

## Research



**Cite this article:** Trimmer PC, Ehlman SM, Sih A. 2017 Predicting behavioural responses to novel organisms: state-dependent detection theory. *Proc. R. Soc. B* **284**: 20162108. <http://dx.doi.org/10.1098/rspb.2016.2108>

Received: 25 September 2016

Accepted: 7 December 2016

**Subject Category:**

Ecology

**Subject Areas:**

behaviour, ecology, theoretical biology

**Keywords:**

HIREC, signal detection, state dependence, ecotourism

**Author for correspondence:**

Pete C. Trimmer

e-mail: [pete.trimmer@gmail.com](mailto:pete.trimmer@gmail.com)

Electronic supplementary material is available online at <https://dx.doi.org/10.6084/m9.figshare.c.3655577>.

# Predicting behavioural responses to novel organisms: state-dependent detection theory

Pete C. Trimmer<sup>1</sup>, Sean M. Ehlman<sup>1,2</sup> and Andrew Sih<sup>1</sup>

<sup>1</sup>Department of Environmental Science and Policy, and <sup>2</sup>Animal Behaviour Graduate Group, University of California-Davis, 1 Shields Avenue, Davis, CA 95616, USA

PCT, 0000-0003-0914-4150; SME, 0000-0001-6981-9549

Human activity alters natural habitats for many species. Understanding variation in animals' behavioural responses to these changing environments is critical. We show how signal detection theory can be used within a wider framework of state-dependent modelling to predict behavioural responses to a major environmental change: novel, exotic species. We allow thresholds for action to be a function of reserves, and demonstrate how optimal thresholds can be calculated. We term this framework 'state-dependent detection theory' (SDDT). We focus on behavioural and fitness outcomes when animals continue to use formerly adaptive thresholds following environmental change. In a simple example, we show that exposure to novel animals which *appear* dangerous—but are actually safe—(e.g. ecotourists) can have catastrophic consequences for 'prey' (organisms that respond as if the new organisms are predators), significantly increasing mortality even when the novel species is not predatory. SDDT also reveals that the effect on reproduction can be greater than the effect on lifespan. We investigate factors that influence the effect of novel organisms, and address the potential for behavioural adjustments (via evolution or learning) to recover otherwise reduced fitness. Although effects of environmental change are often difficult to predict, we suggest that SDDT provides a useful route ahead.

## 1. Introduction

In times of rapid change, experience could be your worst enemy.

—J. Paul Getty

Human-induced rapid environmental change (HIREC) has devastating effects on many species. By contrast, other species are thriving despite HIREC. A key issue is explaining this variation. The initial response to HIREC often involves behavioural plasticity [1,2]. Some animals exhibit adaptive behavioural responses to HIREC [3,4], while others show strikingly maladaptive behaviours; e.g. falling into evolutionary or ecological traps [5,6]. If the initial plastic response to HIREC is sufficiently effective for survival of a species over numerous generations, then populations can potentially evolve to cope better with HIREC. The effectiveness of initial behavioural responses before subsequent evolution is thus critically important for species persistence. Although hundreds of empirical studies have documented changes in behavioural responses following HIREC [4], few have used explicit, mathematical theory to predict why some species (or individuals) are responding well to HIREC while others are imperilled [7–9].

Here, we model behavioural responses to novel organisms [2,4]. Numerous studies show that naive prey sometimes respond appropriately to novel predators, but in other cases, they do not [10–13], which can be fatal. Numerous studies also show that naive consumers sometimes adopt novel resources or suitable novel habitats but, in other cases, ignore or even avoid novel options that could be used beneficially [14,15]. In particular, animals vary in their

tendency to avoid humans: some animals have moved in with humans (e.g. urbanized pests) while others appear to treat humans and human-altered habitats as highly dangerous even when we pose little or no danger [16–18]. Despite myriad case studies, theory that predicts responses to novel organisms is lacking, and explanations of variation in responses are typically constructed post hoc.

While standard optimality theory can potentially explain adaptive behavioural responses to HIREC (e.g. appropriate responses to novel predators or novel resources), our goal here is to explain *both adaptive and maladaptive* responses to novel organisms. Recent reviews [2,6,8] have emphasized that theory to explain variation in behavioural responses to HIREC should consider both the nature of the environmental change and the cognitive ‘rules’ used by focal species (which will typically depend on evolutionary history). Following that suggestion, we develop explicit theory to examine how behavioural rules that were formerly adaptive (before HIREC) affect the behavioural responses and fitness of individuals exposed to novel situations. Of particular interest are novel situations that result in evolutionary mismatches; e.g. situations that are safe but appear dangerous (SBAD).

In a separate paper, we examine responses to novel situations that are dangerous but might appear safe (e.g. introduced predators [12]). Here, we focus on the converse: responses to novel organisms that are SBAD; e.g. ecotourists. Organisms that have perfect information would not avoid them as they are safe, but because they appear potentially dangerous, ‘prey’ might avoid them, and this avoidance often has costs; e.g. reduced energy intake rate, growth and reproductive rates. Interestingly, many experiments aiming to measure avoidance of (and costs of avoiding) *actual* predators have ‘tricked’ animals into avoiding what are, in reality, SBAD. That is, numerous experimental studies have used various methods to induce animals to exhibit antipredator behaviour in the absence of actual dangerous predators; e.g. by exposing prey to predator playbacks [19] or predator chemicals [20] with no actual predators present, or to predators that have had their mouthparts disabled [21]. These studies (that were meant to quantify prey responses to actual predators) show that prey can be misled by their evolved cue-response systems into responding unnecessarily to what are, in fact, SBAD. We thus know that fear responses to SBAD can be persistent (i.e. animals often do not learn quickly to ignore SBAD and may never habituate to disturbance stimuli [16]), and that they can result in long-term costs: reduced feeding rates, energy reserves, growth rates and even fecundity that have been termed the non-consumptive effects of predators [22]. However, we also know that there is substantial unexplained variation in animal responses to SBAD (either ecotourists, other novel SBAD in nature, or simulated predators in experiments). Here, we develop models that predict when organisms should be more likely to avoid SBAD and when the costs of avoidance might be particularly large.

