An adaptive response to uncertainty can lead to weight gain during dieting attempts

Journal:	*Evolution, Medicine, and Public Health*
Manuscript ID	EMPH-2016-024.R1
Manuscript Type:	Original Research Article
Date Submitted by the Author:	13-Oct-2016
Complete List of Authors:	Higginson, Andrew; University of Bristol, Biology McNamara, John; University of Bristol, Biology
Keywords:	obesity, optimal foraging, contrast effect, low calorie diets, yo-yo dieting, weight cycling

SCHOLARONE™
Manuscripts

1

2   An adaptive response to uncertainty can lead to

3   weight gain during dieting attempts

4   A. D. Higginson[1,2], J. M. McNamara[3]

5

6   [1] Centre for Research in Animal Behaviour, College of Life and Environmental Sciences, University of

7   Exeter, Exeter EX4 4QG, UK, adhigginson@gmail.com Tel: +44 (0) 117 394 1296, Fax: +44 (0)117 331

8   7985

9   [2] Previous address: School of Biological Sciences, Life Sciences Building, University of Bristol, 24 Tyndall

10 Avenue, Bristol, BS8 1TQ, UK.

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

12 john.mcnamara@bristol.ac.uk

13

14 **Header**: Uncertainty and weight gain

15 **Keywords**: obesity, optimal foraging, contrast effect, low calorie diets, hunger, yo-yo dieting, weight cycling

16 **Abstract**: 209 words

17 **Lay summary**

18 Repeated dieting may lead to weight gain because the brain learns that the food supply is unreliable. Animals

19 respond to food shortage by storing fat. Our model of learning shows that if the food supply is restricted (as

20 in dieting) an optimal animal should gain excess weight between diets.

21

22   **ABSTRACT**

23

24   **Background and objectives**: Peoples' attempts to lose weight by low calorie diets often result in weight

25   gain because of over-compensatory overeating during lapses. Animals usually respond to a change in food

26   availability by adjusting their foraging effort and altering how much energy reserves they store. But in many

27   situations the long term availability of food is uncertain, so animals may attempt to estimate it to decide the

28   appropriate level of fat storage.

29   **Methodology**: We report the results of a conceptual model of feeding in which the animal knows whether

30   food is currently abundant or limited, but does not know the proportion of time there will be an abundance in

31   the long term and has to learn it.

32   **Results**: If the food supply is limited much of the time, such as during cycles of dieting attempts, the optimal

33   response is to gain a lot of weight when food is abundant.

34   **Conclusions and implications**: This implies that recurring attempts to diet, by signalling to the body that the

35   food supply is often insufficient, will lead to a greater fat storage than if food was always abundant. Our

36   results shed light on the widespread phenomenon of weight gain during weight cycling, and indicate possible

37   interventions that may reduce the incidence of obesity.

38   **INTRODUCTION**

39

40   Overweight and obese people are frequently able to lose weight but are unable to maintain such losses long-

41   term [1], which is why a large proportion of individuals are on diets at any given time [2]. Repeated weight

42   loss and gain is referred to as yo-yo dieting or weight cycling [2]. Whilst most people can lose weight during

43   diets, weight gain between diets is proportional to the weight lost [3] and may even lead to weight gain in the

44   long-term [4–7]. Whilst weight cycling *per se* is not associated with health issues [8,9], the weight gain has

45   many health implications [10]. There are many mechanisms underpinning eating behaviour that may

46   contribute to weight gain [11]. Some research has focussed on the physiological mechanisms that cause long

47   term weight gain in response to repeated dieting attempts, such as changes in the production of regulatory

48   hormones [5,7] which may shift the body's response to signals from adipose tissue [12].

49

50   Whilst it is essential to understand the mechanisms, the search for treatments for obesity will involve

51   achieving a holistic understanding of regulatory systems. A descriptive model that mimics the cycling

52   phenomena [13] assumes that weight gain stops at some maximum and weight loss stops at some minimum.

53   But this model does not elucidate why, in evolutionary terms, a system would be designed as it is supposed.

54   An evolutionary perspective can help to elucidate the causes of being overweight and obese [14].

55   Evolutionary arguments centre around the usefulness of fat as a source of energy under food shortage and the

56   costs of carrying stored fat [15]. Models of adaptive behaviour that consider fat as a means to reduce the risk

57   of starvation have been highly successful at predicting energy storage in animals [16–23]. These models

58   typically do not try to capture the complexities of physiological and psychological mechanism that control

59   eating [11,24,25], but provide functional explanations for the values of states that arise from such

60   mechanisms, such as the quantity of energy that is stored [26]. Evolutionary approaches to understanding

61   obesity [27,28] typically assume that humans will have physiological and cognitive systems that evolved in

62   natural (ancestral) environments and have not changed since then, and we know that maladaptive behaviours

63   of various kinds can emerge from strategies that are adaptive in natural environments [29]. Evidence

64   suggests that energy use in Western environments is similar to that for hunter-gatherers [30], suggesting that

65   excessive food consumption rather than sedentary lifestyles causes obesity.

66

67   Humans appear to have sophisticated controls on fat storage that act to maintain weight at some target, but

68   the variation in body weight within populations indicates that this target must differ between individuals

69   [31]. It has not been fully elucidated why individuals might differ in this way. Existing data show that whilst

70   a significant proportion of the variation in BMI is attributable to genetic factors [32] there are strong effects

71   of socioeconomic factors [33]. This indicates that learning may play an important role in determining the

72   individuals' targets. Here, we assess how weight gain after dieting attempts could be an adaptive response

73   involving learning about the environment. Our model provides proof of the concept that weight gain may be

74   a response to an environment to which the evolved subconscious system for controlling energy storage is no

75   longer adapted.

76

77   **THE MODEL**

78

79   We assume that humans have evolved in environments where the food supply fluctuates between limited and

80   abundant, but also that there are times, years or seasons, where the proportion of time that food is abundant is

81   greater or lesser [30,34,35]. The current level of food availability is therefore not sufficient to infer the long-

82   term food availability. It is a ubiquitous feature of natural environments that food availability varies over

83   time and shows such positive autocorrelation and our formulation captures this in the simplest possible way.

84   We model a hypothetical animal that uses energetic reserves to meet all its needs and tries to learn about the

85   long-term food availability from observing the short-term fluctuations. This animal is adapted to conditions

86   over evolutionary history in which the food supply fluctuated. We are interested in the consequences if

87   dieting attempts are interpreted by the subconscious brain as such fluctuations.

88

89   **The animal and its environment**

90   We model time as a sequence of discrete epochs in which the animal makes a decision and its state variables

91   may change from one epoch to the next. The animal is characterized by four state variables [36]. The first is

92   its level of energetic reserves x. There are two external states: the current food condition C where food

93   availability is higher in the rich condition ($C=R$) than the poor condition ($C=P$), and the current state of the

94   world W which can be good ($W=G$) or bad ($W=B$), which differ in the average durations of rich and poor

95   periods. The animal knows the current conditions without error, but does not directly know whether the

96   world is good or bad. The final state variable is the animal's current estimated probability that the world is

97   good (ρ). Note that we do not assume that any animal has a perfect system for calculating probabilities, but

98   that evolution has selected for a cognitive system that behaves as though it tracks a probability. At the end of

99   a decision epoch, the world changes from its current state W to the alternative state with probability θ_W.

100   When the world is in state W conditions change from the current condition C to the alternative condition with

101   probability $\lambda_{W,C}$. We fix these probabilities so that conditions are predominantly rich in the good world and

102   predominantly poor in the bad world, and that conditions change much more frequently than the state of the

103   world ($\theta_W \ll \lambda_{W,C}$). Examples of food availability over time in good and bad worlds are shown in Figure A1.

