PROCEEDINGS OF
THE ROYAL SOCIETY B

BIOLOGICAL SCIENCES

Trust your gut: using physiological states as a source of information is almost as effective as optimal Bayesian learning

Journal:	*Proceedings B*
Manuscript ID	RSPB-2017-2411.R1
Article Type:	Research
Date Submitted by the Author:	n/a
Complete List of Authors:	Higginson, Andrew; University of Exeter, Centre for Research in Animal Behaviour; Fawcett, Tim; University of Exeter, Centre for Research in Animal Behaviour Houston, Alasdair; University of Bristol, School of Biological Sciences McNamara, John; Bristol University, Mathematics
Subject:	Behaviour < BIOLOGY, Ecology < BIOLOGY, Evolution < BIOLOGY
Keywords:	behavioural gambit, cognition, computational costs, decision-making, information use, optimal foraging
Proceedings B category:	Behaviour

SCHOLARONE™
Manuscripts

1   Trust your gut: using physiological states as a source of information

2   is almost as effective as optimal Bayesian learning

3

4

5   A. D. Higginson[1], Tim W. Fawcett[1], Alasdair I. Houston[2], John M. McNamara[3]

6

7   [1] Centre for Research in Animal Behaviour, College of Life and Environmental Sciences, University

8   of Exeter, Exeter EX4 4QG, UK, adhigginson@gmail.com, t.w.fawcett@exeter.ac.uk.

9   [2] School of Biological Sciences, Life Sciences Building, Tyndall Avenue, Bristol, BS8 1TQ, U.K.,

10   a.i.houston@bristol.ac.uk

11   [3] School of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW,

12   john.mcnamara@bristol.ac.uk

13

14   **Running title**: Information in physiology

15   **Correspondence**: Andrew Higginson, Tel: +44 (0) 1392 724658, Fax: +44 (0)1392 724623

16   **Abstract**: 200 words

17   **Main text**: 5530 words

18   **References**: 49

19 **ABSTRACT**

20

21 Approaches to understanding adaptive behaviour often assume that animals have perfect information

22 about environmental conditions or are capable of sophisticated learning. If such learning abilities are

23 costly, however, natural selection will favour simpler mechanisms for controlling behaviour when

24 faced with uncertain conditions. Here we show that, in a foraging context, a strategy based only on

25 current energy reserves often performs almost as well as a Bayesian learning strategy that integrates

26 all previous experiences to form an optimal estimate of environmental conditions. We find that

27 Bayesian learning gives a strong advantage only if fluctuations in the food supply are very strong and

28 reasonably frequent. The performance of both the Bayesian and the reserve-based strategy are more

29 robust to inaccurate knowledge of the temporal pattern of environmental conditions than a strategy

30 that has perfect knowledge about current conditions. Studies assuming Bayesian learning are often

31 accused of being unrealistic; our results suggest that animals can achieve a similar level of

32 performance to Bayesians using much simpler mechanisms based on their physiological state. More

33 broadly, our work suggests that the ability to use internal states as a source of information about

34 recent environmental conditions will have weakened selection for sophisticated learning and decision-

35 making systems.

36

39   **INTRODUCTION**

40

41   "*Il meglio è nemico del bene*" [*The best is enemy of the good*]

42   Italian proverb

43

44   The study of animal decision-making has typically taken an optimisation approach in which the

45   animal is assumed to have perfect knowledge of current and long-term conditions [1–4]. In reality,

46   animals will be uncertain about conditions [5]. Such uncertainty can be incorporated into evolutionary

47   models using Bayes' rule, which updates knowledge given new information in a logically consistent

48   way [6,7], invoking the behavioural gambit [8] that animals will behave *as though* they can perform

49   Bayesian calculations [9,10]. However, it remains unclear how most animals could approximate

50   Bayesian learning without invoking implausible computational abilities or excessively costly

51   physiological or cognitive mechanisms that would require a large brain. For a mechanism to be

52   favoured by selection, there needs to be sufficient advantage to the animal in terms of reproductive

53   success to offset the costs of the mechanism. In many situations a simpler but less accurate

54   mechanism, determining a 'rule of thumb' or heuristic, might be advantageous if it has a smaller cost

55   [11]. An example is simple learning rules based on a linear operator [12]. Such rules may also be

56   more robust than Bayesian learning, in that their performance is less affected if information is

57   imperfect [12,13]

58

59   One of the best-studied situations in decision-making is searching for food [4,14,15]. Described rules

60   of thumb include the 'two-strikes' rule that bees (*Bombus lapidarius*) appear to follow in making

61   patch-quitting decisions [16]; the animal acts as though it has a fixed memory window for foraging

62   success, such as remembering whether or not it found food on the last few occasions that it looked.

63   Another example is the constant time in patches used by caddis fly larvae (*Plectrocnemia conspersa*)

64   [17]; here the animal acts as though it keeps track of time and ignores changes in conditions. Both

65   methods may lead to behaviour that is similar to a more sophisticated system that tracks food

66   availability explicitly [16,17].

67

68   To behave optimally in different conditions, the animal needs some way of assessing current

69   conditions. In the case of foraging, the animal discovers food items stochastically, which does not

70   necessarily reflect the overall food abundance at that point in time. Animals therefore need some way

71   to integrate past events, but acquiring and processing information in a Bayesian way is likely to be

72   costly [18]. Instead, natural selection could exploit the fact that animals have internal states that are a

73   potential source of information about conditions. All else being equal, energetic reserves tend to

74   increase if food is abundant and fall if food is scarce. Since conditions are positively autocorrelated

75   over time in most natural environments, conditions in the recent past are informative of current

76   conditions [5]. As such, reserves could act as a physiological 'memory' of environmental conditions

77   and so indicate current conditions [19].

78

79   Here, we show that energy reserves, a physiological state, provides a simple yet surprisingly effective

80   cue to decide how intensively to forage for food. For clarity, we use a simple model of survival in a

81   fluctuating environment (i.e. the generalised risk allocation model of [20]), where food availability

82   varies over time. We characterise the animal's environment in terms of the distribution, variability

83   and abundance of food items. We investigate under what conditions we expect animals to behave as

84   though they have sophisticated learning mechanisms for assessing current conditions, when they

85   should have simpler mechanisms, and when they should ignore fluctuations in conditions altogether.

86   To predict the outcome of natural selection it would be necessary to quantify the cost of mental

87   mechanisms, but this is currently not possible. We therefore compare the survivorship of various

88   candidate mechanisms to understand when sophisticated mechanisms give large benefits, in which

89   case animals are unlikely to have simple mechanisms. We find that, across a wide range of situations,

90   a strategy based only on the level of reserves performs almost as well as optimal Bayesian learning,

91   despite being much simpler, because reserve level acts as a memory. We discuss how such

92   mechanisms may operate in non-foraging contexts too, and suggest that physiological states acting as

93   'memories' may be ubiquitous.

94

95   **THE MODEL**

96

97   We are interested in the foraging strategy that maximizes survival in a temporally changing

98   environment where death can occur through starvation or predation. One possible response to harsh

99   conditions is to cease activity and wait for better times but the consequences of this for the forager's

100 survival and future state will depend on its current reserves. We therefore use a state-dependent model

101 in which the optimal action is allowed to depend on both the current conditions and the current level

102 of reserves. We model behaviour over a long sequence of discrete time steps. The animal and its

103 environment are characterized by two states: its level of reserves x ($x \geq 0$) and the current

104 environmental conditions E where food availability is higher in good conditions ($E=G$) than bad

105 conditions ($E=B$). Food availability differs only in the maximum probability of finding food when

106 foraging (γ_G and γ_B, where $\gamma_G \geq \gamma_B$).

