainty. Nevertheless, this will not prevent selection acting on the timing of any given behavioural change. After all, predictable environmental cues and/or endogenous circannual clocks can shape the timing of just about everything from flowering to migration, without the organisms understanding why these changes arise.

The current model assumes that each time the novel object is handled, it is either profitable or unprofitable. However, the model is readily extended to allow objects to have more than two properties, such as profitable, unprofitable and neutral (see Supplementary Material). Computationally, the task is made easier if one employs the Dirichlet distribution to represent the prior (see Supplementary Fig. S3), since it is the conjugate of the multinomial likelihood model (Aubier, Joron, & Sherratt, 2017). As we show (see Supplementary Fig. S4a, b), this more general model yields the same qualitative results with regard to the age-related tendency to explore. We have also assumed for tractability that b and c are known. However, it is likely that an agent's behavioural responses and/or the priors that generate them, have been at least in part shaped by selection in the agents' ancestral environment (Berger-Tal & Avgar, 2012; Gregory, 2006; Stamps & Frankenhuis, 2016).

Perhaps most importantly, it is entirely conceivable that b and c change throughout the agent's life. For example, individuals may get better at securing benefits (or avoiding costs) the more times they handle the object. Alternatively or in addition, as an animal ages, it may get to know where the reliable food sources are, so any new information about food (say) may be intrinsically less valuable to it. More generally, although we have assumed that all agents make decisions over a fixed time horizon, in reality there may be substantial changes in the nature of selection on individuals as they age (Baltes, 1997). In particular, if handling costs of unprofitable objects were paid in terms of survival, then younger individuals would necessarily pay higher costs from exploring new objects than older individuals because they are risking more of their reproductive lives. In these cases, the ‘asset protection principle’ (Clark, 1994) would apply. In our Supplementary Material we have recast our model to incorporate a chance of death from handling an unprofitable object. When individuals place their entire future reproductive output at risk by handling an unprofitable object, then there is a range of parameter values in which only young and old individuals are exploratory (Supplementary Fig. S5a, b), comparable to the nonlinear responses reported by Lafaille and Feron (2014). Here young individuals are exploratory because they have much to gain from the information, but old individuals are also exploratory because they have little to lose in terms of future reproductive output.

Allowing b and c to vary with age would almost certainly make the model more realistic, and (as seen above) it could take predictions in almost any direction. Note, however, that we have not needed to invoke these effects to explain the observed age-related changes in exploratory tendency, showing that variation in the value of information (Stephens, 2007) is entirely sufficient. Quantifying the functional form of age-dependent changes in b and c is likely to be challenging. However, evaluating how individuals respond to familiar profitable and unprofitable objects, such as common palatable and unpalatable prey, would reveal how the costs and benefits of exploiting known resources change over time.

Our information-centred model makes several predictions that appear to match observations reasonably well. In particular, the model predicts that if agents encounter several novel items of a given type at one time, then they should be more prepared to investigate them, whatever their age, because the benefits of acquiring this information will be higher (see also Sherratt, 2011). Several experiments have reported that predators consume more unfamiliar unpalatable prey before ceasing the higher the density of prey (Beatty, Beirinckx, & Sherratt, 2004; Greenwood, Cotton, & Wilson, 1989; Lindstrom, Alatalo, Lyytinen, & Mappes, 2001; Rowland, Wiley, Ruxton, Mappes, & Speed, 2010). Likewise, if the novel object is placed out every day (as opposed to once a week, say), then all agents will increase their estimate of the probability of seeing the object per time step, and the model predicts that the initial reluctance of older adults can be overcome through repeated presentation. In fact, this type of behavioural response to novel prey is often assumed (rather than predicted) in foraging models (e.g. Puurtinen & Kaitala, 2006; Speed, 2001), while experiments by Marples, Quinlan, Thomas, and Kelly (2007) showed that dietary wariness of domestic chicks could indeed be inactivated with increasing exposure. Skelhorn, Holmes, Hossie, and Sherratt (2016) similarly found that domestic chicks reduce their latency to attack artificial imperfect snake mimics over 3 days of repeated exposure. Although the prediction that neophobia should decrease after repeated exposures is somewhat intuitive (and may even explain the lower predator neophobia in captive species; Crane & Ferrari, 2017), it strongly suggests at least some form of informational cost–benefit trade-off. Put simply, if the object appears common (and has not yet proved harmful), it will be worth checking out.