To predict not just behavioural responses to SBAD, but also the effects of these responses on fitness, we develop a new modelling framework: ‘state-dependent detection theory’ (SDDT) which makes use of signal detection theory [23,24] in a state-dependent manner [25]. The method assesses: (i) how the optimal response threshold (governing behaviour) in each time unit depends on the organism’s state (e.g. energy reserves); (ii) how the resulting behavioural response affects the organism’s survival and subsequent

state; (iii) the optimal state for reproduction (i.e. it assesses a joint optimal behavioural and life-history strategy); (iv) how these outcomes over a lifetime determine reproductive success; and (v) can thus be used to evaluate the effects of environmental change, such as the fitness cost of SBAD. Here, we use state-dependent signal detection to generate predictions on responses to novel SBAD; however, this new modelling method has exciting potential to be useful for numerous issues involving behaviour under uncertainty (e.g. food choices or habitat selection).

In the next section (§2), we summarize signal detection theory. In §3, we describe how costs and benefits required by the approach can have meaningful values assigned to them in the context of energy reserves. In §4, we provide a simple example to show how this framework can be used to look at decisions that maximize reproductive success under uncertainty. In §5, we use SDDT to consider the effects of introducing novel animals that appear dangerous but are actually safe.

## 2. Signal detection theory

Signal detection theory (also known simply as detection theory [23,24]) allows optimal behavioural decisions to be identified in a range of circumstances. The standard model assumes that there are two (mutually exclusive) environmental conditions and two possible actions, with payoffs depending on the condition and the choice of action. For instance, consider a scenario in which an individual is uncertain about whether a predator is present (in which case, it should run away). If the signal received from a predator is typically higher (e.g. louder, larger) than the sensory noise when there is no predator, it can be optimal to set a detection threshold whereby the individual reacts as though a predator is present for signals above a critical level (the threshold), otherwise reacting as though safe. Analogous situations relate to food selection [26], mate selection [27], etc.; we focus on predator detection as an example.

Table 1 summarizes the four possible payoffs,  $V$ ; each denotes the expected value of life for having reached that combination of environmental condition and action taken; the meaning we assign to these values is discussed in subsequent sections.

Prior to receiving a cue, the probability of each environmental condition is assumed to be known by the individual; i.e. we assign a probability,  $p_D$ , to the world being dangerous. The individual then receives a cue and must decide what action to take. The standard assumptions are that, for a given environmental condition (here, dangerous or safe), the cue has a normal distribution with a particular variance, and the signal mean depends on the environmental condition. We assume that dangerous situations have a greater mean signal level than safe ones and, for simplicity, that the variance of each type of cue is the same. Egan [24] shows that the position of the optimal threshold,  $x_T$ , is such that

$$\frac{P(x_T|D)}{P(x_T|S)} = \frac{(1 - p_D)(V_{SF} - V_{SR})}{p_D(V_{DR} - V_{DF})}$$

where  $p_D$  is the probability that the animal is dangerous; for other parameters, see table 1.

**Table 1.** General signal detection theory payoff matrix.

		environmental condition	
		dangerous, D	safe, S
action taken	run away, R	$V_{DR}$	$V_{SR}$
	forage, F	$V_{DF}$	$V_{SF}$

For signals greater than  $x_T$ , the individual should run, rather than forage. Numerous authors have made use of this rule in explaining behavioural effects in humans and other animals, such as the ‘smoke detector principle’ [28,29], ‘error management theory’ [30] and the evolution of ‘paranoid optimists’ [31].

Qualitative statements are easy to make and can, to an extent, be checked and quantified using signal detection theory. However, even if the cue distributions are perfectly known, the payoff parameters,  $V$ , must have values assigned to them to make any predictions at a quantitative level; the payoffs are not always easily known.

In this paper, we set signal detection theory within a larger framework of state-dependent modelling, whereby the payoff values can become meaningful. Rather than assume that animals have a single fixed threshold, we allow the threshold to vary as a function of their internal state, producing state-dependent decision rules. For example, we allow animals to be less responsive to signs of potential danger when their energy reserves are low. This modelling framework, described below, can result in more meaningful (and sometimes more unexpected and subtle) predictions than standard signal detection theory.

### 3. Payoff values depend on current state

Signal detection theory requires the expected payoffs to be known for each combination of situation (environmental condition) and action. However, each ‘reproductive value’ (RV, a technical term for the ‘value of life’; see [32]) can be very difficult to quantify as it depends on the current circumstance of the individual. The RV of an individual can often be usefully approximated by the expected reproductive success (i.e. number of offspring) or, in some circumstances, by expected survival time, using dynamic programming [25].

Crucially, the payoffs will generally depend on the reserve levels of the individual [25]. For instance, consider the expected payoff for mistakenly responding (running and hiding) when there is no predator. If responding uses one unit of reserves, then at high reserve levels the individual can afford such a loss; it will still have a high RV. However, at low-reserve levels, responding can severely reduce the survival prospects; at critically low reserves, the individual must forage if responding would result in death through starvation. To incorporate state, we allow the individual to set a response threshold at each reserve level. The expected payoffs then depend on the thresholds at each reserve level.

For a given set of thresholds, the RV at a given level of reserves can be calculated (i.e. updated) using the threshold at that reserve level and the values at other reserve levels. Having updated each of the values, we calculate the optimal threshold at each reserve level by using the values at

surrounding reserve levels in the payoff matrix. By repeating this process until both the values and thresholds have converged, we identify the optimal strategy, as described in electronic supplementary material, S1. By so doing, the emergent payoff values are also meaningful as RVs.

The outcome of our SDDT approach could, in theory, have been reached using standard state-dependent modelling techniques without recourse to signal detection theory, by applying a very fine grid of possible thresholds at each reserve level (effectively treating the threshold for running as the action that is being chosen), as described in electronic supplementary material, S2. However, the number of calculations required increases as the resolution of thresholds becomes finer, showing a clear benefit of using the signal detection theory equation within the state-dependent framework. As run-time can be an issue for state-dependent approaches, SDDT thus increases the scope of state-dependent modelling; especially in situations involving ambiguous cues.