107

108 The food availability of the environment is assumed to fluctuate over time. Incorporating

109 environmental heterogeneity into models of adaptive behaviour requires the inclusion of an

110 environmental state variable [5]. Often we can capture sufficient complexity with just two possible

111 environmental states A and B, such as high and low food availability. Next, we characterise stochastic

112 transitions between the two environmental states. The simplest case is where the probability of

113   transition (per unit time) between states depends only on the current state. At the end of a time step,

114   we assume that the environment changes from the current conditions E to the alternative conditions

115   with probability λ_E. Thus a good environment becomes a bad environment with probability λ_G, while a

116   bad environment becomes a good environment with probability λ_B. The duration of good and bad

117   periods both follow a geometric distribution whose mean is the reciprocal of the transition

118   probabilities, which we term t_G and t_B respectively. Note that this environment will show positive

119   temporal autocorrelation if $\lambda_B + \lambda_G < 1$ because then conditions are more likely to stay the same than

120   to change [5].

121

122   The aspect of behaviour we are interested in is foraging intensity, which we call f $(0 \leq f \leq 1)$.

123   Increasing f increases the probability of finding food but also increases exposure to predators and

124   hence the probability of being attacked. We assume that while the animal is not foraging, it is safe

125   from predation. We also assume that predation risk when foraging increases with energy reserves x

126   because of decreasing manoeuvrability [21]. Regardless of the exact cost, some cost needs to be

127   assumed if long-term adaptive fat levels are to be stable [22]. In a given time step, the probability of

128   mortality of the animal due to predation (μ) is given by

129    $$\mu(x, f) = f^c d \left(\frac{1}{2} + \frac{x}{s} \right) \tag{1}$$

130   where c controls how the risk increases with f, d is the maximum probability of predation attack, and s

131   is the maximum reserve level . We assume that the forager uses m units of energy per time step on

132   metabolism and finds a food item with probability $\gamma_E f$. For computational reasons there is some

133   variance in the energy content of food items (see Online Appendix): food items contain either b_1 or b_2

134   units of energy; for the results shown in the main text we assume that items with energy $b_1 = 5$ and

135    $b_2 = 6$ occur with equal probability. The reserves at the next time step are therefore

136    $$x_{t+1} = x_t + b_j - m$$

137   after a successful discovery of food item of type j $(j=1,2)$, and

138    $$x_{t+1} = x_t - m$$

139   after a failure to find food. If $x_t = 0$ then $x_{t+1} = 0$ because the animal is dead.

140

141   A strategy specifies how f depends on circumstances (e.g. reserves, information). We find optimal

142   strategies of various classes, all of which minimise the mortality rate and so maximise the survival

143   probability over a long time period. The classes of strategy differ in the constraints on the information

144   available to the forager. Where the current environmental state E is known (perfect information) this

145   is the generalised risk allocation model [20]. Where E is not known, the forager may be able to

146   estimate it based on available cues. To model this, we include a state variable ρ to represent the

147   forager's estimated probability that conditions are currently good (i.e. that $E = G$). Here we find the

148   optimal strategy f^* from two classes of strategy in which information is imperfect: (i) the animal

149   estimates the probability ρ that conditions are currently good directly from its foraging experiences,

150   using Bayesian updating; (ii) the animal does not monitor its foraging experiences directly but is

151   sensitive to its current energy reserves, and can take into account the fact that the level of reserves is

152   informative of recent conditions to estimate ρ. Assuming that the forager is optimally adapted to

153   minimize its long-term mortality rate, we use dynamic programming to find optimal solutions given

154   the constraints on information (see Appendix A). We set other parameter values (m, d, c, b_j) such that

155   the risk of mortality over some long time period is realistic. If each time step is thought of as around

156   one hour then 2000 time steps represent around 100 days of winter, over which the animals try to

157   survive. Small birds in temperate regions survive winter with 50-70% probability [23–25], so we tune

158   the parameters such that the survival at the baseline parameter values is around this range. As

159   mortality is far from both zero and one this ensures that the model can make clear predictions about

160   the effects of the parameter values of interest on the performance of the various strategies.

161

162   We compare the performance of these constrained optimal strategies to two other classes of strategy

163   that would be optimal if the environmental conditions were unchanging:

164   (1) A 'pessimistic' class of strategy that behaves as though the food availability is constantly low (γ_B).

165   (We do not show results for the alternative 'optimistic' strategy that behaves as though food

166   availability is constantly high (γ_G), because it performs very poorly in all non-trivial conditions.) .

167   (2) An optimally biased strategy that behaves as though the food availability is high with a fixed

168   probability and low otherwise, where the fixed probability is that which is optimal, and so will have

169   been naturally selected for in the absence of any attempt to track food availability.

170

171   Thus, in summary we compare the performance of five classes of strategy:

- **Perfect (P):** Forager has perfect knowledge about current food availability.

- **Bayesian (L):** Forager uses Bayes' theorem to estimate current food availability directly from its foraging experiences.

- **Reserves (R):** Forager does not monitor its foraging experiences but can base its decisions on its current reserve level; note that, through natural selection, the response to reserves will be influenced by the conditional probability that food availability is high given the reserve level.

- **Pessimist (S):** Forager behaves as though the current food availability is always low.

- **Optimal bias (U):** Forager behaves as though the current food availability is high with a fixed probability ρ^*, which is the estimate that minimizes the long-term mortality rate.

181

182   For each class, we find the optimal foraging strategy as a function of reserves and information state.

183   We then assess the resulting survival over 2000 time steps starting from the stationary distribution of

184   x in the population. To do this, we simulate a population following the optimal strategy until the

185   distribution of individuals stops changing, rescale so the size of the population is unity, and then run

186   for 2000 time steps to determine the survival probability $Q(i)$, where i indicates one of the strategy

187   classes as shown above. All parameters and their baseline values are shown in Table 1.

188

189   **RESULTS**

190

191   When using the reserve-based strategy (class R) the probability that conditions are good as a function

192   of reserves x is shown in Figure 1. For all parameter settings the probability follows a sigmoid curve,

193   with a low probability that conditions are good at low reserves and a high probability at high reserves,

194   because reserves gradually build up when food is abundant and decrease when food is scarce. The

195   curve shifts to the right as the difference between γ_G and γ_B increases because the optimal strategy is to

196   store more reserves in good conditions to prepare for bad conditions. The steepness of the sigmoid

197   curve depends on the fluctuation rate (Figure B1).

198

199   The optimal foraging intensity f^* for all five strategy classes is shown in Figure 2 for the baseline

200   parameter values (with the differences in foraging intensity plotted in Figure B2). As we have shown

201   previously [20] there is a crossover point in the optimal intensity of foraging under perfect

202   information f^*_P (grey lines), with more intense foraging when food availability is low if reserves are

203   low [$f^*_P(x,B) > f^*_P(x,G)$ when $x < 30$], but less intense foraging when food availability is low if reserves

204   are high [$f^*_P(x,B) < f^*_P(x,G)$ when $x \geq 30$]. A pessimist has f^*_S that is too high because it doesn't expect

205   good conditions to occur at all. For the reserve-based optimal strategy, foraging intensity f^*_R is similar

206   to $f^*_P(x,B)$ when reserves are low and closer to $f^*_P(x,G)$ when reserves are high (cf. grey and dotted

207   lines). This is intuitive, because the lower the reserve level, the more likely it is that conditions are

208   bad, hence the animal should behave as though conditions are bad; whereas if reserves are high it is

209   likely that conditions are good, hence the animal should behave as though conditions are good. For

210   the Bayesian learning strategy, f^*_L is similar to $f^*_P(x,B)$ when the posterior probability that conditions

211   are currently good ρ is zero and similar to $f^*_P(x,G)$ when ρ is unity, with a gradual change in f^*_L for

212   intermediate ρ (Figure B3).