104   Each decision epoch the probability that the world is currently good (ρ) is updated using Bayes' rule. Figure

105   A2 illustrates how probabilities are updated for the baseline parameter values.

106

107   The aspect of behaviour we are interested in is the proportion of time the animal spends foraging per

108   decision epoch, which we call f. Increasing f increases the probability of finding food. Poor and rich

109   conditions differ only in the maximum probability of finding food per decision epoch when foraging (γ_R and

110   γ_P, where $\gamma_R > \gamma_P$); the animal finds food during unit time with probability $\gamma_C f$. For computational reasons

111   there is some variance in the energy content of food items (see Online Appendix) and they contain on

112   average b units of energy.

113

114   In natural environments are a variety of costs of carrying fat reserves. In modelling fat regulation in small

115   birds it is usual to assume that energy expenditure increases with the amount (and hence weight) of fat

116   carried. It is also often assumed that predation risk when foraging increases with increasing fat load because

117   of decreasing manoeverability [37]. Regardless of the exact cost, some cost needs to be assumed if long-term

118   adaptive fat levels are to be stable [38]. In humans it seems reasonable to assume that the rate of energy

119   expenditure during activity increases with increasing fat load. This would then impose a cost since increased

120   expenditure requires increased time finding food, resulting in less time that is available to spend on other

121   activities. Our model is based on such a cost. We assume that the animal's rate of energy expenditure $m(x)$

122   increases with energy reserves x – representing the energetic costs of carrying fat in humans [39] and animals

123   [40] – according to:

124
$$m(x) = m_0 \left[1 + m_x \frac{x}{x_{max}} \right]$$
(1)

125 where m_x (>0) controls how the cost increases with reserves, and m_0 controls the magnitude of costs. For the

126 baseline parameter values (Table 1), this means that an animal with maximum fat stores would use energy at

127 twice the rate of an animal with no fat. A consequence is that the benefit of building up energetic reserves

128 will diminish so we never predict that stores should be near the maximum. We set other parameter values so

129 that the expected net rate of energy gain at f=1 in bad conditions is slightly positive; thus there is a risk of

130 starving to death, but animals are expected to survive sufficiently long that the model makes clear predictions

131 about the effects of other parameters.

132

133 We assume that there are two source of mortality [41]. If the energy reserves of the animal reach $x = 0$ the

134 animal dies of starvation. During each epoch there is also a probability μ of death from external sources that

135 is independent of state and behaviour. We assume that the time that the animal does not spend foraging is

136 invested in increasing its reproductive success, such as in courting potential mates. This reproductive payoff

137 is instantaneous and subject to diminishing returns so that foraging for a proportion f of a single decision

138 epoch increases the animal's lifetime reproductive success by $\sqrt{1-f}$. There is therefore a trade-off

139 between immediate investment in reproduction and increasing the future investment by finding food to

140 increase the expected lifespan. A strategy specifies how the value of f depends on the three state variables x,

141 ρ and C (W is not directly known). The optimal strategy f^* maximises the total lifetime reproductive success

142 of the animal. Under this strategy the proportion of time spent foraging when the combination of state

143 variables is (x, ρ, C) is $f^*(x, \rho, C)$. We use standard methods of stochastic dynamic programming [36] to find

144 this strategy. See Online Appendix for full details.

145

146 **Cost of being active**

147 Thus far we have assumed that the rate of energy use is the same whether the individual is foraging or not,

148 but fitness-promoting activities may be sedentary (e.g. grooming) or active (e.g. singing). To allow for the

149 dependence of energy use on activity we set the rate of energy expenditure to be:

150
$$m(x,f) = m_0 \left(m_f f + \{1 - m_f\} \right) \left[1 + \left(m_{x,f} f + \{1 - m_{x,f}\} \right) m_x \frac{x}{x_{max}} \right] \quad (2)$$

151 where m_f controls the dependence of energy expenditure on activity and $m_{x,f}$ controls the dependence of the

152 costs of energy reserves on energy use when active (i.e. the interaction). Note that if $m_f = m_{x,f} = 0$ we recover

153 equation (1). If all else were equal, the extra costs of activity would decrease average energy expenditure

154 (because $f \leq 1$), and so average costs and type of costs would be confounded in any comparison. To minimise

155 the effect of average costs we adjusted the value of m_0. The approximate mean value of reserves under

156 normal conditions for baseline parameter values (Table 1) is 25, so the average energy use will be around

157 $\frac{1}{2}\left[1 + \frac{25}{x_{max}} \right] = \frac{5}{8}$. We took average f to be 0.5, and so use a value of m_0 given by:

158
$$m_0 = \frac{5/8}{\left(\frac{m_f}{2} + \{1 - m_f\} \right)\left[1 + \left(\frac{m_{x,f}}{2} + \{1 - m_{x,f}\} \right) m_x \frac{25}{x_{max}} \right]} . \quad (3)$$

159

Assessment of behaviour

161 The dynamic programming procedure calculates the reproductive value of the animal in all states $V(x, \rho, C)$,

162 which is the expected contributions to reproductive success before death. We use V to assess the strength of

163 the urge to add to fat stores by calculating the risk that would be tolerated to gain the equivalent of two items

164 of food. Specifically, we calculate the extra mortality risk μ' at which the animal is indifferent between its

165 current situation and gaining 10 extra units of reserves at risk μ'. This mortality risk satisfies

166
$$(1 - \mu')V(x + 10, \rho, C) = V(x, \rho, C) . \quad (4)$$

167 Rearranging gives

168
$$\mu' = 1 - \frac{V(x, \rho, C)}{V(x + 10, \rho, C)} . \quad (5)$$

169

170 We calculate the average amount of energy stored when following the optimal strategy in four conditions.

171 Firstly, under normal conditions in the good world with conditions changing between poor and rich

172 according to the values of $\lambda_{G,P}$ and $\lambda_{G,R}$. Secondly, for constant rich conditions, which we refer to as 'glut'.

173 Thirdly, when conditions switch slowly between poor and rich, referred to as 'slow diet'. Fourthly, when

174   conditions switch rapidly between poor and rich, referred to a 'quick diet'. Thus, we simulate different

175   dieting patterns. We are interested in the predicted energy storage and the belief that the world is good (ρ)

176   under these four conditions.

177

178   **RESULTS**

179

180   Figure 1 shows the optimal strategy for the baseline parameter values (Table 1). Generally, the optimal

181   foraging rate f^* is higher in the bad world, because the animal must attempt to have greater insurance against

182   the risk of going without food and starving. In both worlds, f^* is greater at low reserves in poor conditions

183   than in rich conditions because it is crucial to find food before starvation, whilst at high reserves f^* is greater

184   in rich conditions than in poor conditions (even in the good world) because it is worth trying to build up the

185   insurance when food is abundant (for more exploration of conditions see [42]). The target level of reserves

186   in rich conditions is higher in the bad world than the good world because more insurance is needed as the

187   period of food shortage is likely to be longer.

188

189   Constant glut conditions lead to greater energy reserves than under normal conditions, but the response to

190   periods of poor conditions leads to overcompensation when conditions become rich (Figure 2a). This results

191   in greater energy reserves after dieting attempts than in constant glut conditions. This occurs because the

192   animal becomes convinced that the world is bad (Figure 2b), and that it must take advantage of rich

193   conditions whilst they last. If conditions fluctuate quickly, reserves are lower in the short-term (Figure 2a)

194   but the animal becomes more convinced the world is bad over the longer term (Figure 2b). The extra

195   mortality risk that would be tolerated to get two food items is plotted for as a function of current reserves

196   when following the slow diet (Figure 2c). Because the animal is convinced the world is bad, it is willing to

197   risk up to 2x greater than in a constant glut when reserves become low. However, this increase depends on

198   the combination of reserves being low and the belief that the world is bad: lower values of μ' than for glut

199   conditions are predicted at low reserves and believing the world is *good* (grey dashed line) and believing the

200   world is bad at *high* reserves (black dashed line).