Our model also predicts that individuals or species that live a long time should be prepared to explore until a more advanced age, all else being equal. Manipulating the perceived time horizon of agents might be one way to test this prediction. Intriguingly, events that shorten the perceived time horizon of young humans (such being found positive for HIV) tends to alter their social preferences, making them more similar to those of elders (reviewed in Carstensen, 1995). However, care should be exercised when artificially shortening life expectancy because the immediate benefits and costs of exploratory behaviour may concomitantly change. For example, wild adult great tits with artificially reduced life expectancy (following experimental manipulation of the sex ratio and size of their broods; Nicolaus et al., 2012a) were found to explore novel environments more rapidly than control adults (Nicolaus et al., 2012b). As the authors suggest, this result may well have arisen because hungry or otherwise desperate individuals have more to immediately gain (or less to immediately lose) from taking risks (see for example Sherratt, 2003; Wolf, van Doorn, Leimar, & Weissing, 2007). While there will no doubt be multiple confounding factors, it may also ultimately be possible to evaluate the relationship between longevity and exploration across species through comparative analyses.

Our work parallels research on observed reductions in phenotypic plasticity over time, and many models invoked to explain this phenomenon also assume a form of Bayesian updating (English, Fawcett, Higginson, Trimmer, & Uller, 2016; Fawcett & Frankenhuis, 2015; Stamps & Frankenhuis, 2016; Stamps & Krishnan, 2017). Nevertheless, in the above models the agents tend to be considered relatively passive recipients of external information about the state of their environments, but here we focus on the dilemma of whether agents should actively seek further information or exploit their current information. Exploration–exploitation models are tailor made for addressing circumstances where there is a potential trade-off between gaining more information and exploiting current information. Collectively, they serve to emphasize that the decision to approach a given novel object, conspecific or microhabitat (or even purchase an unfamiliar type of wine in the supermarket) does not simply depend on the anticipated immediate reward associated with the object, but also the future rewards that could be reaped with this new-found information.

Acknowledgments

We thank Thomas Aubier, Tal Avgar and two anonymous referees for some very helpful comments on the manuscript. T.N.S. and J.M.-F. are supported by Discovery Grants from the Natural Sciences and Engineering Research Council of Canada.

Supplementary Material

Supplementary material related to this article can be found at https://doi.org/10.1016/j.anbehav.2018.01.025.

References

Abbott, K. R., & Sherratt, T. N. (2011). The evolution of superstition through optimal use of incomplete information. *Animal Behaviour, 82*, 85–92. https://doi.org/10.1016/j.anbehav.2011.04.002.

Aubier, T. G., Joron, M., & Sherratt, T. N. (2017). Mimicry among unequally defended prey should be mutualistic when predators sample optimally. *American Naturalist, 189*(3), 267–282. https://doi.org/10.1086/690121.

Baltes, P. B. (1997). On the incomplete architecture of human ontogeny. *American Psychologist, 52*, 366–380.

Beatty, C. D., Beirinckx, K., & Sherratt, T. N. (2004). The evolution of müllerian mimicry in multispecies communities. *Nature, 431*(7004), 63–67. https://doi.org/10.1038/nature02818.

Benson-Amram, S., & Holekamp, K. E. (2012). Innovative problem solving by wild spotted hyenas. *Proceedings of the Royal Society B: Biological Sciences, 279*(1744), 4087–4095. https://doi.org/10.1098/rspb.2012.1450.

Benson-Amram, S., Weldele, M. L., & Holekamp, K. E. (2013). A comparison of innovative problem-solving abilities between wild and captive spotted hyaenas, *Crocuta crocuta*. *Animal Behaviour, 85*, 349–356. https://doi.org/10.1016/j.anbehav.2012.11.003.

Berger-Tal, O., & Avgar, T. (2012). The glass is half-full: Overestimating the quality of a novel environment is advantageous. *PLoS One, 7*(4), e34578. https://doi.org/10.1371/journal.pone.0034578.

Berger-Tal, O., Nathan, J., Meron, E., & Saltz, D. (2014). The exploration–exploitation dilemma: A multidisciplinary framework. *PLoS One, 9*(4), e95693. https://doi.org/10.1371/journal.pone.0095693.

Bergman, T. J., & Kitchen, D. M. (2009). Comparing responses to novel objects in wild baboons (*Papio ursinus*) and geladas (*Theropithecus gelada*). *Animal Cognition, 12*(1), 63–73. https://doi.org/10.1007/s10071-008-0171-2.