213

214   We assess the probability of surviving 2000 time steps for each optimal strategy under various

215   conditions (Figure 3; shown for $\gamma_B = 0.25$ and $\gamma_G = 0.75$, for other values see Figure B4). For clarity we

216   first show survival under perfect knowledge (P, which always does best) and then the differences

217   between the various strategies. Survival always increases with the mean duration of good periods and

218   decreases with the mean duration of bad periods because mortality mostly occurs in bad periods, and

219   the length of these therefore determines survival (Figure 3a, Figure B4a-e). Survival decreases as the

220   difference in food availability increases because that determines the severity of bad periods, except

221   that survival increases with the difference in food availability if conditions are good most of the time

222   (cf. Figure B4a, d), because the increased rate of gain in good periods more than compensates for this

223   and risk allocation has a large benefit.

224

225   In general, the difference in survival between perfect knowledge (*P*) and the information-constrained

226   strategies (*L, R*) is much less than 5% for most conditions. *L* (Bayesian learning) does worst compared

227   to *P* when periods are short because it is impossible to learn fast enough to perform risk allocation

228   effectively (Figure 3b, B4f-j); this is exacerbated when food availability differs markedly between

229   good and bad conditions (Figure B4j). Across parameter space there is strikingly little difference

230   between *L* and the reserve-based strategy *R* (Figure 3c, B4k-o), except when periods are moderately

231   short (around 20 time steps) and the difference in food availability between conditions is very large

232   (Figure B4o). *R* does much better than *U* (optimal bias) when periods are long and of roughly equal

233   duration, because then it is most important to do the correct thing (Figure 3d, B4p–t). The optimal

234   estimate ρ^* under the *U* strategy is always smaller than the actual ρ (Figure B5). This is because

235   eating too much in good conditions is less deleterious than eating too little in poor conditions.

236

237   In Figure 4 we clarify the conditions under which a learning (*L*) or reserve-based (*R*) strategy should

238   evolve, under the arbitrary assumption that *L* is twice as costly as *R*. We expect sophisticated learning

239   to be worth this additional cost when periods are moderately short and food availability changes

240   greatly (bottom-left of Figure 4b, d) or when the fluctuations are subtle and infrequent (top-right

241   Figure 4a). We expect the reserve-based strategy to be favoured if the world is not predominantly

242   poor or rich (i.e. along the main diagonal of Figure 4) and does not change too quickly (not the

243   bottom-left). This is because *R* does not adapt fast enough when conditions turn bad and so the animal

244   is more likely to die; in this situation, either *L* or *U* does better. In all other cases, decisions based

245   solely on the current reserve level allow the animal to perform almost as well as a sophisticated

246   Bayesian learning strategy, with differences less than 1% in most of parameter space, and 0.04% for

247   the baseline parameter values.

248

249   For the results above we assumed that $\gamma_G + \gamma_B = 1$. However, the difference between *L* and *R* remains

250   small for almost all combinations of γ_G and γ_B (Figure B6). We have also confirmed that the results are

251   not sensitive to our assumptions about the variance in energy consumption over time (Figure B7). We

252   did this by increasing the energy content of food items b_j while decreasing their rate of discovery γ_G

253   and γ_B, such that the total amount of energy in the environment remained constant but the variance

254 increased (implying longer periods without eating). The results are almost unchanged across the full

255 range of the proportion of food that occurs under good conditions (Figure B7).

256

257 In addition to having imperfect knowledge about current conditions, a forager's perception of the

258 pattern of environmental change may be prone to error. This may be the case because of dispersal or

259 because anthropogenic change is altering environments faster than animals can adapt [26]. To

260 investigate this, we assess the performance of the same five strategy classes in an environment that

261 fluctuates on a different timescale from that to which the forager is adapted. In Figure 5 we present

262 the survivorship relative to the *P* case when the strategy is mismatched for the change probabilities

263 (for absolute values see Figure B8). Overall the survival of *P* is poorer than that of *R* and *L* if the

264 perceived rate of environmental change is different to the actual rate. This occurs because the optimal

265 decision depends on the forager's current state and its expectations about the future; if those

266 expectations are wrong then performance will be poor. This is ameliorated if the forager can adjust its

267 expectations via learning or other changes in state, which are influenced by the real conditions. At the

268 extreme, if the forager expects periods to be long then the performance of *P* worsens as the actual

269 period durations decrease (Figure B8), whereas performance improves for *L* and *R* (Figure 5g. If the

270 actual duration of periods is much longer than expected then it would be better to act as though

271 conditions are always poor (*S*) (Figure 5b, d, f), but there is always a range of perceived durations

272 where *L* and *R* outperform *P*. When the expected durations are quite inaccurate, the actual durations

273 determine whether *R* outperforms *L* or vice versa: if the actual durations are long, reserves become a

274 reliable cue of current conditions (Figure 5a, d), whereas if the actual durations are short the Bayesian

275 strategy performs better (Figure 5f, h).

276

277 The maintained reserve level is similar under *L* and *R* but slightly shifted to lower reserves compared

278 to *P* for baseline parameter values (Figure B9). Storing a lower level of reserves is predicted across

279 most of parameter space (Figure B10), except where there is a very strong difference in food

280 availability between good and bad conditions and conditions change slowly (Figure B10b, f) or when

281 conditions are more often good (Figure B10d, h). Across all of parameter space, reserves under *L* are

282 closer to those under *P* than *R*, explaining the slightly better performance of *L*.

283

284 **DISCUSSION**

285

286 The need to track and respond appropriately to environmental conditions generates an important

287 selective pressure on sensory and cognitive systems. Animals typically do not have perfect knowledge

288 [27]. While foraging they may learn about the current food availability, but because food discovery is

289 stochastic there is uncertainty. Given this uncertainty, animals are likely to have decision rules that

290 perform well in most conditions [8,11,28]. The level of sophistication of these rules will depend on

291 their associated costs and the benefit of tracking the environment. Here we have compared the

292 performance of a number of implementations of possible foraging mechanisms in an environment

293 with fluctuating food availability. Our findings suggest that a Bayesian learning strategy—a

294 commonly used paradigm in research on learning [6,9,12,29] but one which is arguably implausible

295 for real organisms [6,7] (but see [9,10])—is unlikely to evolve under most conditions, because a

296 simpler decision rule based solely on current energy reserves could allow the animal to perform

297 almost as well. The greatest benefit to distinguishing between conditions occurs when the

298 environment fluctuates slowly, but in this case there is ample time for energetic reserves to respond to

299 current conditions before they change, and so most of the time the reserve level will be a sufficiently

300 reliable indicator of current conditions. The ability to behave appropriately using only energy reserves

301 as a cue is likely to have greatly reduced the selective pressure for sophisticated learning systems.

302

303 Bayesian learning might still be advantageous if other classes of strategy are very expensive, if there

304 is a strong difference between conditions (making it more important to adjust behaviour accordingly)

305 and if conditions change sufficiently fast that reserves are an unreliable cue to current conditions. This

306 perspective suggests that animals in strongly and quickly fluctuating environments might be better at

307 learning, which contradicts the suggestion that learning is favoured under intermediate rates of change

308 [30,31]; note that these previous studies did not consider simpler alternative mechanisms. Strikingly,

309 we predict that animals should be insensitive to some types of environmental fluctuations, such as if

310 the fluctuations are not very large, or fluctuations are very quick, or if the world is usually in one state

311 or the other. The latter result is predicted because if food conditions are dominated by one level of

312 availability, then animals can just behave as though this is always the case. With fast changes or

313 changes of small magnitude it is less important to be sensitive to changes in food availability because

314 current conditions do not provide much information about future conditions [20]. In experiments that

315 have found no response to changing conditions [32], it is important to consider whether the study

316 organism is adapted to an environment in which there is limited benefit of responding to changes.

317 In some situations, such as when the level of food availability changes frequently, it may be that the

318 animal should do the same thing in the different conditions [20]. In such cases, an evolved mechanism

319 may implement some simpler rule that does not try to track conditions (U). This may underlie state-

320 dependent valuation of food sources because an animal's state may reflect what conditions were

321 generally like when particular sources were exploited [33].