201

202   The reasons for the weight gain after the period of poor conditions can be understood by considering the

203   optimal strategy (Figure 1). When the individual has a period of poor conditions then switches to rich, it

204   'believes' that the world is bad, so the gain in reserves is greater than it would have been if conditions were

205   always rich. Thus, the gain in weight after repeated dieting comes about because an animal with high

206   reserves should forage more in rich conditions especially when it believes that there is a strong possibility

207   that conditions will turn poor. The reserves stored under dieting approach an asymptote over a longer period

208   of time whereas under constant glut they drop down to that stored under normal conditions (Figure A3),

209   because the individual becomes convinced the world is good so there is no need to store much energy.

210

211   The mean reserves stored in the good world increases with the duration of bad periods and with the duration

212   of periods in the bad world (Figure 3a), due to the insurance effect. To illustrate this effect, we present the

213   optimal strategy in Figure A4 for nine of the 21 parameter value combinations used to make Figure 3.

214   Reserves in a glut are greater relative to under normal conditions when poor periods in the good world are

215   longer (Figure 3b). Hence, the greatest gain after dieting attempts, relative to glut conditions, is when poor

216   conditions are short in the good world (Figure 3c), because this causes a greater difference in the target level

217   of reserves in rich conditions between the bad and good world (cf. Figure A4b, h). After 256 decision epochs

218   lower reserves are stored if dieting fluctuations are quicker (cf. Figure 3c, 1) for most situations, and exceed

219   reserve level expected in a glut only if poor periods are very short in the good world.

220

221   One explanation for difficulty in losing weight is that lighter bodies require less energy so food consumption

222   needs to progressively reduce [7]. In Figure 4 we show the effect of changing the magnitude of the

223   dependence of energy expenditure on the level of reserves. In all conditions more fat is stored in constant

224   glut compared than under normal conditions. Although larger values of m_x than unity tend to either decrease

225   (g/n) or increase ($s/g, q/g$) relative reserve levels, the overall pattern is unchanged. However, when m_x is zero

226   – meaning that energy use does not increase with energy storage – we do not predict dieting to cause weight

227   gain ($s/g<1$ and $q/g<1$), suggesting that the energetic cost of fat storage is essential to the increase in body

228   weight due to weight cycling.

229

230   The optimal strategy is influenced by changing the various costs (m_x, m_f, $m_{x,f}$, see Figure A5). The effect of

231   dieting is considerably weaker when $m_{x,f}>0$ because the extra cost of activity means that the animal gains

232   more reserves in rich conditions. However, the effect is stronger in the good world, and so results in less of a

233   difference between glut and dieting conditons. m_f has very small effects on predicted energy storage (cf. left

234   and right panels of Figure 4) because the individual can save costs in poor conditions by being inactive,

235   which reduces the advantage to storing fat, and this cancels out the selective pressure to store more fat in

236   response to increased costs. There is discrepancy between the level of reserves that individuals should try to

237   store ('target') and the reserves that can be built up ('realised'), which differ due to stochasticity (Figure A6)

238   and the difference depends on the types of costs (Figure A6e, f). Note that in all cases the discrepancy in

239   dieting conditions is much smaller than in glut conditions. Based on the discrepancy between the target and

240   the realised state, the urge to eat strongest when the rate of energy use is constant (0, 0) or increases with

241   reserves and this is at a greater rate when foraging (1,1). The urge to eat will be weakest when the rate of

242   energy use only depends on reserves, but strongly (2,0). Again, the effect of an overall cost of foraging (m_f)

243   is small and constant across other costs (cf. Figure A6e, f).

244

245   **DISCUSSION**

246

247   Weight cycling is common in people that are attempting to lose weight, but many people gain weight in the

248   long term. The functional reasons that our energy storage systems might respond in this way to dieting

249   attempts has not been elucidated. We have used a simple generic model of feeding to demonstrate how a

250   reserve-control system following an ecological rational strategy [43] could gain weight over the long-term if

251   periods of food shortage are frequent, even if they are associated with short-term weight loss. Our work

252   therefore proposes a potential cause of the association between weight cycling and weight gain [5,44,45]:

253   that dieting attempts *cause* weight gain via providing (misleading) information about the environment to the

254   subconscious systems that control body mass. That is, even in the "constant glut" [46] conditions in the

255   developed world where food is always abundant, the subconscious decision-making systems that underpin

256   our behaviour may interpret dieting attempts as indicative of an environment with common food shortages,

257   and this triggers the (previously) appropriate behavioural responses.

258

259   Our model predicts that energy reserves should respond to repeated attempts to diet by weight cycling and

260   becoming greater from one cycle to the next. The more reliable food was when the world was good, the

261   greater the relative fat storage during repeated dieting attempts is predicted to be, because these dieting

262   attempts cue that the world is more likely to be bad. Thus, the very conditions that cause weight gain initially

263   – a glut of food – causes further weight gain once cyclical dieting begins. There is evidence that among

264   weight cycling people those that switch between dieting and binge-eating more frequently gain more weight

265   [45]. By contrast we found that quick oscillations tended to lead to less weight gain for the period we

266   studied, but over the longer-term the duration of dieting periods has little effect on the average energy

267   storage (Figure A3). We concentrate on outcomes after a relatively short period of dieting (256 time steps,

268   Figure 2 and 3), partly because people do not diet forever but also because energy storage tends to level off

269   (Figure A3). We note that our model predicts that fat storage under constant glut conditions that persist for a

270   long time will actually not be substantially greater than under normal conditions (Figure A3), suggesting that

271   the abundance of energy-rich food is not a complete explanation for the obesity epidemic. Fat storage over a

272   long period of dieting attempts will be greater than under constant glut conditions, implying a critical role of

273   informational constraints and learning.

274

275   Our results suggest that the magnitude of the weight gain between diets will depend on the cost of the non-

276   foraging activity. Our rescaling (by adjusting m_0) means that we compare predictions depending on the

277   *relative* cost of the other activity. When m_f and $m_{x,f}$ are small foraging is much more costly than other

278   activities; when unity other activities are equally costly. For some species and situations the activities that

279   enhance reproduction may be energetically inexpensive, such as grooming in primates. In other cases,

280   activities essential to reproduction may be equally as energetic as foraging, such as maintaining a territory. It

281   is difficult to know what best applies to humans. However, it may be possible to quantify the relative costs of

282   foraging across species, which would offer possibilities for testing our predictions. Foraging may be

283   relatively more costly than non-foraging in a small bird (i.e. small m_f) compared to a rodent (i.e. large m_f). If

284   we could expose laboratory birds (e.g. zebra finch) and rodents (e.g. mice) to a yo-yo diet regime, we would

285   predict that the rodent would gain more weight. A very large scale project could try to estimate our cost

286   parameters (m_x, m_f, $m_{x,f}$) for several populations or closely related species in order to assess the responses to

287   'dieting', and then measure the target and/or realised level of reserves (Figure A6).

288

289   Lowe [3] argues that yo-yo dieting does not cause weight gain but is merely a correlate of the potential for

290   weight gain, which may arise if people who know they often overeat take steps to avoid weight gain. This

291   argument is based on the assumption that causation is one-way, but our perspective shows that causation may

292   be two-way between food restriction and overeating, leading to a spiral of dieting and weight gain. We

293   suggest that the interpretation of data is hampered by a lack of robust theory, and hope that our work may

294   cause a re-evaluation of observations of weight cycling. For instance, Lowe [3] suggests that weight regain is

295   *caused* not by dieting but by increased binge eating and increased reward value of food. From our

296   perspective, these are proximate mechanisms that implement the behavioural strategy that we have

297   identified; thus both explanations can be true.