Biondi, L. M., Bo, M. S., & Vassallo, A. I. (2010). Inter-individual and age differences in exploration, neophobia and problem-solving ability in a Neotropical raptor (*Milvago chimango*). *Animal Cognition, 13*(5), 701–710. https://doi.org/10.1007/s10071-010-0319-8.

Biondi, L. M., Guido, J. M., Bo, M. S., Muzio, R. N., & Vassallo, A. I. (2015). The role of stimulus complexity, age and experience in the expression of exploratory behaviour in the chimango caracara, *Milvago chimango*. *Animal Cognition, 18*(1), 139–150. https://doi.org/10.1007/s10071-014-0785-5.

Biondi, L. M., Guido, J., Madrid, E., Bo, M. S., & Vassallo, A. I. (2013). The effect of age and sex on object exploration and manipulative behavior in a Neotropical raptor, the chimango caracara, *Milvago chimango*. *Ethology, 119*(3), 221–232. https://doi.org/10.1111/eth.12056.

Carstensen, L. L. (1995). Evidence for a life-span theory of socioemotional selectivity. *Current Directions in Psychological Science, 4*, 151–156.

Clark, C. W. (1994). Anti-predator behaviour and the asset protection principle. *Behavioral Ecology, 5*, 159–170.

Clark, C. W., & Mangel, M. (2000). *Dynamic state variable models in ecology*. Oxford, U.K.: Oxford University Press.

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences, 362*(1481), 933–942. https://doi.org/10.1098/rstb.2007.2098.

Costa, P. T., Jr., McCrae, R. R., Zondeerman, A. B., Barbano, H. E., Lebowitz, B., & Larson, D. M. (1986). Cross-sectional studies of personality in a national sample: 2. Stability in neuroticism, extraversion, and openness. *Psychology and Aging, 1*, 144–149.

Crane, A. L., & Ferrari, M. C. O. (2017). Patterns of predator neophobia: A meta-analytic review. *Proceedings of the Royal Society B: Biological Sciences, 284*, 20170583.

DeGroot, M. H. (1970). *Optimal statistical decisions*. Hoboken, NJ: J. Wiley.

Donnellan, M. B., & Lucas, R. E. (2008). Age differences in the big five across the life span: Evidence from two national samples. *Psychology and Aging, 23*, 558–566.

Eliassen, S., Jorgensen, C., Mangel, M., & Giske, J. (2007). Exploration or exploitation: Life expectancy changes the value of learning in foraging strategies. *Oikos, 116*(3), 513–523.

English, S., Fawcett, T. W., Higginson, A. D., Trimmer, P. C., & Uller, T. (2016). Adaptive use of information during growth can explain long-term effects of early life experiences. *American Naturalist, 187*(5), 620–632. https://doi.org/10.1086/685644.

Fawcett, T. W., & Frankenhuis, W. E. (2015). Adaptive explanations for sensitive windows in development. *Frontiers in Zoology, 12*(Suppl. 1), S3. https://doi.org/10.1186/1742-9994-12-s1-s3.

Foster, K. R., & Kokko, H. (2009). The evolution of superstitious and superstition-like behaviour. *Proceedings of the Royal Society B: Biological Sciences, 276*(1654), 31–37. https://doi.org/10.1098/rspb.2008.0981.

Gershman, S. J., & Niv, Y. (2015). Novelty and inductive generalization in human reinforcement learning. *Topics in Cognitive Science, 7*(3), 391–415. https://doi.org/10.1111/tops.12138.

Green, R. F. (1980). Bayesian birds: A simple example of Oaten's stochastic model of optimal foraging. *Theoretical Population Biology, 18*, 244–256.

Greenberg, R., & Mettke-Hofmann, C. (2001). Ecological aspects of neophobia and neophilia in birds. In V. Nolan, & C. F. Thompson (Eds.), *Current ornithology* (pp. 119–178). Boston, MA: Springer.

Greenwood, J. J. D., Cotton, P. A., & Wilson, D. M. (1989). Frequency-dependent selection on aposematic prey: Some experiments. *Biological Journal of the Linnean Society, 36*(1–2), 213–226.

Greggor, A. L., Thornton, A., & Clayton, N. S. (2015). Neophobia is not only avoidance: Improving neophobia tests by combining cognition and ecology. *Current Opinion in Behavioral Sciences, 6*, 82–89. https://doi.org/10.1016/j.cobeha.2015.10.007.

Gregory, R. L. (2006). Editorial essay. Bayes window (3): Where do prior probabilities come from? *Perception, 35*, 289–290.