322

323 The reserve-based strategy class may be the most likely evolutionary outcome in most situations.

324 Even the simple rule (U) requires a basic sensitivity to reserve level to avoid starvation, and the

325 reserve-based strategy is unlikely to involve significant additional costs. Thus, animals will not

326 necessarily carry the level of reserves predicted by standard models that assume perfect knowledge or

327 Bayesian learning, but instead may make systematic deviations because they are using reserves as a

328　source of information. We predict that these deviations will be positive (more reserves than predicted

329　by perfect information models) when conditions change slowly (Figure B10a) but negative when

330　conditions change quickly (Figure B10c). There may be no need for a cognitively encoded memory of

331　recent foraging experiences; natural selection will simply exploit information by favouring an

332　adaptive response to energetic reserves. In effect, the animal's reserves act as a physiological memory

333　of past events. This suggestion could be tested empirically in systems where foraging experiences can

334　be decoupled from the perceived level of reserves, for example through experimental manipulation of

335　hormones such as ghrelin and leptin that are involved in the regulation of feeding behaviour. By

336　manipulating hormone levels and foraging experiences independently of each other, it should be

337　possible to determine whether foraging behaviour is controlled by a cognitively encoded memory, a

338　reserve-based memory or some combination of the two.

339

340　Lea and colleagues [34] assessed the performance of cognitive mechanisms for solving the explore-

341　exploit trade-off. They found that a simple decision rule can perform better than more sophisticated

342　strategies in some conditions, such as where there is insufficient time to learn about current

343　conditions, which is comparable to the poor performance of our Bayesian learning strategy when

344　fluctuations are frequent. However, the choice of foraging currency is likely to be crucial for the

345　insights obtained [15], and often maximisation of net rate as assumed by Lea et al. [34] will make

346　substantially different predictions to currencies that incorporate the risk of mortality that most

347　foragers face [15,35]. Future theoretical work should consider how a foraging rule based on

348　physiological state, such as a hormone level, performs relative to a cognitive mechanism that attempts

349　to learn about the level of predation risk from direct experiences (e.g. sightings of predators).

350

351　Learning rules that maximise long-term reward rate by learning about conditions can perform much

352　better than ignorant rules [36,37]. But these rule sets have not accounted for the fact that internal state,

353　such as the level of energy reserves or body temperature, always provides animals with some

354　information and we expect natural selection to have formed strategies that exploit all sources of

355　information about the external conditions. Several models have shown that an animal's state should

356　influence decision-making to the extent that behaviour may appear irrational [38–41]. Here, we have

357　identified that the effect of energetic reserves may be more complex still: animals with equal levels of

358　reserves may differ in their response if they are adapted to different environments, such as different

359　rates of change, because of how this affects the information content [20].

360

361　The marginal value theorem predicts that the marginal capture rate for leaving patches of prey should

362　be higher when the overall prey abundance is higher, but this is often not observed [42]. A simple rule

363　of thumb of a constant giving-up time results in behaviour that approximates the optimal solution

364　much of the time [17,43,44]. Such a rule may be driven by some internal physiological state,

365   involving feedback from the gustatory system, which reflects the time since the last prey item was

366   consumed. Nonacs [45] showed that including a forager's energy reserves alters the predictions of the

367   marginal value theorem, but he also assumed animals could keep track of foraging success in a perfect

368   way. We suggest that a better approach may be to model a gustatory state, such as stomach contents,

369   which the animal can use as a cue of foraging success. Our reserve-based approach could be used to

370   incorporate information constraints in many established models of animal behaviour and decision-

371   making.

372

373   There is currently much interest and concern about the ability of organisms to cope with human-

374   induced rapid environmental change [46]. Such rapid changes will cause there to be a mismatch

375   between the conditions that animals have evolved to deal with and those they actually experience. Our

376   results (Figure 5) suggest that the details of how the environment has changed will determine how

377   organisms respond. Interestingly, if environmental change causes conditions to fluctuate more quickly

378   or more slowly than in the evolutionary past—for example, because it leads to more extreme weather

379   patterns—then organisms that can perceive the current conditions directly (P) may in fact perform

380   worse than those that use simple rules to estimate current conditions (Figure 5). Which strategy class

381   performs best depends on whether fluctuations are more or less frequent: if conditions now change

382   more quickly than in the past then learning does best (left of Figure 5c), whereas if conditions change

383   more slowly then simpler (e.g. reserve-based) strategy classes not based on learning do best (right of

384   Figure 5c).

385

386   We have shown that, in a foraging context, a behavioural strategy based only on an internal

387   physiological state (R) can perform so well that more sophisticated strategies, such as learning directly

388   from foraging outcomes (L) or accurately perceiving current conditions (P), might not provide

389   sufficient advantages to offset their costs. It is striking that a reserve-based strategy is more robust to

390   error in the pattern of environmental fluctuations than a rule based on perfect information about

391   current food availability. Therefore, if the information about the environment is unreliable, we expect

392   selection to favour simpler strategy classes. So far, we have been unable to prove that our

393   methodology for finding the best-performing reserve-based strategy actually converges on the global

394   optimum, rather than a local optimum (See Figure B11). However, if it is just a local optimum, then

395   our conclusions would be strengthened: the performance of the reserve-based strategy at its global

396   optimum (elsewhere in N-dimensional space) would be even better than the one we have described

397   here, and hence even closer to the performance of the Bayesian learning strategy.

398

399   Similar principles could well apply in other (non-foraging) contexts: any physiological or

400   psychological state variable that is altered by experience might function as an efficient integrator (a

401   'memory') of past experiences. An obvious candidate is emotions and moods, which have been

402 modelled mechanistically [47] and may help an animal to adjust its behaviour adaptively when

403 conditions are uncertain [48,49]. In fact, in non-foraging contexts, the state variable may have greater

404 flexibility to act as a cue because (unlike energy reserves) the animal doesn't necessarily depend on it

405 for survival, so it could potentially evolve to be more informative than energy reserves are in the

406 foraging case. One intriguing possibility is that emotional states were initially unavoidable

407 consequences of levels of neurotransmitter activity, but have been modified by selection to provide

408 more reliable information about recent experiences and thereby influence cognitive decisions. If the

409 principle we have highlighted applies to most physiological states then organisms may often appear to

410 be cognitively sophisticated despite basing their decisions on relatively simple mechanisms. Since

411 internal states can summarise a great deal of information about the environmental conditions, they

412 will reduce the selective pressure to learn directly from the immediate outcomes of decisions. Animals

413 are therefore likely to be cognitively unsophisticated when they are able to perform well using simple

414 mechanisms.

415

416 **ETHICS STATEMENT**

417 This work is entirely theoretical so there are no ethical considerations.

418 **DATA ACCESSIBILITY STATEMENT**

419 The code used to generate the results have been uploaded as part of the supplementary material.

420 **COMPETING INTERESTS STATEMENT**

421 We have no competing interests

422 **AUTHORS' CONTRIBUTIONS**

423 All authors conceived the study. ADH designed the study, wrote the code and generated the results.

424 All authors wrote the manuscript.

425 **ACKNOWLEDGMENTS**

REFERENCES

431

432   1.   Charnov, E. E. L. 1976 Optimal foraging: the marginal value theorem. *Theor. Popul. Biol.* **9**,
433   129–136. (doi:10.1016/0040-5809(76)90040-X)

434   2.   McNamara, J. M. & Houston, A. I. 1990 The value of fat reserves and the tradeoff between
435   starvation and predation. *Acta Biotheor.* **38**, 37–61.

436   3.   Parker, G. A. & Maynard Smith, J. 1990 Optimality theory in evolutionary biology. *Nature*
437   **348**, 27–33.