298

299   Not all individuals acquire excess weight after dieting [47]. Our results suggest that variation among

300   individuals could occur if people have different subconscious expectations of the pattern of food availability

301   (Figure 2c, d). For instance, weight cycling does not promote extra weight gain if the system 'expects'

302   conditions to change very slowly or very rapidly when the world is bad. If such 'expectations' were

303   determined by natural selection in different environments and encoded in genes, then this effect may underlie

304   effects of ethnicity on the risk of obesity [33,48]. On the other hand, this 'expectation' may be learnt during a

305   lifetime, which may underlie the effects of age on the apparent heritability of obesity [32]. Furthermore, this

306   provides a possibility for testing our predictions: if young mice occasionally experience periods where food

307   is restricted but is available with various rates of fluctuations (e.g. every other hour; every other day) then

308   when older they should show different responses to intermediate frequencies of food restriction. Specifically,

309   those that were exposed to intermediate rate of fluctuations may gain the most weight (peaks in Figure 4),

310   and those used to constant glut conditions would gain more weight relative to control individuals than those

311   subject to occasional food shortage when young (cf. different lines in Figure 4).

312

313   Experiments that use various protocols of food restrictions could be used to assess the predictions around

314   foraging intensity (Figure 1, A5), provided there was an appropriate continuous measure of the behaviour of

315   subjects. Since our predictions are state-dependent, repeated measures of the same individuals after their fat

316   stores have been manipulated through food restriction or gluts would provide a powerful test. For instance,

317   the crossover points in the strategies mean that at low reserves we predict higher intensity foraging when

318   food is scarce (e.g. low fixed ratio schedule) than abundant, and the converse when at high reserves. A more

319   challenging experiment could try to manipulate the subjects' beliefs about not only current conditions, but the

320   'world': the long term conditions. Under some parameter combinations (e.g. Figure A4b), we predict there

321   will be a crossover in foraging intensity when food is currently abundant: a low reserves subjects should

322   show lower intensity foraging if the world is bad but the converse when at high reserves.

323

324   The additional risk of mortality that would be incurred to obtain food can be seen as a surrogate for the

325   strength of motivation to eat. Our results on this risk explain why people's motivation systems strongly push

326   them to eat high calorie food, and why this urge will be especially strong during a diet [49]. Interestingly, we

327   predict that this urge will not gradually diminish over dieting attempts (although calories consumed will be

328   lower) despite weight being gained, because the system becomes more and more convinced the world is bad.

329   People who attempt to diet for a very long time will not continue to gain weight but reach an asymptote

330   (Figure A3), seemingly much higher than those who never diet (constant glut). Real people are much more

331   complex than our model, but it seems likely that for some people who have been dieting for a long time may

332   benefit from trying to maintain their body weight for some time rather than reduce calorie intake, to

333   'convince' their regulatory systems that the food supply is reliable.

334

335   Our cognitive systems will have evolved to reflect the fact that current conditions are informative of future

336   conditions (i.e. the world is temporally autocorrelated)[50]. This is a contrast effect [51], a seemingly

337   irrational behavioural phenomenon seen in many animals [52–54], including humans [55] that can arise due

338   to uncertainty about the long-term state of the world [56], which could underlie several other psychological

339   phenomena [29]. Current conditions in the developed world are constant glut [10], but any uncertainty could

340   make people gain further weight, because learning about food availability from dieting attempts alters

341   expectations about food availability in the future. That optimal behaviour depends on future expectations is

342   well established [41], but weight gain between diets is another possible example of behaviour being affected

343   by past experience in seemingly irrational ways [57].

344

345 We cannot capture all the complexities of weight cycling in a simple model, so we assume that there are two

346 levels of food availability and study a single cycle, finding when the level of fat should be greater at the end

347 of the cycle than it would otherwise have been. In reality, people are learning over a long term. However, we

348 find that the weight gain slows as more fat is stored (Figure A3), which is consistent with the observation

349 that obese people do not gain further weight as a result of dieting [7], so we expect that a more long-term

350 model would not lead to further insights. Our model only captures the function of fat storage, and we have

351 not attempted to specify the psychological or physiological mechanisms that bring it about; one possible

352 mechanism is an alteration of the sensitivity of anabolic responses to adiposity signals [12].

353

354 Further developments of our model could include decision-making about how much lean mass should be

355 stored and when protein might be catabolised for energy, as we have shown this flexibility may affect

356 decisions about fat storage [58,59]. However, even our simple model demonstrates the principle that

357 understanding weight gain during yo-yo dieting does not require recourse to explanations based around the

358 feeding control system malfunctioning [1,11] or being overwhelmed by modern food stimuli [10,11]. The

359 feeding system could be functioning perfectly, but uncertainty about the food supply triggers the adaptive

360 response to gain weight.

361

362 **ACKNOWLEDGMENTS**

363

367

368 **REFERENCES**

369 1. Stunkard A, Penick S. Behavior-modification in the treatment of obesity - problem of maintaining weight-

370 loss. *Arch Gen Psychiatry* 1979;**36**:801–6.

371 2. Blackburn G, Wilson G, Kanders B *et al.* Weight cycling - the experience of human dieters. *Am J Clin

372 Nutr* 1989;**49**:1105–9.

373 3. Lowe MR. Dieting: proxy or cause of future weight gain? *Obes Rev* 2015;**16**:19–24.

374   4. Delahanty LM, Meigs JB, Hayden D *et al.* Psychological and Behavioral Correlates of Baseline BMI in

375   the Diabetes Prevention Program (DPP). *Diabetes Care* 2002;**25**:1992–8.

376   5. Dankel SN, Degerud EM, Borkowski K *et al.* Weight cycling promotes fat gain and altered clock gene

377   expression in adipose tissue in C57BL/6J mice. *Am J Physiol Endocrinol Metab* 2014;**306**:E210-24.

378   6. Summermatter S, Mainieri D, Russell AP *et al.* Thrifty metabolism that favors fat storage after caloric

379   restriction: a role for skeletal muscle phosphatidylinositol-3-kinase activity and AMP-activated protein

380   kinase. *FASEB J* 2008;**22**:774–85.

381   7. Dulloo AG, Jacquet J, Montani J-P. How dieting makes some fatter: from a perspective of human body

382   composition autoregulation. *Proc Nutr Soc* 2012;**71**:379–89.

383   8. Simkin-Silverman L, Wing R, Plantinga P *et al.* Lifetime weight cycling and psychological health in

384   normal-weight and overweight women. *Int J Eat Disord* 1998;**24**:175–83.

385   9. Bartlett SJ, Wadden TA, Vogt RA. Psychosocial consequences of weight cycling. *J Consult Clin Psychol*

386   1996;**64**:587–92.

387   10. O'Dea K. Obesity and diabetes in the land of milk and honey. *Diabetes Metab Rev* 1992;**8**:373–88.

388   11. Lowe MR. Eating patterns, diet quality and energy balance: from the macro- to the microscopic. *Physiol*

389   *Behav* 2014;**134**:123–5.

390   12. Ferrannini E, Rosenbaum M, Leibel RL. The threshold shift paradigm of obesity: evidence from

391   surgically induced weight loss. *Am J Clin Nutr* 2014;**100**:996–1002.

392   13. Goldbeter A. A model for the dynamics of human weight cycling. *J Biosci* 2006;**31**:129–36.

393   14. Wells JCK. The evolution of human fatness and susceptibility to obesity: an ethological approach. *Biol*

394   *Rev* 2006;**81**:183–205.

395   15. Wells JCK. Thrift: a guide to thrifty genes, thrifty phenotypes and thrifty norms. *Int J Obes*

396   2009;**33**:1331–8.

397   16. McNamara JM, Houston AI. The value of fat reserves and the tradeoff between starvation and predation.

398   *Acta Biotheor* 1990;**38**:37–61.