The general principle is one of identifying behaviour that maximizes RV. In an evolved stable world, the average RV of individuals must be one (for the population not to be increasing or decreasing). We therefore assume that, before any environmental change, between being born and becoming viable there is a linear rescaling that brings the initial reproductive rate of a parent down to one viable offspring on average, prior to the offspring having any opportunity to reproduce. In some systems, expected lifespan is a good proxy for RV; we show how this approach can be used to generate meaningful results in electronic supplementary material, S3. Here, we focus on a more interesting case, where individuals wait until they have attained sufficient reserves before reproducing. We then show that looking at expected lifespan for such individuals following HIREC can be highly misleading.

### 4. Maximizing reproductive success

We discretize time into a series of periods, with the environmental condition being either dangerous (predator present) or safe (predator absent) in any given period, and independent of other periods. We assume that, in each period, the animal either: (i) lives cautiously (hides more, forages less), in which case its reserves decrease by one unit; this corresponds to ‘responding’ (table 1), (ii) forages more actively, in which case reserves increase by one unit, or (iii) reproduces if reserves are high enough. We allow reserves to change by one unit as the simplest example for illustrative purposes; altering the model to affect reserves by more than a single unit does not change qualitative results. We assume that death (be it through starvation or predation) occurs at the end of a period. Setting optimal thresholds at each reserve level,  $i$ , depends on the payoff values of each action.

For simplicity, we assume asexual reproduction, and assume that the focal individual will reproduce when it reaches a particular level of reserves,  $L$ , thereby adding one unit to its payoff of reproductive success, but decreasing its reserves in the process by some amount,  $c$ . We have used  $c = 6$  for illustration. The animal would die if reserves fell to zero. Thus, reserves take an integer value,  $\{1, \dots, L\}$ .

With our baseline parameter values (figure 2), animals live for about 260 time units, reproducing about four times. This could thus represent an invertebrate where each time

unit corresponds to a day, producing a new batch of eggs every couple of months, or a longer-lived vertebrate where each time unit represents a small number of weeks and reproduction occurs once a year.

The payoff matrix at most reserve levels depends on the values at higher and lower reserves, combined with the probabilities of survival. At maximum reserves,  $L$ , the payoff is  $V(L - c) + 1$ , on the assumption that the animal is hidden from the risk of predation while reproducing; i.e. we assume that at maximum reserves,  $L$ , the individual spends  $c$  units of reserves to produce one offspring (and thus leaving the parent with a future reproductive value of  $V(L - c)$ ). The escape probability for animals that respond to predators is  $e$ , and the probability of being missed by the predator despite foraging is  $m$ ;  $e > m$ . Thus, because reserves alter by one at each time step in this simple example, starting from reserves  $i < L$ , the payoffs are  $V_{DR} = V(i - 1)e$ ,  $V_{SR} = V(i - 1)$ ,  $V_{DF} = V(i + 1)m$ ,  $V_{SF} = V(i + 1)$ .

We assume that individuals act to maximize their expected number of offspring. An iterative method of calculating that value and the optimal thresholds is described in electronic supplementary material, S1. If individuals are born with reserves of one unit then, for a given value of  $L$ , the converged payoff value at reserves of one is the expected number of offspring produced by an individual during their lifetime.  $L$  governs the risk of not reproducing (through dying before reproducing if  $L$  is large) and the danger associated with always having low reserves (if  $L$  is small). Trading off these effects, the optimal reproductive level is the value of  $L$  that maximizes  $V(1)$ .

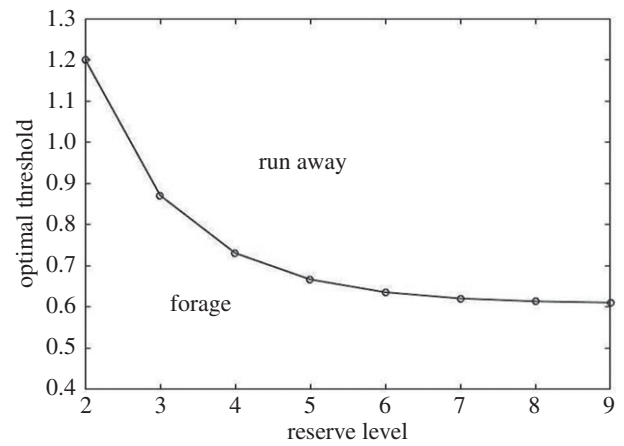
Figure 1 shows the optimal thresholds with respect to reserves that maximize reproductive success. As expected, a stronger cue is required to induce a response when individuals have lower reserves.

## 5. Behavioural responses to novel safe organisms that appear dangerous

Here, we examine situations where naive prey encounter novel organisms that are SBAD. For example, humans often appear dangerous but are actually safe for many organisms; animals often appear to over-avoid humans and human-altered situations [16–18].

We first consider a scenario where the prevalence of the focal organism's real predators is unchanged. A novel species that will not attack our focal organism enters and is encountered with probability  $p_N$ . Despite the novel organisms being safe, their mean signal,  $\mu_N$ , may make them appear dangerous; if  $\mu_N = \mu_D$  then the novel organisms are indistinguishable from actual predators. We assume no change (yet) in the organism's formerly adaptive response threshold. As the prevalence of novel organisms increases, prey unnecessarily run away more often, which reduces reserves and thus survival and fecundity, as shown in figure 2. The reduced survival is interesting because in our model, prey do not starve to death (as prey always forage when reserves get dangerously low) and the new SBAD organisms never kill our naive prey. Survival decreases because when reserves get low, the increased pressure to feed exposes them to an increased risk of predation by the real predators.

In parallel with the predator similarity hypothesis for novel predators [12], negative effects of novel SBAD are



**Figure 1.** Optimal decision threshold as a function of reserves. There is no decision to be made at reserves of one as it is best to forage regardless of the risk of predation. There is also no decision to be made at maximum reserves ( $L = 10$ ), as the individual reproduces at this size. At any intermediate reserve level, there is an optimal threshold whereby a signal greater than that threshold should cause the individual to run away, otherwise they should continue to forage. Note that the threshold decreases with reserves, showing that those with more to lose are more prone to avoid potential danger. (Parameters:  $p_D = 0.1$ ,  $\mu_S = 0$ ,  $\mu_D = 2$ ,  $\sigma = 1$ ,  $e = 1$ ,  $m = 0$ ,  $c = 6$ .)