438   4.   Pyke, G. H., Pulliam, H. R. & Charnov, E. L. 1977 Optimal Foraging: A Selective Review of
439   Theory and Tests. *Q. Rev. Biol.* **52**, 137. (doi:10.1086/409852)

440   5.   Fawcett, T. W., Fallenstein, B., Higginson, A. D., Houston, A. I., Mallpress, D. E. W.,
441   Trimmer, P. C. & McNamara, J. M. 2014 The evolution of decision rules in complex
442   environments. *Trends Cogn. Sci.* **18**, 153–61. (doi:10.1016/j.tics.2013.12.012)

443   6.   McNamara, J. M., Green, R. F. & Olsson, O. 2006 Bayes' theorem and its applications in
444   animal behaviour. *Oikos* **112**, 243–251. (doi:10.1111/j.0030-1299.2006.14228.x)

445   7.   Trimmer, P. C., Houston, A. I., Marshall, J. A. R., Mendl, M. T., Paul, E. S. & McNamara, J.
446   M. 2011 Decision-making under uncertainty: biases and Bayesians. *Anim. Cogn.* **14**, 465–76.
447   (doi:10.1007/s10071-011-0387-4)

448   8.   Fawcett, T. W., Hamblin, S. & Giraldeau, L.-A. 2012 Exposing the behavioral gambit: the
449   evolution of learning and decision rules. *Behav. Ecol.* **24**, 2–11. (doi:10.1093/beheco/ars085)

450   9.   J. Valone, T. 2006 Are animals capable of Bayesian updating? An empirical review. *Oikos*
451   **112**, 252–259. (doi:10.1111/j.0030-1299.2006.13465.x)

452   10.   Biernaskie, J. M., Walker, S. C. & Gegear, R. J. 2009 Bumblebees learn to forage like
453   Bayesians. *Am. Nat.* **174**, 413–23. (doi:10.1086/603629)

454   11.   Hutchinson, J. M. C. & Gigerenzer, G. 2005 Simple heuristics and rules of thumb: where
455   psychologists and behavioural biologists might meet. *Behav. Processes* **69**, 97–124.
456   (doi:10.1016/j.beproc.2005.02.019)

457   12.   Lange, A. & Dukas, R. 2009 Bayesian approximations and extensions: optimal decisions for
458   small brains and possibly big ones too. *J. Theor. Biol.* **259**, 503–16.
459   (doi:10.1016/j.jtbi.2009.03.020)

460   13.   Trimmer, P. C., McNamara, J. M., Houston, A. I. & Marshall, J. A. R. 2012 Does natural
461   selection favour the Rescorla-Wagner rule? *J. Theor. Biol.* **302**, 39–52.

462   14.   Stephens, D. W. & Krebs, J. J. R. 1986 *Foraging theory*. Princeton: Princeton University
463   Press.

464   15.   Houston, A. I. & McNamara, J. M. 2014 Foraging currencies, metabolism and behavioural
465   routines. *J. Anim. Ecol.* **83**, 30–40. (doi:10.1111/1365-2656.12096)

466   16.   Goulson, D. 2000 Why do pollinators visit proportionally fewer flowers in large patches?
467   *Oikos* **91**, 485–492. (doi:10.1034/j.1600-0706.2000.910309.x)

468   17.   Townsend, C. R. & Hildrew, A. G. 1980 Foraging in a patchy environment by a predatory net-
469   spinning caddis larva: A test of optimal foraging theory. *Oecologia* **47**, 219–221.
470   18.   Bowers, J. S. & Davis, C. J. 2012 Bayesian just-so stories in psychology and neuroscience.
471   *Psychol. Bull.* **138**, 389–414. (doi:10.1037/a0026450)
472   19.   Charnov, E. L. 1976 Optimal foraging: attack strategy of a mantid. *Am. Nat.* **110**, 141–151.
473   20.   Higginson, A. D., Fawcett, T. W., Trimmer, P. C., McNamara, J. M. & Houston, A. I. 2012
474   Generalized optimal risk allocation: foraging and antipredator behavior in a fluctuating
475   environment. *Am. Nat.* **180**, 589–603. (doi:10.1086/667885)
476   21.   Witter, M. S. & Cuthill, I. C. 1993 The ecological costs of avian fat storage. *Philos. Trans. R.*
477   *Soc. London, Ser. B* **340**, 73–92.
478   22.   Houston, A. I., Welton, N. J. & McNamara, J. M. 1997 Acquisition and maintenance costs in
479   the long-term regulation of avian fat reserves. *Oikos* **78**, 331–340.
480   23.   Schekkerman, H., Teunissen, W. & Oosterveld, E. 2009 Mortality of Black-tailed Godwit
481   *Limosa limosa* and Northern Lapwing *Vanellus vanellus* chicks in wet grasslands: influence of
482   predation and agriculture. *J. Ornithol.* **150**, 133–145. (doi:10.1007/s10336-008-0328-4)
483   24.   Gullett, P., Evans, K. L., Robinson, R. A. & Hatchwell, B. J. 2014 Climate change and annual
484   survival in a temperate passerine: partitioning seasonal effects and predicting future patterns.
485   *Oikos* **123**, 389–400. (doi:10.1111/j.1600-0706.2013.00620.x)
486   25.   Gullett, P. R., Hatchwell, B. J., Robinson, R. A. & Evans, K. L. 2015 Breeding season weather
487   determines long-tailed tit reproductive success through impacts on recruitment. *J. Avian Biol.*
488   **46**, 441–451. (doi:10.1111/jav.00560)
489   26.   Sih, A., Ferrari, M. C. O. & Harris, D. J. 2011 Evolution and behavioural responses to human-
490   induced rapid environmental change. *Evol. Appl.* **4**, 367–387. (doi:10.1111/j.1752-
491   4571.2010.00166.x)
492   27.   McNamara, J. & Houston, A. 1980 The application of statistical decision theory to animal
493   behaviour. *J. Theor. Biol.* **85**, 673–690. (doi:10.1016/0022-5193(80)90265-9)
494   28.   McNamara, J. M. & Houston, A. I. 2009 Integrating function and mechanism. *Trends Ecol.*
495   *Evol.* **24**, 670–675.
496   29.   Krebs, J. R., Kacelnik, A. & Taylor, P. 1978 Test of optimal sampling by foraging great tits.
497   *Nature* **275**, 27–31. (doi:10.1038/275027a0)
498   30.   Stephens, D. W. 1991 Change, regularity, and value in the evolution of animal learning.
499   *Behav. Ecol.* **2**, 77–89. (doi:10.1093/beheco/2.1.77)
500   31.   Dunlap, A. S. & Stephens, D. W. 2009 Components of change in the evolution of learning and
501   unlearned preference. *Proc. R. Soc. London B Biol. Sci.* **276**.
502   32.   Ferrari, M. C. O., Sih, A. & Chivers, D. P. 2009 The paradox of risk allocation: a review and
503   prospectus. *Anim. Behav.* **78**, 579–585. (doi:10.1016/j.anbehav.2009.05.034)
504   33.   McNamara, J. M., Trimmer, P. C. & Houston, A. I. 2012 The ecological rationality of state-

505   dependent valuation. *Psychol. Rev.* **119**, 114–119. (doi:10.1037/a0025958)

506   34.   Lea, S. E. G., McLaren, I. P. L., Dow, S. M. & Graft, D. A. 2012 The cognitive mechanisms of

507   optimal sampling. *Behav. Processes* **89**, 77–85. (doi:10.1016/j.beproc.2011.10.004)

508   35.   Higginson, A. D. & Houston, A. I. 2015 The influence of the food-predation trade-off on the

509   foraging behaviour of central-place foragers. *Behav. Ecol. Sociobiol.* **69**, 551–561.

510   36.   Gross, R., Houston, A. I., Collins, E. J., McNamara, J. M., Dechaume-Moncharmont, F.-X. &

511   Franks, N. R. 2008 Simple learning rules to cope with changing environments. *J. R. Soc.*

512   *Interface* **5**, 1193–202. (doi:10.1098/rsif.2007.1348)

513   37.   Hamblin, S. & Giraldeau, L.-A. 2009 Finding the evolutionarily stable learning rule for

514   frequency-dependent foraging. *Anim. Behav.* **78**, 1343–1350.