399   17. Krams I, Cirule D, Suraka V *et al.* Fattening strategies of wintering great tits support the optimal body

400   mass hypothesis under conditions of extremely low ambient temperature. *Funct Ecol* 2010;**24**:172–7.

401   18. Gosler A. Environmental and social determinants of winter fat storage in the great tit Parus major. *J Anim*

402   *Ecol* 1996;**65**:1–17.

403   19. McNamara JM, Barta Z, Houston AI *et al.* A theoretical investigation of the effect of predators on

404   foraging behaviour and energy reserves. *Proc R Soc B-Biological Sci* 2005;**272**:929–34.

405   20. Cresswell W, Clark JA, Macleod R. How climate change might influence the starvation-predation risk

406   trade-off response. *Proc Biol Sci* 2009;**276**:3553–60.

407   21. Houston, McNamara JM. A theoretical investigation of the fat reserves and mortality levels of small

408   birds in winter. *Ornis Scand* 1993;**24**:205–19.

409   22. Pravosudov V V, Grubb TC, Doherty PF *et al.* Social dominance and energy reserves in wintering

410   woodland birds. *Condor* 1999;**101**:880–4.

411   23. Higginson ADD, McNamara JMJM, Houston AIAI. The starvation-predation trade-off predicts trends in

412   body size, muscularity and adiposity between and within taxa. *Am Nat* 2012;**179**:338–50.

413   24. Illius AW, Tolkamp BJ, Yearsley J. The evolution of the control of food intake. *Proc Nutr Soc*

414   2002;**61**:465–72.

415   25. Rozin P. Sweetness, sensuality, sin, safety, and socialization: some speculations. In: Dobbing J (ed.).

416   *Sweetness*. Berlin: Springer-Verlag, 1987.

417   26. Houston AI, McNamara JM. What current models can and cannot tell us about adaptive energy storage.

418   In: Stephens DW, Brown JS, Ydenberg RC (eds.). *Foraging: Behavior and Ecology*. Chicago: University of

419   Chicago Press, 2007, Box 7.3.

420   27. Higginson AD, McNamara JM, Houston AI. Fatness and fitness: Exposing the logic of evolutionary

421   explanations for obesity. *Proc R Soc L B* 2016;**2443**.

422   28. McNamara JM, Higginson AD, Houston AI. Costs of foraging predispose animals to obesity-related

423   mortality when food is constantly abundant. *PLoS One* 2015;**141811**.

424   29. Fawcett TW, Fallenstein B, Higginson AD *et al.* The evolution of decision rules in complex

425   environments. *Trends Cogn Sci* 2014;**18**:153–61.

426   30. Pontzer H, Raichlen DA, Wood BM *et al.* Hunter-Gatherer Energetics and Human Obesity. Chehab FF

427   (ed.). *PLoS One* 2012;**7**:e40503.

428   31. Speakman JR. Obesity: The integrated roles of environment and genetics. *J Nutr* 2004;**134**:2090S–

429   2105S.

430   32. Elks CE, den Hoed M, Zhao JH *et al.* Variability in the heritability of body mass index: a systematic

431   review and meta-regression. *Front Endocrinol (Lausanne)* 2012;**3**:29.

432   33. Min J, Chiu DT, Wang Y. Variation in the heritability of body mass index based on diverse twin studies:

433   a systematic review. *Obes Rev* 2013;**14**:871–82.

434   34. Marlowe FW, Berbesque JC. Tubers as fallback foods and their impact on Hadza hunter-gatherers. *Am J

435   Phys Anthropol* 2009;**140**:751–8.

436   35. Sherry DS, Marlowe FW. Anthropometric data indicate nutritional homogeneity in Hadza foragers of

437   Tanzania. *Am J Hum Biol* 2007;**19**:107–18.

438   36. Houston AI, McNamara JM. *Models of Adaptive Behaviour*. Cambridge, U.K.: Cambridge University

439   Press, 1999.

440   37. Witter MS, Cuthill IC. The ecological costs of avian fat storage. *Philos Trans R Soc London, Ser B*

441   1993;**340**:73–92.

442   38. Houston AI, Welton NJ, McNamara JM. Acquisition and maintenance costs in the long-term regulation

443   of avian fat reserves. *Oikos* 1997;**78**:331–40.

444   39. Hall KDD, Jordan PNN. Modeling weight-loss maintenance to help prevent body weight regain. *Am J

445   Clin Nutr* 2008;**88**:1495–503.

446   40. Nilsson J-A. Metabolic consequences of hard work. *Proc Biol Sci* 2002;**269**:1735–9.

447   41. McNamara JM, Houston AI. The common currency for behavioral decisions. *Am Nat* 1986;**127**:358.

448   42. Higginson AD, Fawcett TW, Trimmer PC *et al.* Generalized optimal risk allocation: foraging and

449   antipredator behavior in a fluctuating environment. *Am Nat* 2012;**180**:589–603.

450   43. Todd PM, Gigerenzer G. Environments That Make Us Smart: Ecological Rationality. *Curr Dir Psychol

451   Sci* 2007;**16**:167–71.

452   44. Cereda E, Malavazos AE, Caccialanza R *et al.* Weight cycling is associated with body weight excess and

453   abdominal fat accumulation: a cross-sectional study. *Clin Nutr* 2011;**30**:718–23.

454   45. Marchesini G, Cuzzolaro M, Mannucci E *et al.* Weight cycling in treatment-seeking obese persons: data

455   from the QUOVADIS study. *Int J Obes Relat Metab Disord* 2004;**28**:1456–62.

456   46. O'Dea K. Obesity and diabetes in "the land of milk and honey." *Diabetes / Metab Rev* 1992;**8**:373–88.

457   47. Bosy-Westphal A, Schautz B, Lagerpusch M *et al.* Effect of weight loss and regain on adipose tissue

458   distribution, composition of lean mass and resting energy expenditure in young overweight and obese adults.

459   *Int J Obes (Lond)* 2013;**37**:1371–7.

460   48. Ogden CL, Carroll MD, Curtin LR *et al.* Prevalence of overweight and obesity in the United States,

461   1999-2004. *Jama-Journal Am Med Assoc* 2006;**295**:1549–55.

462   49. Thomas EA, Bechtell JL, Vestal BE *et al*. Eating-related behaviors and appetite during energy imbalance

463   in obese-prone and obese-resistant individuals. *Appetite* 2013;**65**:96–102.

464   50. Aoki K, Feldman MW. Evolution of learning strategies in temporally and spatially variable

465   environments: A review of theory. *Theor Popul Biol* 2014;**91**:3–19.

466   51. Flaherty CF. *Incentive Relativity*. Cambridge, U.K.: Cambridge University Press, 1996.

467   52. Freidin E, Cuello MI, Kacelnik A. Successive negative contrast in a bird: starlings' behaviour after

468   unpredictable negative changes in food quality. *Anim Behav* 2009;**77**:857–65.

469   53. Couvillon PA, Bitterman ME. The overlearning-extinction effect and successive negative contrast in

470   honeybees (Apis mellifera). *J Comp Psychol* 1984;**98**:100–9.

471   54. Crespi LP. Quantitative Variation of Incentive and Performance in the White Rat. *Am J Psychol*

472   1942;**55**:467.

473   55. Kobre KR, Lipsitt LP. A negative contrast effect in newborns. *J Exp Child Psychol* 1972;**14**:81–91.

474   56. McNamara JM, Fawcett TW, Houston AI. An adaptive response to uncertainty generates positive and

475   negative contrast effects. *Science* 2013;**340**:1084–6.

476   57. Tversky A, Simonson I. Context-Dependent Preferences. *Manage Sci* 1993;**39**:1179–89.

477   58. Higginson AD, McNamara JM, Houston AI. The starvation-predation trade-off shapes the strategic use

478   of protein for energy during fasting. *J Theor Biol* 2014;**359**:208–19.