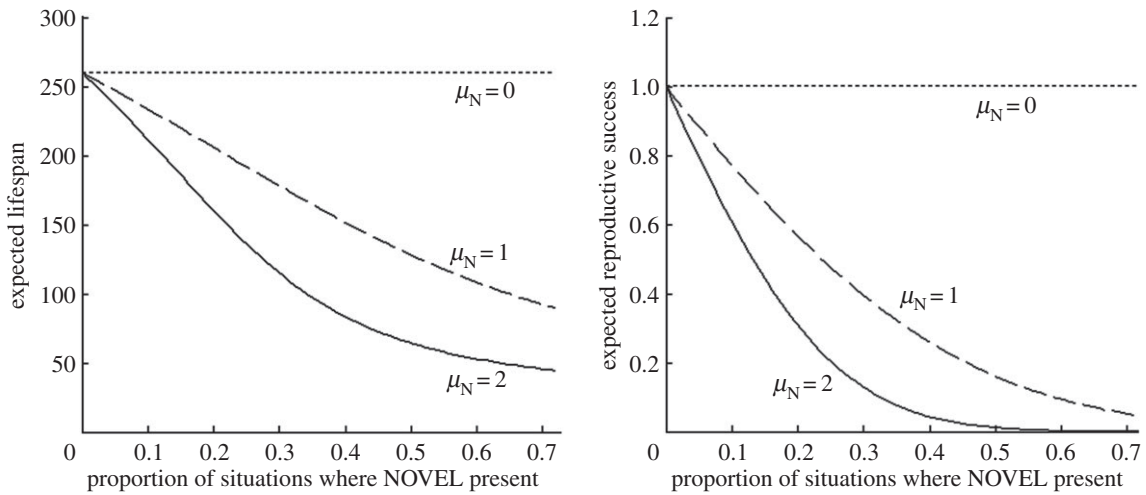
reduced if the novel organisms are less similar to real dangers (contrast  $\mu_N = 2, 1$  and  $0$ ). However, even if the mean signal is halfway between safe and dangerous ( $\mu_N = 1$ ), the cost of SBAD can be large (figure 2). Thus, introducing any novel organism that does not appear entirely safe can have significant negative impacts.

Figure 3 illustrates the interaction of escape probability when running from a predator,  $e$ , and the introduction of SBAD. With a high probability of escape when running (e.g.  $e = 1$ ), it is more important to run away when there are signs of danger, so SBADs have a greater effect. In electronic supplementary material, S4, we discuss the effects of other parameters, including the prevalence of predators ( $p_D$ ) before the introduction of SBAD, and the probability of mortality even if prey respond to predators ( $m$ ).

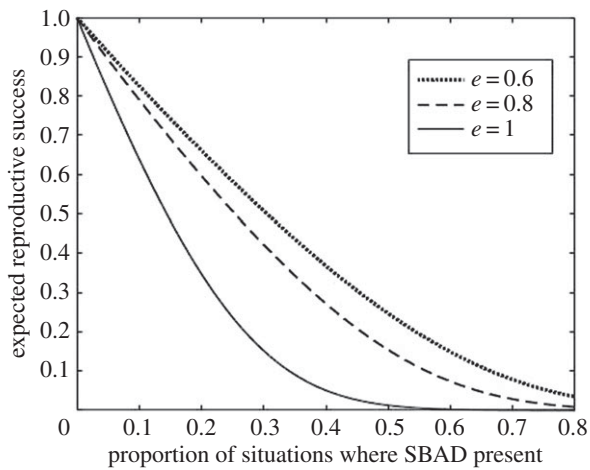
The fitness cost of SBAD can be particularly large if there was previously little overlap between safe and dangerous cue distributions (see electronic supplementary material, S4; figure D3). When exposed to SBAD that closely resemble dangerous predators, prey that evolved in situations where they could easily distinguish between safe and dangerous situations take evasive action even more frequently until they have low reserves, at which point they are more likely to expose themselves to real predators.

The introduction of SBAD can be accompanied by a decrease in the prevalence of dangerous animals. This could happen, for instance, when humans populate a new area and kill some existing predators. If humans are SBAD, so the environment is objectively less dangerous, the focal organism's evolved behavioural response can still produce deleterious effects. Here, we consider the situation where HIREC halves the number of real predators, and assume that the focal species continues to use the same thresholds for decision-making throughout life (i.e. for now, we ignore the effect of learning).

Because the environment is now less dangerous, not surprisingly, if SBADs (humans in this case) are rarely



**Figure 2.** The reduction in expected lifespan and reproductive success associated with an increase in the prevalence of organisms that are safe but appear dangerous (SBAD). As the mean signal of the novel animal,  $\mu_N$ , increases to look more like a predator, the effect on expected lifespan and reproductive success of the focal individual decreases. This occurs even though the novel species is not predatory. (Parameters:  $p_D = 0.1$ ,  $\mu_S = 0$ ,  $\mu_D = 2$ ,  $\sigma = 1$ ,  $e = 1$ ,  $m = 0$ ,  $c = 6$ .)



**Figure 3.** As the probability of escaping a predator when responding,  $e$ , increases, the effect of SBAD also increases. This is because there is ever more reason to run from predator-like signals when  $e$  is large, so under such circumstances, SBADs have a more significant effect on behaviour. Note that each line has been 'normalized' such that without SBAD, the expected reproductive success is 1, so the population size is, on average, unchanging prior to SBAD being introduced. (Parameters:  $p_D = 0.1$ ,  $\mu_S = 0$ ,  $\mu_D = 2$ ,  $\mu_N = 2$ ,  $\sigma = 1$ ,  $e = 1$ ,  $m = 0$ ,  $c = 6$ .)

encountered then focal individuals tend to live longer and reproduce more (figure 4). However, as the prevalence of SBAD increases, reserve levels of the focal species tend to be lower and they thus reproduce less. For a range of SBAD prevalence (between the vertical lines in figure 4), the environmental change reduces reproduction despite increasing survival (i.e. expected lifespan). With an even higher rate of SBAD encounters, focal animals both reproduce less and have shorter lives. Thus, introducing SBAD can have disastrous consequences for the focal species, even if the environmental change significantly reduces the number of real predators.