Griffin, A. S., & Guez, D. (2014). Innovation and problem solving: A review of common mechanisms. *Behavioural Processes, 109*, 121–134. https://doi.org/10.1016/j.beproc.2014.08.027.

Herborn, K. A., Macleod, R., Miles, W. T. S., Schofield, A. N. B., Alexander, L., & Arnold, K. E. (2010). Personality in captivity reflects personality in the wild. *Animal Behaviour, 79*, 835–843. https://doi.org/10.1016/j.anbehav.2009.12.026.

Hobbs, N. T., & Hooten, M. B. (2015). *Bayesian models: A statistical primer for ecologists*. Princeton, NJ: Princeton University Press.

Hopper, L. M., Price, S. A., Freeman, H. D., Lambeth, S. P., Schapiro, S. J., & Kendal, R. L. (2014). Influence of personality, age, sex, and estrous state on chimpanzee problem-solving success. *Animal Cognition, 17*(4), 835–847. https://doi.org/10.1007/s10071-013-0715-y.

Houston, A., Clark, C., McNamara, J., & Mangel, M. (1988). Dynamic models in behavioural and evolutionary ecology. *Nature, 332*(6159), 29–34.

Jones, P. W. (1978). On the two-armed bandit with one probability known. *Metrika, 25*, 235–239.

Joyal, C. C., Beaudin, S., & Lalonde, R. (2000). Longitudinal age-related changes in motor activities and spatial orientation in CD-1 mice. *Archives of Physiology and Biochemistry, 108*, 248–256.

Kendal, R. L., Coe, R. L., & Laland, K. N. (2005). Age differences in neophilia, exploration, and innovation in family groups of callitrichid monkeys. *American Journal of Primatology, 66*, 167–188.

Krebs, J. R., Kacelnik, A., & Taylor, P. (1978). Test of optimal sampling by foraging great tits. *Nature, 275*, 27–31.

Kruschke, J. K. (2015). *Doing Bayesian data analysis: A tutorial with R, JAGS and Stan* (2nd ed.). London, U.K: Elsevier.

Lafaille, M., & Feron, C. (2014). U-shaped relationship between ageing and risk-taking behaviour in a wild-type rodent. *Animal Behaviour, 97*, 45–52. https://doi.org/10.1016/j.anbehav.2014.08.018.

Lamberty, Y., & Gower, A. J. (1990). Age-related changes in spontaneous behavior and learning in NMRI mice from maturity to middle age. *Physiology & Behavior, 37*, 1137–1144.

Lindstrom, L., Alatalo, R. V., Lyytinen, A., & Mappes, J. (2001). Strong antiapostatic selection against novel rare aposematic prey. *Proceedings of the National Academy of Sciences of the United States of America, 98*(16), 9181–9184.

Loepelt, J., Shaw, R. C., & Burns, K. C. (2016). Can you teach an old parrot new tricks? Cognitive development in wild kaka (*Nestor meridionalis*). *Proceedings of the Royal Society B: Biological Sciences, 283*(1832), 20153056. https://doi.org/10.1098/rspb.2015.3056.

Logan, C. J. (2016). Behavioral flexibility and problem solving in an invasive bird. *PeerJ, 4*, e1975. https://doi.org/10.7717/peerj.1975.

Mangel, M., & Clark, C. W. (1988). *Dynamic modeling in behavioral ecology*. Princeton, NJ: Princeton University Press.

Marples, N. M., Quinlan, M., Thomas, R. J., & Kelly, D. J. (2007). Deactivation of dietary wariness through experience of novel food. *Behavioral Ecology, 18*, 803–810.

Mata, R., & von Helversen, B. (2015). Search and the aging mind: The promise and limits of the cognitive control hypothesis of age differences in search. *Topics in Cognitive Science, 7*(3), 416–427. https://doi.org/10.1111/tops.12139.

Mata, R., Wilke, A., & Czienskowski, U. (2013). Foraging across the life span: Is there a reduction in exploration with aging? *Frontiers in Neuroscience, 7*(53), 1–7. https://doi.org/10.3389/fnins.2013.00053.

McNamara, J. M., Green, R. F., & Olsson, O. (2006). Bayes' theorem and its applications in animal behaviour. *Oikos, 112*, 243–251.

Menzel, E. W. (1966). Responsiveness to objects in free-ranging Japanese monkeys. *Behaviour, 26*(1/2), 130–150.