479   59. McNamara JM, Higginson AD, Verhulst S. The influence of the starvation-predation trade-off on the

480   relationship between ambient temperature and body size among endotherms. *J Biogeogr* 2016;**43**:809–19.

481

482

483 **Figure 1:** Optimal strategy of foraging intensity f^* for reserves x and poor ('P', grey) and rich ('R', black)

484 conditions for $\rho=0$ ('B',dashed) and $\rho=1$ ('G', solid) for the baseline parameter values shown in Table 1. f^*

485 changes smoothly for intermediate values of ρ (not shown). Dotted lines indicate the value of f necessary to

486 maintain a constant level of reserves long-term in Rich (black) and Poor (grey) conditions. Hence, where the

487 strategy lines of the same shade intersect the dotted lines is the target level of reserves. The target level of

488 reserves in rich conditions is higher in the bad world than the good world.
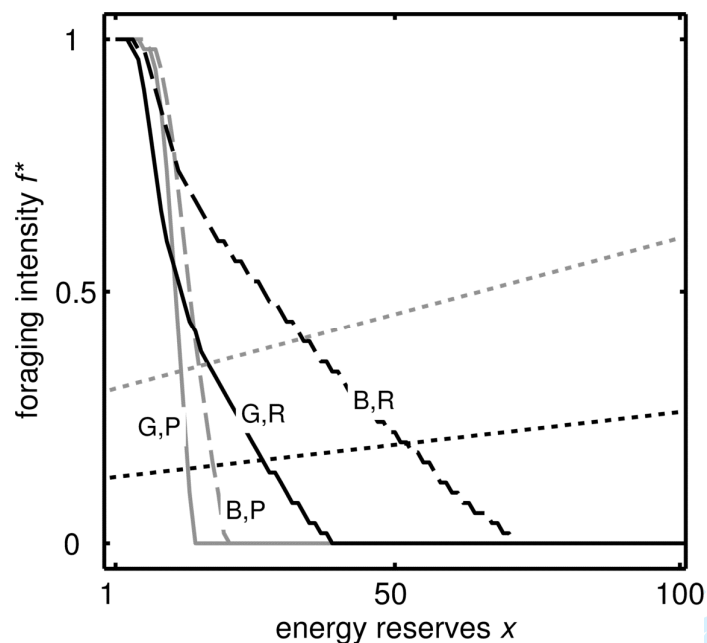
489

490

491   **Figure 2**: Effect of three 'treatments' compared to control conditions. (a) Mean energy reserves x over time

492   when conditions always Rich ('glut': g, dashed line) or when conditions switch between Poor and Rich every

493   32 epochs ('slow dieting attempts periods': s, solid grey line), or when conditions change between Poor and

494   Good every 8 epochs ('quick dieting attempts': q, solid black line), compared to the mean across Poor and

495   Rich conditions in the Good world ('control': n, dotted line). (b) Belief that the world is Good ρ for the same

496   period and treatments. Under normal conditions ρ settles down at a high level, whereas during a glut

497   conditions are always rich so learning is slower as $\lambda_{B,R} \approx \lambda_{G,R}$. (c) Selective pressure to eat food. We plot over

498   the course of the slow dieting periods the mortality risk that would be tolerated to get 10 units of energy μ',

499   as a multiple of what μ' would be tolerated under control conditions (solid grey line), and for comparison the

500   same metric for: reserves in the control conditions and belief under diet conditions (dashed black line);

501   reserves under diet conditions and belief under control conditions (dashed grey line); reserves and belief

502   under constant Rich conditions (dotted line).
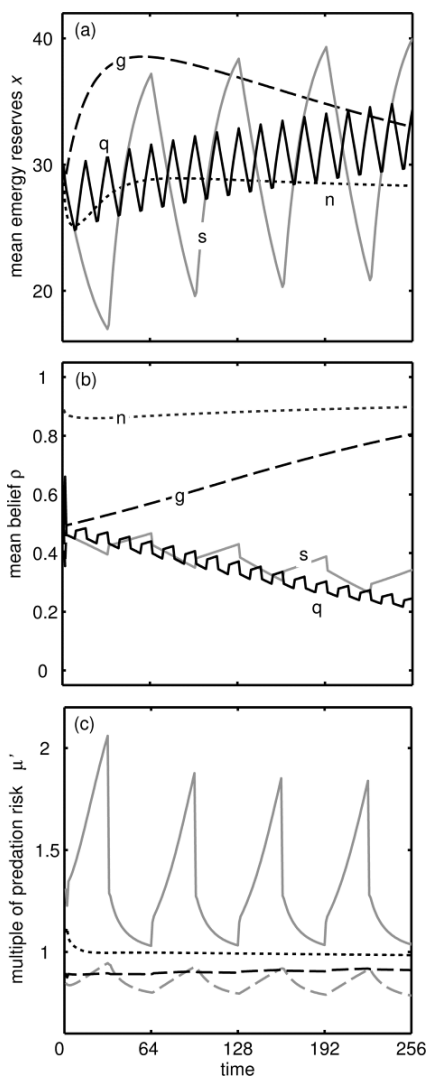
503

504 **Figure 3**: Effect of mean duration of both poor and rich periods in the Bad world ($t_{B,R}=1/\lambda_{B,R}=t_{B,P}=1/\lambda_{B,P}$, x-

505 axis) and mean duration of Poor periods in the Good world ($t_{G,P}$, shown on lines) on (a) mean reserve level in

506 the good world, (b) extra reserves storage during a glut as a proportion of reserves under normal conditions,

507 (c) extra reserves storage after a slow dieting attempt as a proportion of reserves under glut conditions, (d)

508 extra reserves storage after a quick switching dieting attempt as a proportion of reserves under glut

509 conditions.
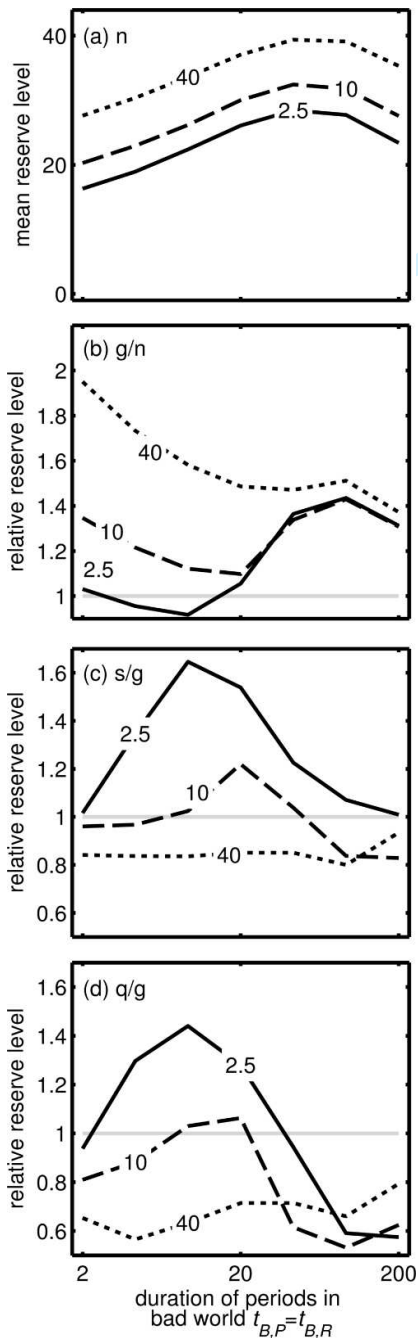
510

511 **Figure 4**: Effect of mean duration of both poor and rich periods in the Bad world ($t_{B,R}=1/\lambda_{B,R}=t_{B,P}=1/\lambda_{B,P}$, x-

512 axis) on energy storage for 10 realisations of the dependence of energy use on reserves and activity. Panels