### (a) Adjustments to behaviour post-human-induced rapid environmental change

In the post-HIREC environment, either learning or evolution can adjust both (i) the optimal thresholds for when to run and

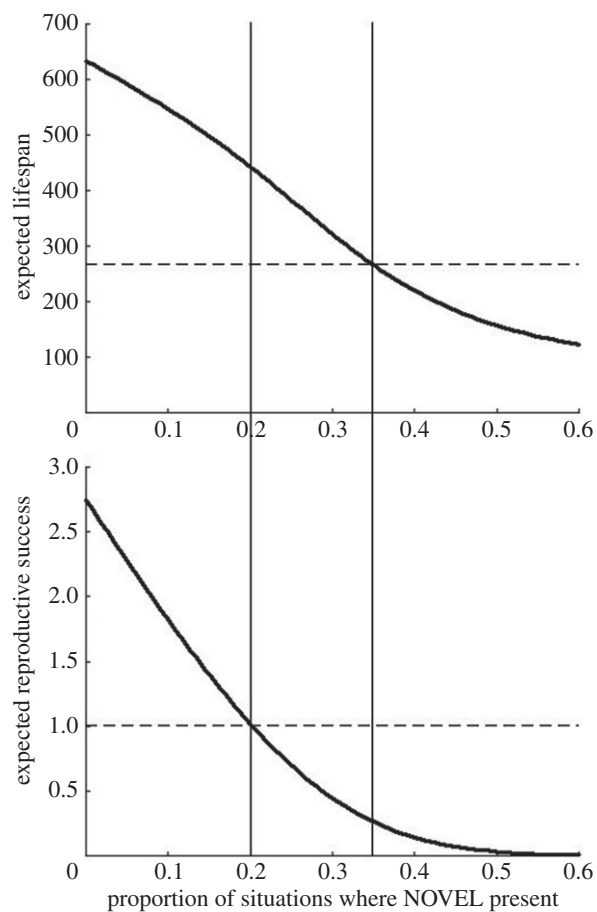
(ii) the reserve levels at which to reproduce ( $L$ ). For instance, in the situation analysed above, if HIREC results in SBAD being seen 25% of the time, then individuals would be better off reproducing at reserves of  $L = 8$  (rather than 10), with less tendency to run at all reserve levels. However, even after learning or evolving the new optimal thresholds, the expected number of offspring is still reduced from the original level. This occurs because with a substantial number of SBAD organisms in the environment, prey have less ability to discriminate between safe and dangerous (as the novel organisms appear dangerous), and even after adjusting thresholds optimally, the uncertainty *per se* reduces lifetime reproductive success.

Figure 5 illustrates that HIREC can have a greater impact on reproductive success than any subsequent evolution of behavioural thresholds and decisions about when to reproduce. In the example, the introduction of an abundant SBAD reduces expected reproductive success by about 75%; however, even after the organism evolves (or learns) to account for a substantial number of organisms that are safe, but appear dangerous, the inability to distinguish between the types still reduces reproductive success by more than 60%.

Figure 5 is based on a single probability of encountering the new species, and this could be argued to be relatively high ( $p_N = 0.225$ ). However, the finding that the ability to discern predators has a greater effect than evolving thresholds is a relatively general one; with a smaller value of  $p_N$ , the effect of the HIREC on reproductive success would be reduced—but the effect of evolving new thresholds would similarly be reduced.

## 6. Discussion

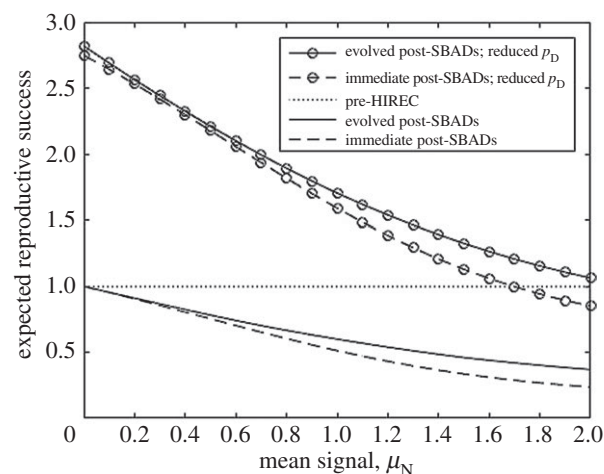
Initial behavioural responses to HIREC can be crucial to species' success. Although numerous empirical studies have quantified variation in behavioural responses to various aspects of HIREC (e.g. novel species, habitat change and climate change), relatively few theoretical papers have generated explicit predictions. One approach to understanding responses to HIREC involves assuming that past selection pressures have shaped formerly adaptive traits that will at



**Figure 4.** The effect of HIREC on expected lifespan and reproductive success, as a function of the proportion of safe situations where a novel animal is experienced, when HIREC kills half of the predators. The horizontal (dashed) lines show the pre-HIREC performance. When the SBADs are rarely present, focal individuals do better than before the HIREC change, due to half the predators being killed; when SBADs are common, they do worse. Between the vertical lines, HIREC increases expected lifespan but decreases lifetime reproduction. (Parameters:  $\mu_S = 0$ ,  $\mu_D = 2$ ,  $\mu_N = 2$ ,  $\sigma = 1$ ,  $e = 1$ ,  $m = 0$ ,  $c = 6$ ,  $p_D = 0.1$  initially, resulting in  $L = 10$ . Post-HIREC, dangerous situations have been halved, so  $p_D$  becomes 0.05.)

least initially be used immediately after HIREC [2,6,8]. We used this approach to address an aspect of behavioural responses to novel organisms: the effects of novel safe organisms that appear dangerous. Specifically, we developed a new modelling framework, SDDT, to address decision-making and changes in expected lifespan, reproductive success, and life histories in response to an influx of organisms that are SBAD, all within a single coherent framework. Standard signal detection theory and state-dependent modelling (by dynamic programming) are both well-known frameworks that can be used to model behaviour under uncertainty. By combining these paradigms, our approach makes the payoff values (required by standard signal detection theory) meaningful—which should, in the longer term, mean that predictions made using this theory can be more readily tested in the real world.