Mettke-Hofmann, C., Winkler, H., & Leisler, B. (2002). The significance of ecological factors for exploration and neophobia in parrots. *Ethology, 108*(3), 249–272. https://doi.org/10.1046/j.1439-0310.2002.00773.x.

Morand-Ferron, J., Cole, E. F., Rawles, J. E. C., & Quinn, J. L. (2011). Who are the innovators? A field experiment with 2 passerine species. *Behavioral Ecology, 22*(6), 1241–1248. https://doi.org/10.1093/beheco/arr120.

Morand-Ferron, J., Hamblin, S., Cole, E. F., Aplin, L. M., & Quinn, J. L. (2015). Taking the operant paradigm into the field: Associative learning in wild great tits. *PLoS One, 10*(8). https://doi.org/10.1371/journal.pone.0133821.

Nicolaus, M., Michler, S. P., Ubels, R., van der Velde, M., Bouwman, K. M., Both, C., et al. (2012a). Local sex ratio affects the cost of reproduction. *Journal of Animal Ecology, 81*, 564–572.

Nicolaus, M., Tinbergen, J. M., Bouwman, K. M., Michler, S. P., Ubels, R., Both, C., et al. (2012b). Experimental evidence for adaptive personalities in a wild passerine bird. *Proceedings of the Royal Society B: Biological Sciences, 279*, 4885–4892.

O'Hara, M., Mioduszewska, B., von Bayern, A., Auersperg, A., Bugnyar, T., Wilkinson, A., et al. (2017). The temporal dependence of exploration on neotic style in birds. *Scientific Reports, 7*(1), 4742. https://doi.org/10.1038/s41598-017-04751-0.

Press, W. H. (2009). Bandit solutions provide unified ethical models for randomized clinical trials and comparative effectiveness research. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 22387–22392.

Reader, S. M. (2015). Causes of individual differences in animal exploration and search. *Topics in Cognitive Science, 7*(3), 451–468. https://doi.org/10.1111/tops.12148.

Reader, S. M., & Laland, K. N. (2001). Primate innovation: Sex, age and social rank differences. *International Journal of Primatology, 22*(5), 787–805. https://doi.org/10.1023/A:1012069500899.

Ricceri, L., Colozza, C., & Calamandrei, G. (2000). Ontogeny of spatial discrimination in mice: A longitudinal analysis in the modified open-field with objects. *Developmental Psychobiology, 37*, 109–118.

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin, 132*, 1–25.

Rowland, H. M., Wiley, E., Ruxton, G. D., Mappes, J., & Speed, M. P. (2010). When more is less: The fitness consequences of predators attacking more unpalatable prey when more are presented. *Biology Letters, 6*(6), 732–735. https://doi.org/10.1098/rsbl.2010.0207.

Sherratt, T. N. (2003). State-dependent risk-taking by predators in systems with defended prey. *Oikos, 103*(1), 93–100.

Sherratt, T. N. (2011). The optimal sampling strategy for unfamiliar prey. *Evolution, 65*(7), 2014–2025. https://doi.org/10.1111/j.1558-5646.2011.01274.x.

Skelhorn, J., Holmes, G. G., Hossie, T. J., & Sherratt, T. N. (2016). Multicomponent deceptive signals reduce the speed at which predators learn that prey are profitable. *Behavioral Ecology, 27*(1), 141–147. https://doi.org/10.1093/beheco/arv135.

Speed, M. P. (2001). Can receiver psychology explain the evolution of aposematism? *Animal Behaviour, 61*, 205–216.

Stamps, J. A., & Frankenhuis, W. E. (2016). Bayesian models of development. *Trends in Ecology & Evolution, 31*(4), 260–268. https://doi.org/10.1016/j.tree.2016.01.012.

Stamps, J. A., & Krishnan, V. V. (2017). Age-dependent changes in behavioural plasticity: Insights from Bayesian models of development. *Animal Behaviour, 126*, 53–67. https://doi.org/10.1016/j.anbehav.2017.01.013.

Stephens, D. W. (2007). Models of information use. In D. W. Stephens, J. S. Brown, & R. C. Ydenberg (Eds.), *Foraging: Behavior and ecology* (pp. 31–58). Chicago, IL: University of Chicago Press.

Wolf, M., van Doorn, G. S., Leimar, O., & Weissing, F. J. (2007). Life-history trade-offs favour the evolution of animal personalities. *Nature, 447*(7144), 581–584. https://doi.org/10.1038/nature05835.