515   38.   Nevai, A. L., Waite, T. A. & Passino, K. M. 2007 State-dependent choice and ecological

516   rationality. *J. Theor. Biol.* **247**, 471–9. (doi:10.1016/j.jtbi.2007.03.029)

517   39.   Schuck-Paim, C., Pompilio, L. & Kacelnik, A. 2004 State-dependent decisions cause apparent

518   violations of rationality in animal choice. *PLoS Biol.* **2**, e402.

519   (doi:10.1371/journal.pbio.0020402)

520   40.   Trimmer, P. C. 2013 Optimal behaviour can violate the principle of regularity. *Proc. Biol. Sci.*

521   **280**, 20130858. (doi:10.1098/rspb.2013.0858)

522   41.   Houston, A. I., McNamara, J. M. & Steer, M. D. 2007 Violations of transitivity under fitness

523   maximization. *Biol. Lett.* **3**, 365–7. (doi:10.1098/rsbl.2007.0111)

524   42.   Marshall, H. H., Carter, A. J., Ashford, A., Rowcliffe, J. M. & Cowlishaw, G. 2013 How do

525   foragers decide when to leave a patch? A test of alternative models under natural and

526   experimental conditions. *J. Anim. Ecol.* **82**, 894–902. (doi:10.1111/1365-2656.12089)

527   43.   Hutchinson, J. M. C., Wilke, A. & Todd, P. M. 2008 Patch leaving in humans: can a generalist

528   adapt its rules to dispersal of items across patches? *Anim. Behav.* **75**, 1331–1349.

529   (doi:10.1016/j.anbehav.2007.09.006)

530   44.   Ydenberg, R. C. 1984 Great tits and giving-up times: Decision rules for leaving patches.

531   *Behaviour* **90**, 1–24.

532   45.   Nonacs, P. 2001 State dependent behavior and the Marginal Value Theorem. *Behav. Ecol.* **12**,

533   71–83. (doi:10.1093/oxfordjournals.beheco.a000381)

534   46.   Sih, A. 2013 Understanding variation in behavioural responses to human-induced rapid

535   environmental change: a conceptual overview. *Anim. Behav.* **85**, 1077–1088.

536   (doi:10.1016/j.anbehav.2013.02.017)

537   47.   Giske, J., Eliassen, S., Fiksen, Ø., Jakobsen, P. J., Aksnes, D. L., Jørgensen, C. & Mangel, M.

538   2013 Effects of the emotion system on adaptive behavior. *Am. Nat.* **182**, 689–703.

539   (doi:10.1086/673533)

540   48.   Trimmer, P. C., Paul, E. S., Mendl, M. T., McNamara, J. M. & Houston, A. I. 2013 On the

541   evolution and optimality of mood states. *Behav. Sci. (Basel).* **3**, 501–521.

542   49.   Nettle, D. & Bateson, M. 2012 The Evolutionary Origins of Mood and Its Disorders. *Curr.*

543   *Biol.* **22**, R712–R721. (doi:10.1016/j.cub.2012.06.020)

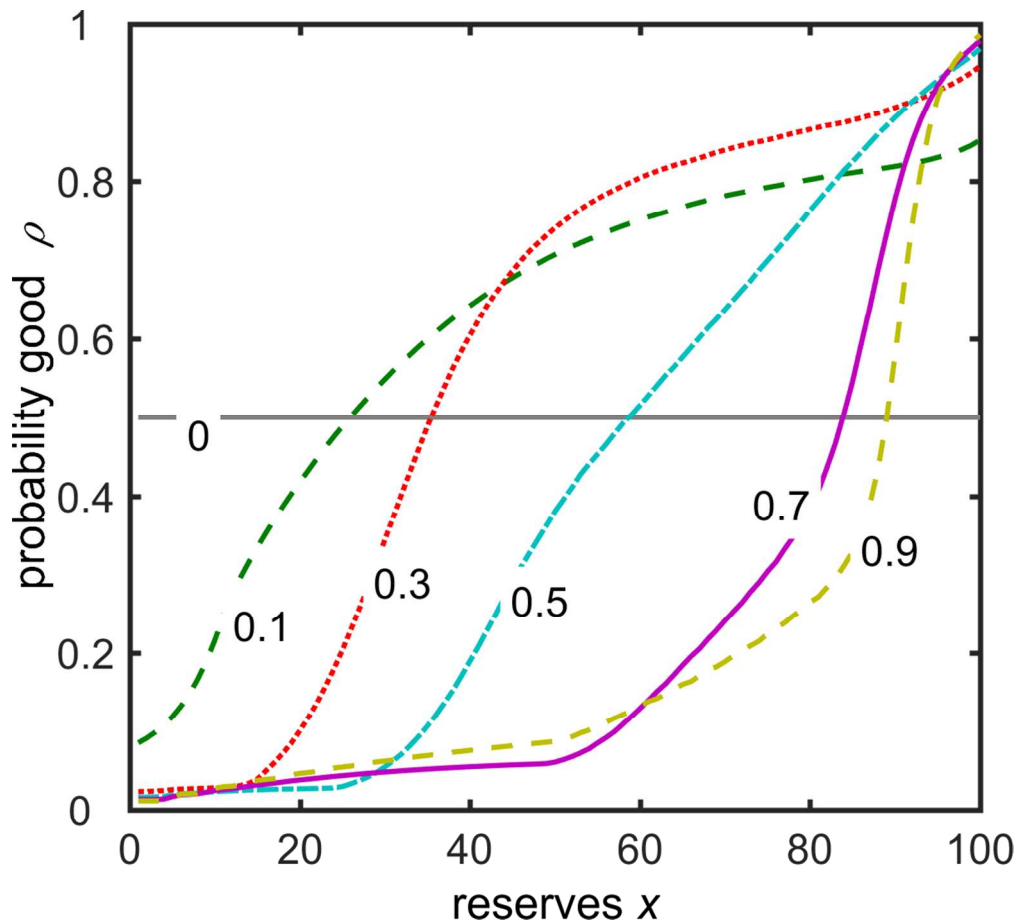544

545

546   **Figure Legends**

547   Figure 1: Probability that conditions are good given current reserves x under the reserve-based

548   strategy R, $\rho(G|x)$. The values on each line indicate the difference between conditions in food

549   availability $\gamma_G - \gamma_B$, where $\gamma_G + \gamma_B = 1$. Mean duration of bad and good periods: $t_B = t_G = 0.005$.

550

551   Figure 2: The optimal foraging intensity (f^*) as a function of energy reserves (x) for the perfect

552   information (grey lines: P_G solid, P_B dashed), reserve-based (R), pessimistic (S) and optimal bias (U)

553   cases. For the optimal Bayesian learning strategy (L), the values of f_L^* are intermediate to those for P_G

554   and P_B (shown in Figure B3).

555

556   Figure 3: Comparison of survival probabilities over 2000 time steps $Q(i)$ for the various methods

557   across parameter space (t_B and t_G shown on axes). We show results for the baseline probability of

558   finding food in the two conditions ($\gamma_B = 0.25$, $\gamma_G = 0.75$); for other values see Figure B4. (a) Survival

559   under perfect information (P). (b, c, d) Differences in survival between strategies (b) P and L

560   (Bayesian learner); (c) L and R (reserve-based); (d) R and U (optimal bias). Note the different scales

561   of the vertical axes.