513 show (a, b) mean reserve level in the good world, (c, d) extra reserves storage during a glut as a proportion of

514 reserves under normal conditions, (e, f) extra reserves storage after a slow dieting attempt as a proportion of

515 reserves under glut conditions, (g, h) extra reserves storage after a quick switching dieting attempt as a

516 proportion of reserves under glut conditions, and (a, c, e, f) no extra costs energy reserves on when active

517 $m_{x,f}=0$, and (b, d, f, g) energy reseves are more costly when active $m_{x,f}=0.5$. Lines are shown for various

518 values of m_x (first value: 0, 1, or 2) and $m_{x,f}$(second value: 0 or 1)
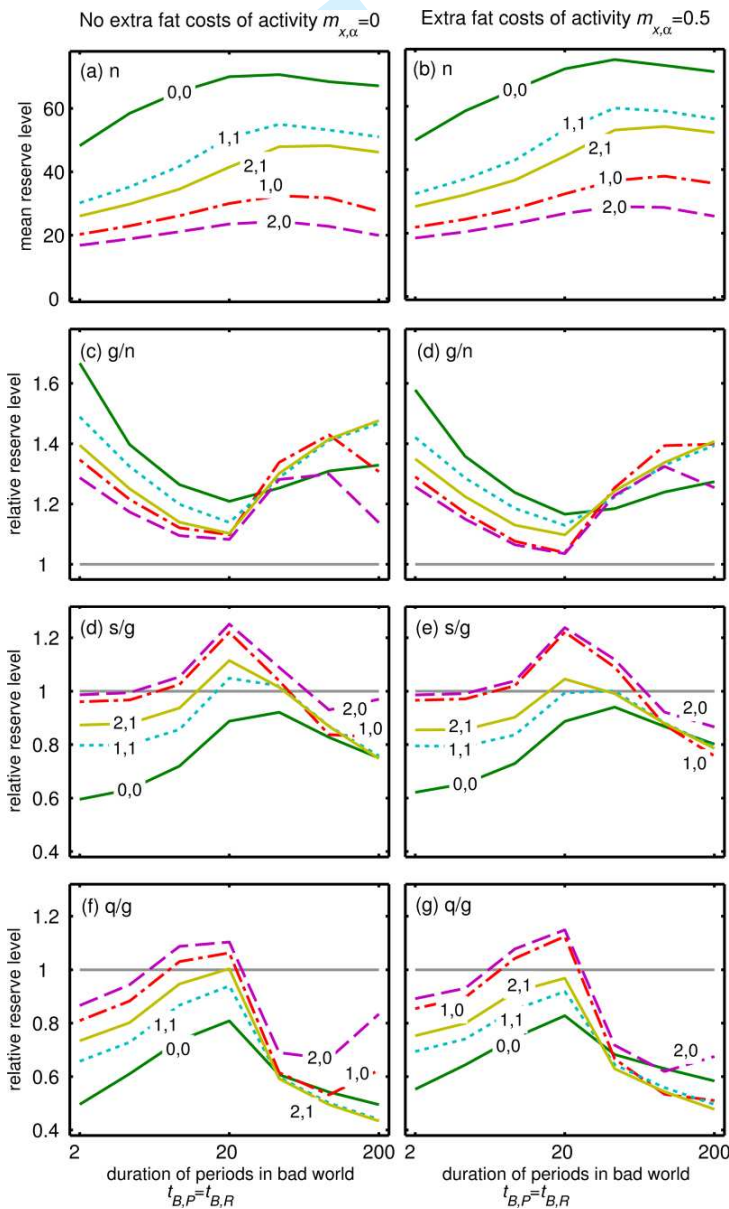
519

520

521

522   Table 1: Parameters and variables in the model and their default values

Symbol	Description	Value
Individual		
x	Energy reserves	$0 - x_{max}$
ρ	Probability that world is Good	$0 \leq \rho \leq 1$
x_{max}	Maximum level of energy reserves	100
V	Value of the animal's life	$V \geq 0$
f	Intensity of foraging	$0 \leq f \leq 1$
m_0	Magnitude of energy use	0.5
m_x	Dependence of energy use on reserves	1
m_f	Dependence of energy use on activity	0
$m_{x,f}$	Dependence of the cost of reserves on activity	0
Environmental		
b	Mean energy in food items	5.5
μ	Probability of mortality per decision epoch	0.00001
θ_W	Probability that world W changes to other world	$\theta_B{=}0.0001,\ \theta_G{=}0.0001$
$\lambda_{W,C}$	Probability that world W in condition C changes to the other condition	$\lambda_{B,P} = 0.05,\ \lambda_{B,R}{=}0.05$ $\lambda_{G,P} = 0.1,\ \lambda_{G,R}{=}0.02$
$t_{W,C}$	Mean number of decision epochs for which world W stays in condition C ($t_{W,C} = 1/\lambda_{W,C}$)	$t_{B,P} = 20,\ t_{B,R} = 20$ $t_{G,P} = 10,\ t_{G,R} = 50$
γ_C	Probability of finding food in condition C per unit time spent foraging	$\gamma_P = 0.3,\ \gamma_R = 0.7$

523

524

1   **ONLINE APPENDIX**

2   **An adaptive response to uncertainty can lead to weight gain during dieting attempts**

3   **A. D. Higginson, J. M. McNamara**

4   **Evolution, Medicine, and Public Health**

5

6   Here, we provide a complete description of the model implementation. The overall framework can be

7   summarized as follows. At a decision epoch reserves are assumed to take values in the range $x = 0, 1, 2, \cdots, s$,

8   the conditions C are either rich ($C=R$) or poor ($C=P$), and the world W can be either good ($W = $ G) or bad

9   ($W = $ B). The animal knows x and C but does not know W, instead storing the probability ρ that the world is

10   good. At each decision epoch the animal chooses the proportion of time foraging. The proportion of time

11   foraging is allowed to depend on x, ρ and C. A strategy f specifies this dependence; under f the proportion of

12   time foraging when reserves are x, the probability that the world is good ρ, and the food conditions are C

13   is $f(x, \rho, C)$. Let $V_f(x, \rho, C)$ denote the expected total future lifetime reproductive success from the current

14   epoch of an individual that is initially in state (x, ρ, C) and follows strategy f until it dies. Let

15     $V^*(x, \rho, C) = \max_f V_f(x, \rho, C),$

16   where the maximum is taken over all strategies f. If a strategy f^* satisfies $V_{f*}(x, \rho, C) = V^*(x, \rho, C)$ for all

17   combination of states (x, ρ, C) then we refer to f^* as an optimal strategy. This strategy then maximises the

18   total lifetime reproductive success of the animal for all possible initial states. Standard results from the

19   theory of Markov decision processes show that such a strategy exists (Puterman 2005). In this appendix we

20   detail how the optimal strategy can be found.

21

22   Model details are as follows. If the proportion of time spent foraging in each decision epoch t is f, then the

23   probability of finding an item of food during this time interval is $\gamma_C f$, To avoid potentially problematic grid

24   effects [36], we assume that food items are of two types (type $j = 1$ or 2) with relative abundance β_j ($\Sigma\beta_j = 1$),

25   and provide a reward of energetic value r_j. Foraging for a proportion f of a single decision epoch increases

26   the animal's lifetime reproductive success by $\sqrt{1 - f}$. There is therefore a trade-off between immediate

27   investment in reproduction and increasing the future investment by finding food to increase the expected

28   lifespan.