We showed that SBAD can cause major reductions in survival and reproduction; the *perception* of risk per se can be very costly. Indeed, even changes that might be expected to benefit a population (e.g. the introduction of safe, non-venomous snakes that reduce the number of venomous snakes that prey on the focal species) can have deleterious



**Figure 5.** Expected reproductive success as a function of the mean signal produced by the novel species following two forms of HIREC. These are shown both as immediate post-HIREC effects and longer term, when thresholds have subsequently evolved following HIREC. Introduction of the novel species (lines without markers) has deleterious consequences. If HIREC also halves the number of predators (circle markers), a positive effect can be obtained when the novel species can be readily distinguished from predators. A mean signal of 0 is indistinguishable from the safe signal and a mean signal of 2 is indistinguishable from the dangerous signal. Note that there is relatively little difference between the lines of effects of SBAD before (dashed lines) versus after (solid lines) the focal organism has had time to evolve (or learn) to adjust their behavioural and life-history strategies. (Parameters:  $p_D = 0.1$  initially,  $\mu_S = 0$ ,  $\mu_D = 2$ ,  $\sigma = 1$ ,  $e = 1$ ,  $m = 0$ ,  $c = 6$ . Post-HIREC,  $p_N = 0.225$ .)

consequences for the focal species. The key aspect is that the novel animals *appear* to be dangerous. Even when some of the predators are killed, so the environment is objectively safer than pre-HIREC, the focal species may live longer but produce fewer offspring because they are hiding more often and thus living their lives in a chronically low-reserve state. Our modelling approach will also allow other (initially non-lethal) life-history effects to be considered. For instance, the quality of offspring can be studied by altering the level of fat of offspring at the point of weaning, thus affecting the offspring's resulting survival chances (cf. [33] for a slightly different modelling approach on this topic).

Introducing SBAD tends to have more deleterious effects on reproduction than lifespan; under some conditions, individuals are predicted to live for nearly as long but hardly reproduce. This occurs because prey with fairly high reserves readily run and hide when they detect strong cues (causing a loss of reserves); in the altered environment, this can often be triggered by the novel (but safe) animals. Consequently, the focal animals rarely reach reproductive reserve levels. At low reserves, however, animals must forage even when the signal indicates possible danger. When SBADs are common, most situations are safe, so foraging is often beneficial. Consequently, animals tend to stay alive for a considerable time but at chronically low to moderate reserve levels.

The prediction that SBAD can reduce survival might seem counterintuitive because in our model SBAD never kill focal prey, and the focal prey never starve to death. Nonetheless, SBADs cause prey to die more frequently because when prey are induced to hide by SBAD, their consequent lower reserves causes them to take more risks to avoid starvation,

which, in turn, results in higher mortality from real predators. In the longer term, this may result in an increase in the number of predators, which would further increase the costs of SBAD.

Not surprisingly, we predict that the costs of SBAD will be higher if the SBAD more closely resemble real predators. This prediction is parallel to the ‘predator similarity hypothesis’ that novel predators are more likely to be recognized as danger if they resemble familiar predators [12] which might be more likely if prey have been exposed to multiple predators [34]. By using an explicit state-dependent model, we add the more detailed suggestion that even when SBADs are moderately different from real predators, they can still have large negative impacts on both prey reproductive success and survival. We are also able to make the less intuitive predictions that the relative costs of SBADs are higher if existing predators are less dangerous (figure 3), and if existing predators and non-predators are more easily distinguished (electronic supplementary material, S4; figure D3). The latter prediction shows that prey that made fewer errors (that seemed more ‘intelligent’) before HIREC might be more susceptible to high costs of HIREC.

Interestingly, the numerous studies using simulated predators (playbacks, predator chemicals, human approach) to assess prey avoidance of real predators, and their costs (the non-consumptive effects) are actually literature on SBAD [17,19,20]. These studies implicitly assume (or hope) that the simulated predators closely resemble real predators. In reality, simulated predators, that are actually SBAD, vary in their similarity to real predators. Our models examine explicitly how the similarity of SBAD to real predators should affect responses to SBAD, and how these responses should be mediated by other cost/benefit considerations. Thus, our models potentially provide a better fit to studies using non-lethal simulated predators than existing theory on optimal predator avoidance. Notably, our approach highlights factors that are rarely considered in experimental studies of non-consumptive effects. We predict that responses to SBAD (whether they are simulated predators in an ecologist’s experiment, or novel organisms, or ecotourists) should depend on not just characteristics of actual dangerous predators, but also on the similarity of familiar *safe* organisms to those predators, the relative prevalence of safe versus dangerous organisms that overlap in cues, and the costs of over-avoiding existing SBAD. That is, beyond generating predictions for animal responses to novel SBAD, our models also provide new insights for interpreting experiments using simulated predators to quantify prey responses to real predators.

Our results accord with an existing literature on prey responses to ecotourists. Previous studies have shown decreases in organismal condition and mass in areas of high ecotourism, and this is likely due to animals’ perceptions of risk rather than any physical disturbance [35–38]. In particular, we predict, and some observations confirm, that ecotourism that does not directly harm focal animals can kill them by causing them to hide less from their real predators. As noted above, our models predict situations when this unfortunate outcome is more likely to occur. The effect of ecotourists on any particular species is likely to depend on their life history: factors such as reproductive demands, stress responses and learning during life. For instance, as Müllner *et al.* [35] identify, the effect of ecotourists on hoatzins is to reduce survival of chicks pre-fledging and to

reduce the body mass of juniors, but, later in life, the flush distances of individuals were reduced in areas with high ecotourism suggesting that learning or habituation has occurred. One method of accounting for learning during life would be to allow the signal distribution associated with a novel species to alter over time (i.e. appearing less dangerous with time as a focal species learns about the novel species).

We see several directions for future modelling. First, our initial models set thresholds for behaviour without considering the effect of adjustments in response thresholds via learning [39]. However, insights can be gleaned on the potential importance of learning by contrasting ‘immediate post-HIREC’ reproductive success (based on the use of previously adaptive response thresholds) against the ‘evolved post-HIREC’ success (based on thresholds optimized to fit the new world now including SBAD). Figure 5 suggests that the effect of learning that SBADs are now present is small compared with the potential effects of an increase in the ability of the focal prey to discriminate that the novel animals are not actually predators. That is, the main benefit of learning comes only if prey can learn to better distinguish SBAD from real predators.