562

563   Figure 4: Best strategy class under varying costs of implementation. For comparison, we assumed that

564   the reserve-based (R) strategy pays an arbitrary survival cost ψ per time step whereas the Bayesian

565   learning (L) strategy pays $k\psi$; the optimal bias (U) and pessimistic (S) strategies pay no cost. We

566   assume that the P strategy cannot evolve. The shaded regions indicate which strategy (L, R or U) has

567   highest survival given these costs for (a, b) $\psi = 0.001$ and (c, d) $\psi = 0.004$, and for (a, c) $\gamma_B = 0.35$,

568   $\gamma_G = 0.65$ and (b, d) $\gamma_B = 0.15$, $\gamma_G = 0.85$, with $k = 2$ in all panels. The results do not qualitatively depend

569   on the values of ψ or k, with an intuitive gradual shrinking of the L region as k (relative cost of

570   Bayesian learning compared to a reserve-based strategy) increases.

571

572   Figure 5: Probability of surviving 2000 time steps $Q(i)$ when the actual fluctuation rates differ from

573   those the animal is adapted to. Survival is plotted as the difference compared to survival under perfect

574   knowledge (P); negative values imply lower survival as a result of imperfect knowledge about current

575   conditions (different lines for strategies L, R, S and U). Left column (a, c, e, g) shows relative survival

576   as a function of the actual mean duration of periods (x-axis) when following the optimal strategy for

577   the duration shown on the panels. Right column (b, d, f, h) shows survival as a function of the mean

578   duration of periods to which the animal is adapted (x-axis) in four environments with different actual

579   mean durations (shown on the panels). (a, b) $t_B = t_G = 2000$; (c, d) $t_B = t_G = 200$; (e, f) $t_B = t_G = 20$; (g, h)

580   $t_B = t_G = 2$.

581 **Table 1:** Parameters in the model and their default values

Symbol	Description	Value
s	Maximum level of reserves	100
m	Energy use per unit time	1
b_j	Energy in food item type j	5, 6
d	Magnitude of predation risk	0.002
c	Power of relationship between foraging and predation risk	2
ψ	Survival cost per time step for reserve-based strategies	0.001, 0.004
k	Relative cost of Bayesian compared to reserve-based strategy	2
γ_E	Probability of finding food per unit time spent foraging in environment in condition E	$\gamma_G = 0.7$, $\gamma_B = 0.3$
λ_E	Probability that environment in condition E changes to the other condition	$\lambda_G = 0.01$, $\lambda_B = 0.01$
t_E	Mean number of time steps for which environment stays in condition E ($t_E = 1/\lambda_E$)	$t_G = 100$, $t_B = 100$

582

583

Figure 1: Probability that conditions are good given current reserves x under the reserve-based strategy R, $\rho(G|x)$. The values on each line indicate the difference between conditions in food availability $\gamma_G - \gamma_B$, where $\gamma_G + \gamma_B = 1$. Mean duration of bad and good periods: $t_B = t_G = 0.005$.
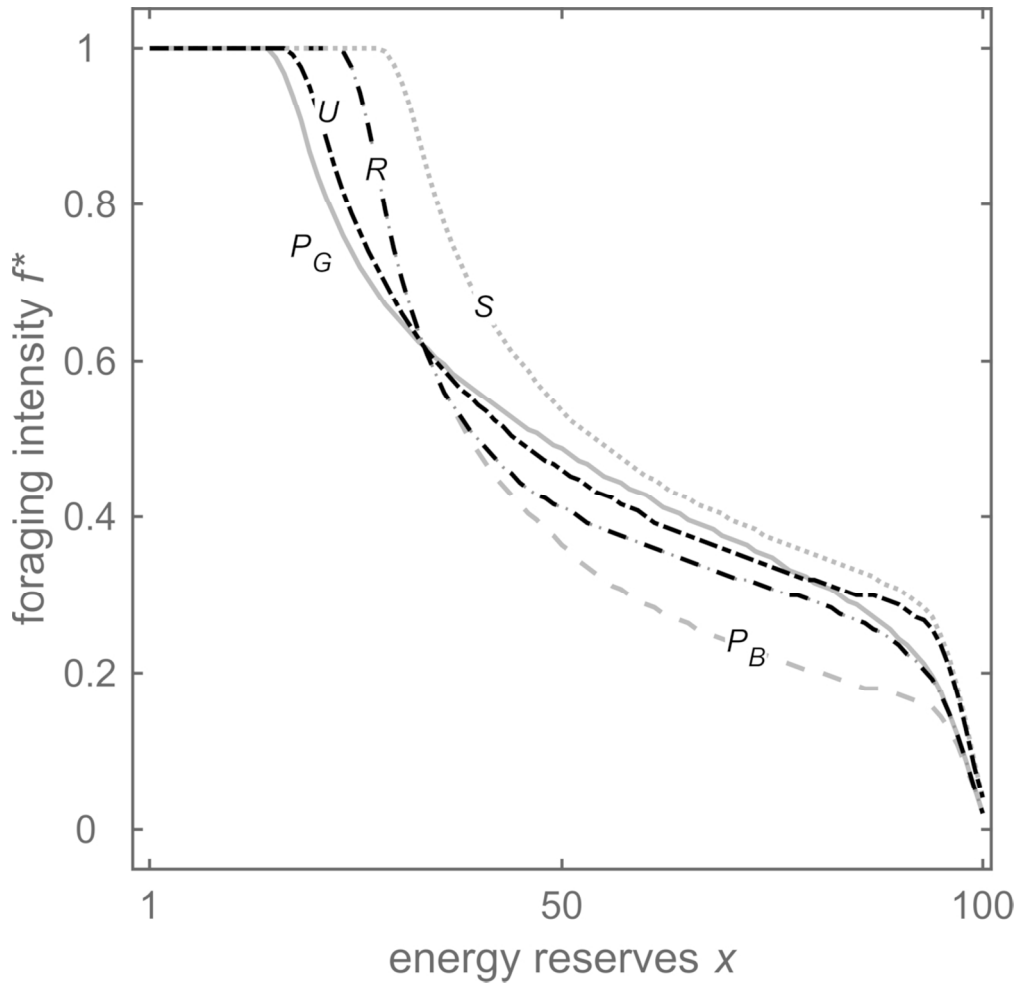
102x92mm (300 x 300 DPI)

Figure 2: The optimal foraging intensity (f*) as a function of energy reserves (x) for the perfect information (grey lines: PG solid, PB dashed), reserve-based (R), pessimistic (S) and optimal bias (U) cases. For the optimal Bayesian learning strategy (L), the values of fL* are intermediate to those for PG and PB (shown in Figure B3).
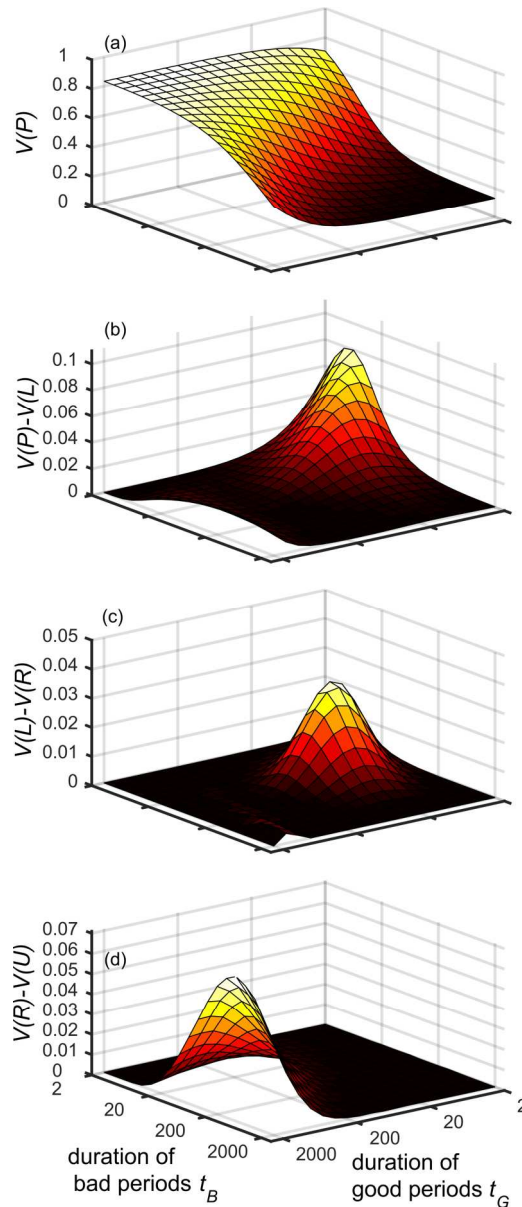
109x107mm (300 x 300 DPI)