We have pitched signals for danger on a single scale. In many real situations, although a novel animal may not initially be distinguished from a predator using current modalities (e.g. visual appearance), new signals (such as sound or smell) may be learned (or senses evolved) over time. Another possibility is multiple thresholds (for numerous possible actions) per reserve level. Sole *et al.* [40] show that pigeon (*Columbidae*) decisions can be accurately described by a signal detection model with more than one threshold, maximizing perceived reward [40]. This modelling has also assumed that the danger posed by predators is independent from one period to another, whereas the world is generally auto-correlated in this respect, which can affect learning and behaviour [41].

State-dependent modelling is often used to consider the effect of reserves on optimal behaviour given stochasticity in foraging success [25,42]. SDDT can include such stochasticity; here, we have simply assumed that food intake is deterministic (as may occur in a grazing animal) and considered the effect of reserves when dealing with uncertainty in cues for danger. Our results concur with various studies that suggest that animals take more evasive action when they are in good condition [43,44]. Rather than directly link reserves and behaviour, this modelling approach links reserves with mental state (preparedness for behaviour), a step towards evolving mental mechanisms [45].

Finally, we note that SDDT can be applied to other scenarios involving novel environmental stimuli (e.g. novel foods or novel habitats). The general scenario is that each signal relates to an item that is either beneficial or deleterious, that organisms experience some uncertainty about any given item, and that the organism’s best response depends on the signals relating to each item and the costs and benefits of responding or not. Thus, the same basic model can be modified to address factors that should affect whether uncertain prey will avoid novel real predators, or whether uncertain foragers will attempt to consume novel items when some familiar options are good food, but others are of low quality or even toxic. Overall, the goal is a more fully developed theory to understand the effect of novel conditions on behavioural responses and species’ success.

**Data accessibility.** All parameter values have been identified in this manuscript and electronic supplementary material.

**Authors' contributions.** P.C.T. conceived the principles and led in direction setting, analysis and writing the manuscript. S.M.E. coded the model and helped with analysis and writing of the manuscript. A.S. helped with direction setting, structuring the writing, choosing appropriate parameter values and writing the manuscript.

**Competing interests.** We have no competing interests.

**Funding.** P.C.T. was supported by the ERC (Evomech Advanced Grant 250209 to A.I. Houston) and the NSF (IOS 1456724 grant to A.S.). S.M.E. was supported by a grant from the UC Davis Animal Behavior Graduate Group and an NSF Graduate Research Fellowship.

**Acknowledgements.** We thank members of the Sih laboratory and the Bristol MAD group for constructive discussions along the way. We also thank Niels Dingemanse, Dan Blumstein and anonymous reviewers for their helpful feedback.