Figure 3: Comparison of survival probabilities over 2000 time steps Q(i) for the various methods across parameter space (tB and tG shown on axes). We show results for the baseline probability of finding food in the two conditions (γB=0.25, γG=0.75); for other values see Figure B4. (a) Survival under perfect information (P). (b, c, d) Differences in survival between strategies (b) P and L (Bayesian learner); (c) L and R (reserve-based); (d) R and U (optimal bias). Note the different scales of the vertical axes.
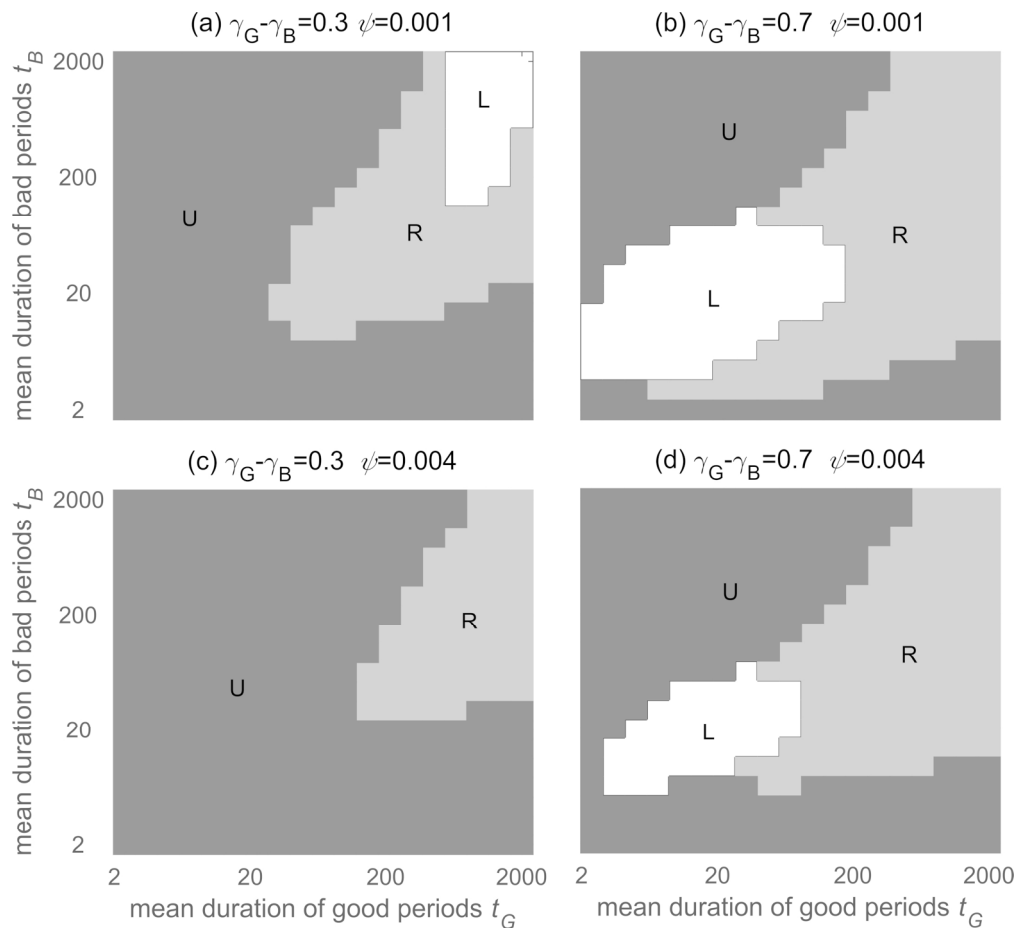
96x225mm (300 x 300 DPI)

Figure 4: Best strategy class under varying costs of implementation. For comparison, we assumed that the reserve-based (R) strategy pays an arbitrary survival cost ψ per time step whereas the Bayesian learning (L) strategy pays kψ; the optimal bias (U) and pessimistic (S) strategies pay no cost. We assume that the P strategy cannot evolve. The shaded regions indicate which strategy (L, R or U) has highest survival given these costs for (a, b) ψ=0.001 and (c, d) ψ=0.004, and for (a, c) γB=0.35, γG=0.65 and (b, d) γB=0.15, γG=0.85, with k = 2 in all panels. The results do not qualitatively depend on the values of ψ or k, with an intuitive gradual shrinking of the L region as k (relative cost of Bayesian learning compared to a reserve-based strategy) increases.
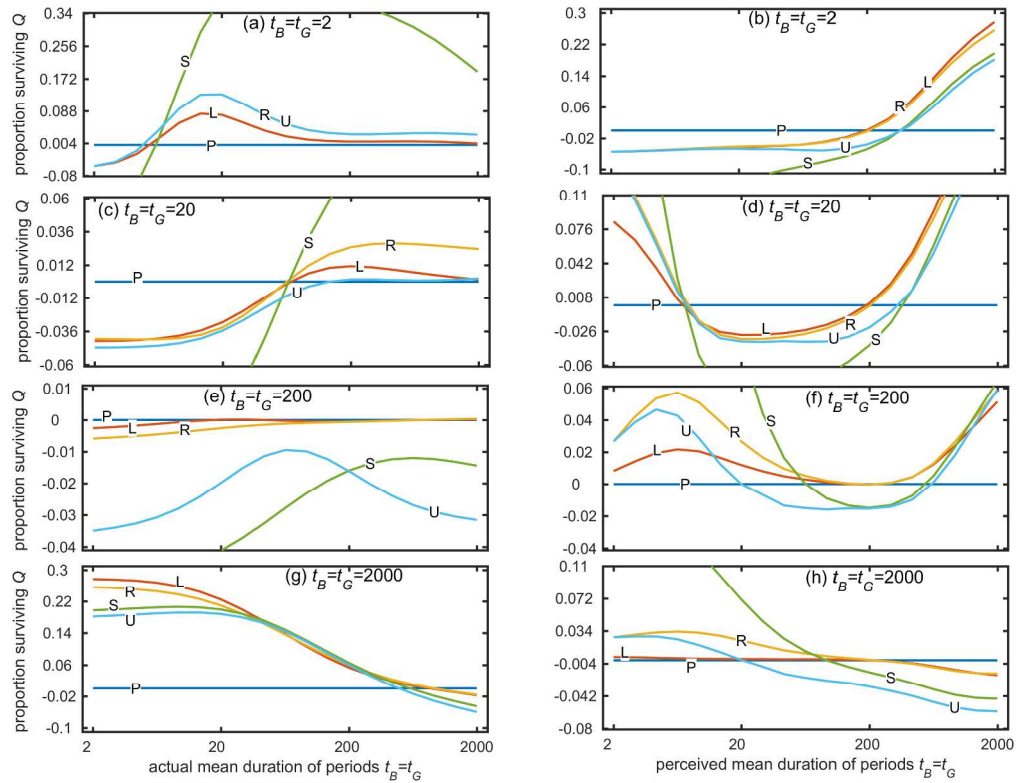
183x167mm (300 x 300 DPI)

Figure 5: Probability of surviving 2000 time steps Q(i) when the actual fluctuation rates differ from those the animal is adapted to. Survival is plotted as the difference compared to survival under perfect knowledge (P); negative values imply lower survival as a result of imperfect knowledge about current conditions (different lines for strategies L, R, S and U). Left column (a, c, e, g) shows relative survival as a function of the actual mean duration of periods (x-axis) when following the optimal strategy for the duration shown on the panels. Right column (b, d, f, h) shows survival as a function of the mean duration of periods to which the animal is adapted (x-axis) in four environments with different actual mean durations (shown on the panels). (a, b) tB=tG=2000; (c, d) tB=tG=200; (e, f) tB=tG=20; (g, h) tB=tG=2.

236x182mm (300 x 300 DPI)