## References

- Hendry AP, Farrugia TJ, Kinnison MT. 2008 Human influences on rates of phenotypic change in wild animal populations. *Mol. Ecol.* **17**, 20–29. (doi:10.1111/j.1365-294X.2007.03428.x)
- Sih A, Ferrari MCO, Harris DJ. 2011 Evolution and behavioural responses to human-induced rapid environmental change. *Evol. Appl.* **4**, 367–387. (doi:10.1111/j.1752-4571.2010.00166.x)
- Tuomainen U, Candolin U. 2010 Behavioural responses to human-induced environmental change. *Biol. Rev.* **86**, 640–657. (doi:10.1111/j.1469-185X.2010.00164.x)
- Candolin U, Wong B. 2012 *Behavioural responses to a changing world: mechanisms and consequences*. Oxford, UK: Oxford University Press.
- Schlaepfer MA, Runge MC, Sherman PW. 2002 Ecological and evolutionary traps. *Trends Ecol. Evol.* **17**, 474–480. (doi:10.1016/S0169-5347(02)02580-6)
- Robertson BA, Rehage JS, Sih A. 2013 Ecological novelty and the emergence of evolutionary traps. *Trends Ecol. Evol.* **28**, 552–560. (doi:10.1016/j.tree.2013.04.004)
- Johansson J, Bolmgren K, Jonzén N. 2012 Climate change and the optimal flowering time of annual plants in seasonal environments. *Glob. Change Biol.* **19**, 197–207. (doi:10.1111/gcb.12006)
- Sih A. 2013 Understanding variation in behavioural responses to human-induced rapid environmental change: a conceptual overview. *Anim. Behav.* **85**, 1077–1088. (doi:10.1016/j.anbehav.2013.02.017)
- McNamara JM, Barta Z, Klaassen M, Bauer S. 2011 Cues and the optimal timing of activities under environmental changes. *Ecol. Lett.* **14**, 1183–1190. (doi:10.1111/j.1461-0248.2011.01686.x)
- Chivers DP, Wildy EL, Kiesecker JM, Blaustein AR. 2001 Avoidance response of juvenile Pacific treefrogs to chemical cues of introduced predatory bullfrogs. *J. Chem. Ecol.* **27**, 1667–1676. (doi:10.1023/A:1010418526991)
- Cox JG, Lima SL. 2006 Naïveté and an aquatic-terrestrial dichotomy in the effects of introduced predators. *Trends Ecol. Evol.* **21**, 674–680. (doi:10.1016/j.tree.2006.07.011)
- Sih A, Bolnick DI, Luttbeg B, Orrock JL, Peacor SD, Pintor LM, Preisser E, Rehage JS, Vonesh JR. 2010 Predator–prey naïveté, antipredator behavior, and the ecology of predator invasions. *Oikos* **119**, 610–621. (doi:10.1111/j.1600-0706.2009.18039.x)
- Nunes AL, Richter-Boix A, Laurila A, Rebelo R. 2013 Do anuran larvae respond behaviourally to chemical cues from an invasive crayfish predator? A community-wide study. *Oecologia* **171**, 115–127. (doi:10.1007/s00442-012-2389-6)
- Gilroy JJ, Sutherland WJ. 2007 Beyond ecological traps: perceptual errors and undervalued resources. *Trends Ecol. Evol.* **22**, 351–356. (doi:10.1016/j.tree.2007.03.014)
- Pearse IS, Harris DJ, Karban R, Sih A. 2013 Predicting novel herbivore–plant interactions. *Oikos* **122**, 1554–1564. (doi:10.1111/j.1600-0706.2013.00527.x)
- Frid A, Dill LM. 2002 Human-caused disturbance stimuli as a form of predation risk. *Conserv. Ecol.* **6**, 11–26. (doi:10.5751/ES-00404-060111)
- Beale CM, Monaghan P. 2004 Human disturbance: people as predation-free predators? *J. Appl. Ecol.* **41**, 335–343. (doi:10.1111/j.0021-8901.2004.00900.x)
- Blumstein DT, Fernández-Juricic E, Zollner PA, Garity SC. 2005 Inter-specific variation in avian responses to human disturbance. *J. Appl. Ecol.* **42**, 943–953. (doi:10.1111/j.1365-2664.2005.01071.x)
- Zanette LY, White AF, Allen MC, Clinchy M. 2011 Perceived predation risk reduces the number of offspring songbirds produced per year. *Science* **334**, 1398–1401. (doi:10.1126/science.1210908)
- Kats LB, Dill LM. 1998 The scent of death: chemosensory assessment of predation risk by prey animals. *Ecoscience* **5**, 361–394. (doi:10.2307/42902443)
- Schmitz OJ, Beckerman AP, O'Brien KM. 1997 Behaviorally mediated trophic cascades: effects of predation risk on food web interactions. *Ecology* **78**, 1388–1399. (doi:10.1890/0012-9658)
- Preisser EL, Bolnick DI, Benard MF. 2005 Scared to death? The effects of intimidation and consumption on predator–prey interactions. *Ecology* **86**, 501–509. (doi:10.1890/04-0719)
- Kingdom F, Moulden B. 1989 Modelling visual detection: luminance response non-linearity and internal noise. *Q. J. Exp. Psychol. A* **41**, 675–696. (doi:10.1080/14640748908402389)
- Egan JP. 1975 *Signal detection theory and ROC analysis*. New York, NY: Academic Press. (doi:10.1234/12345678)
- Houston AI, McNamara JM. 1999 *Models of adaptive behaviour: an approach based on state*. Cambridge, UK: Cambridge University Press.
- Getty T, Krebs JR. 1985 Lagging partial preferences for cryptic prey: a signal detection analysis of great tit foraging. *Am. Nat.* **125**, 39–60. (doi:10.2307/2461607)
- Wiley RH. 2006 Signal detection and animal communication. *Adv. Study Behav.* **36**, 217–247. (doi:10.1016/S0065-3454(06)36005-6)
- Nesse RM. 2001 The smoke detector principle. *Annu. N. Y. Acad. Sci.* **935**, 75–85. (doi:10.1111/j.1749-6632.2001.tb03472.x)
- Nesse RM. 2005 Natural selection and the regulation of defenses. *Evol. Hum. Behav.* **26**, 88–105. (doi:10.1016/j.evolhumbehav.2004.08.002)
- Haselton MG, Buss DM. 2000 Error management theory: a new perspective on biases in cross-sex mind reading. *J. Pers. Soc. Psychol.* **78**, 81–91. (doi:10.1037/0022-3514.78.1.81)
- Haselton MG, Nettle D. 2006 The paranoid optimist: an integrative evolutionary model of cognitive biases. *Pers. Soc. Psychol. Rev.* **10**, 47–66. (doi:10.1207/s15327957pspr1001\_3)
- Fisher RA. 1930 *The genetical theory of natural selection*. New York, NY: Oxford University Press.
- New LF *et al.* 2014 Using short-term measures of behaviour to estimate long-term fitness of southern elephant seals. *Mar. Ecol. Prog. Ser.* **496**, 99–108. (doi:10.3354/meps10547)
- Blumstein DT. 2006 The multipredator hypothesis and the evolutionary persistence of antipredator behavior. *Ethology* **112**, 209–217. (doi:10.1111/j.1439-0310.2006.01209.x)
- Müllner A, Eduard Linsenmair K, Wikelski M. 2004 Exposure to ecotourism reduces survival and affects stress response in hoatzin chicks (*Opisthocomus hoazin*). *Biol. Conserv.* **118**, 549–558. (doi:10.1016/j.biocon.2003.10.003)
- McClung MR, Seddon PJ, Massaro M, Setiawan AN. 2004 Nature-based tourism impacts on yellow-eyed penguins *Megadyptes antipodes*: does unregulated visitor access affect fledging weight and juvenile survival? *Biol. Conserv.* **119**, 279–285. (doi:10.1016/j.biocon.2003.11.012)
- Amo L, López P, Martí NJ. 2006 Nature-based tourism as a form of predation risk affects body condition and health state of *Podarcis muralis* lizards. *Biol. Conserv.* **131**, 402–409. (doi:10.1016/j.biocon.2006.02.015)
- Geffroy B, Samia D, Bessa E, Blumstein DT. 2015 How nature-based tourism might increase prey vulnerability to predators. *Trends Ecol. Evol.* **30**, 755–765. (doi:10.1016/j.tree.2015.09.010)
- McNamara JM, Barta Z, Houston AI, Race P. 2005 A theoretical investigation of the effect of predators on foraging behaviour and energy reserves.



- Proc. R. Soc. B* **272**, 929–934. (doi:10.1098/rspb.2004.3037)
40. Sole LM, Shettleworth SJ, Bennett PJ. 2003 Uncertainty in pigeons. *Psychon. Bull. Rev.* **10**, 738–745. (doi:10.3758/BF03196540)
41. Higginson AD, Fawcett TW, Trimmer PC, McNamara JM, Houston AI. 2012 Generalized optimal risk allocation: foraging and antipredator behavior in a fluctuating environment. *Am. Nat.* **180**, 589–603. (doi:10.1086/667885)
42. Clark CW, Mangel M. 2000 *Dynamic state variable models in ecology*. New York, NY: Oxford University Press.
43. Hilton GM, Ruxton GD, Cresswell W. 1999 Choice of foraging area with respect to predation risk in redshanks: the effects of weather and predator activity. *Oikos* **87**, 295–302. (doi:10.2307/3546744)
44. Beale CM, Monaghan P. 2004 Behavioural responses to human disturbance: a matter of choice? *Anim. Behav.* **68**, 1065–1069. (doi:10.1016/j.anbehav.2004.07.002)
45. McNamara JM, Houston AI. 2009 Integrating function and mechanism. *Trends Ecol. Evol.* **24**, 670–675. (doi:10.1016/j.tree.2009.05.011)