29

30   We assume that the forager pays an energy cost $m(x)$ to meet its metabolic needs in each decision epoch

31    $m(x) = m_0 + m_x x$,

32   where m_x controls how the cost increases with reserves and m_0 is the cost at zero reserves. The probability

33   the animal is not killed by a predator before the next decision epoch is $1 - \mu$. If the change in reserves

34   results in the new reserves being zero or below, the animal is assumed to have died of starvation. If the new

35   reserves would have been greater than the maximum value of s, then reserves are taken to be s. Thus if the

36   animal has reserves x and forages for a proportion of time f its reserves (x') at the next decision epoch given

37   it does not die of predation are

38    $x'_0 = \max(x - m(x), 0)$, **with probabilit y** $p_0 = (1 - \gamma_C f)$,

39    $x'_1 = \min(x + r_1 - m(x), s)$, **with probability** $p_1 = \gamma_C f \beta_1$,

40    $x'_2 = \min(x + r_2 - m(x), s)$, **with probabilit y** $p_2 = \gamma_C f \beta_2$.

41   Computations are based on the values $r_1 = 5$, $r_2 = 6$, $\beta_1 = 0.5$, $\beta_2 = 0.5$; so the mean energetic value of a

42   food item is b=5.5.

43

44   Let ρ denote the current probability that the world is good. This probability is updated as follows. The

45   world may have changed before the animal assesses the conditions, so the intermediate probability is

46    $\hat{\rho} = (1 - \rho)\theta_B + \rho(1 - \theta_G)$.

47   The animal knows what the conditions were at the previous decision epoch and now assesses what the

48   conditions currently are. From this and the prior probability $\hat{\rho}$ the posterior probability that the world is

49   good (ρ') is given by Bayes' rule. If conditions were poor and are still poor:

50    $\rho'_{P,P} = \rho' | P \rightarrow P = \dfrac{(1 - \lambda_{G,P})\hat{\rho}}{(1 - \lambda_{G,P})\hat{\rho} + (1 - \lambda_{B,P})(1 - \hat{\rho})}$

51   And similarly,

52 $$\rho'_{P,R} = \rho' | P \rightarrow R = \frac{\lambda_{G,P}\hat{\rho}}{\lambda_{G,P}\hat{\rho} + \lambda_{B,P}(1-\hat{\rho})}$$

53 $$\rho'_{R,P} = \rho' | R \rightarrow P = \frac{\lambda_{G,R}\hat{\rho}}{\lambda_{G,R}\hat{\rho} + \lambda_{B,R}(1-\hat{\rho})}$$

54 $$\rho'_{R,R} = \rho' | R \rightarrow R = \frac{(1-\lambda_{G,R})\hat{\rho}}{(1-\lambda_{G,R})\hat{\rho} + (1-\lambda_{B,R})(1-\hat{\rho})}$$

55 calculates ρ' if conditions were poor and are now rich; rich and are now poor; rich and are still rich,

56 respectively. See Figure A2 for values for the baseline parameter values (Table 1).

57

58 Given these ingredients, the dynamic programming operator T^* can be expressed follows. Let V be a

59 function, $V(x,\rho,C)$, of energy reserves x and the probability that conditions are good ρ and conditions C

60 satisfying $V(0,\rho,C)=0$. Then T^*V is a new function of reserves and environmental conditions that

61 satisfies $(T^*V)(0,\rho,C)=0$ for all W and $(T^*V)(x,\rho,C)=\max_f H(x,\rho,C;f,V)$ for $x>0$ and all ρ and

62 all C, where

63 $$H(x,\rho,C;f,V) = (1-\mu)\left\langle \sqrt{1-f} + \sum_{j=0}^{2} p_j \left(\begin{array}{c} V(x'_j, \rho'_{C,\neg C}, \neg C)[\lambda_{B,C}(1-\hat{\rho}) + \lambda_{G,C}\hat{\rho}] \\ + V(x'_j, \rho'_{C,C}, C)[(1-\lambda_{B,C})(1-\hat{\rho}) + (1-\lambda_{G,C})\hat{\rho}] \end{array} \right) \right\rangle$$

64 where $\neg C$ indicates the other condition (i.e. R when $C=P$; P when $C=R$).

65

66 To find the optimal strategy, we define a sequence of functions $V_0, V_1, V_2, \cdots$ iteratively as follows. Initially

67 set $V_0(0,\rho,C)=0$ for all ρ and C and $V_0(x,\rho,C)=1$ for all $x>0$ and all ρ and C. Given V_k, set

68 $V_{k+1}=T^*V_k$. Then the sequence of functions converges pointwise to a limit $V^* = \lim_{k\to\infty} V_k$ (Puterman 2005).

69 Convergence was judged to have occurred when $|V_{k+1} - V_k| < 10^{-6}$, which typically happened within 500

70 iterations. Any strategy f^* satisfying

71 $$H(x,\rho,C;f^*(x,\rho,C),V^*) = \max_f H(x,\rho,C;f,V^*) \textbf{ for all } x>0 \textbf{ and all } \rho \textbf{ and } C \quad \text{(A4)}$$

72   necessarily satisfies equation (A2), and is hence optimal (Puterman 2005). This process finds the optimal

73   Bayesian strategy $f^*(x,\rho,C)$. The optimal strategies for nine combinations of switching probabilities are

74   shown in Figure A3.

75   **Reference**

76   Puterman M. L. (2005) Markov decision processes: Discrete stochastic dynamic programming. Wiley, New

77   Jersey.

78   **Figure A1**: Examples of the probability of finding food (γ_C) over time when the world is (a) Bad and (b)

79   Good. The duration of periods between changes in γ_C ('conditions') is variable, but γ_C is more often low

80   (poor conditions) when the world is bad than when it is good.

81

82

83
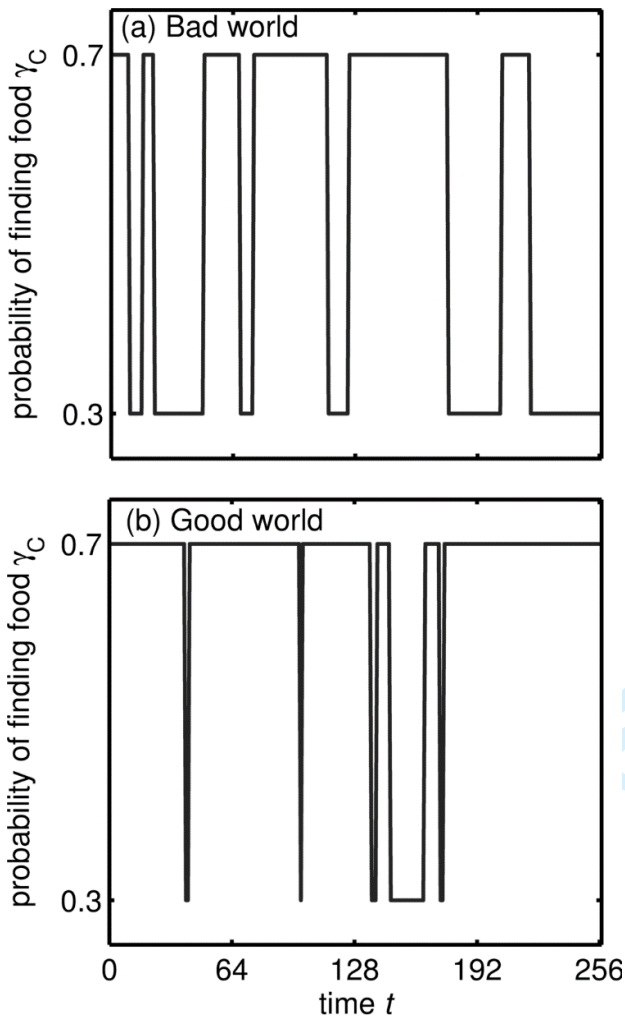
84   **Figure A2**: Probability ρ' that the world is currently good given that the probability at the previous decision

85   epoch was ρ and that conditions have changed as indicated on the lines (e.g. P->R indicates that the

86   conditions have changed from poor to rich;   R->R indicates that the conditions have remained rich).
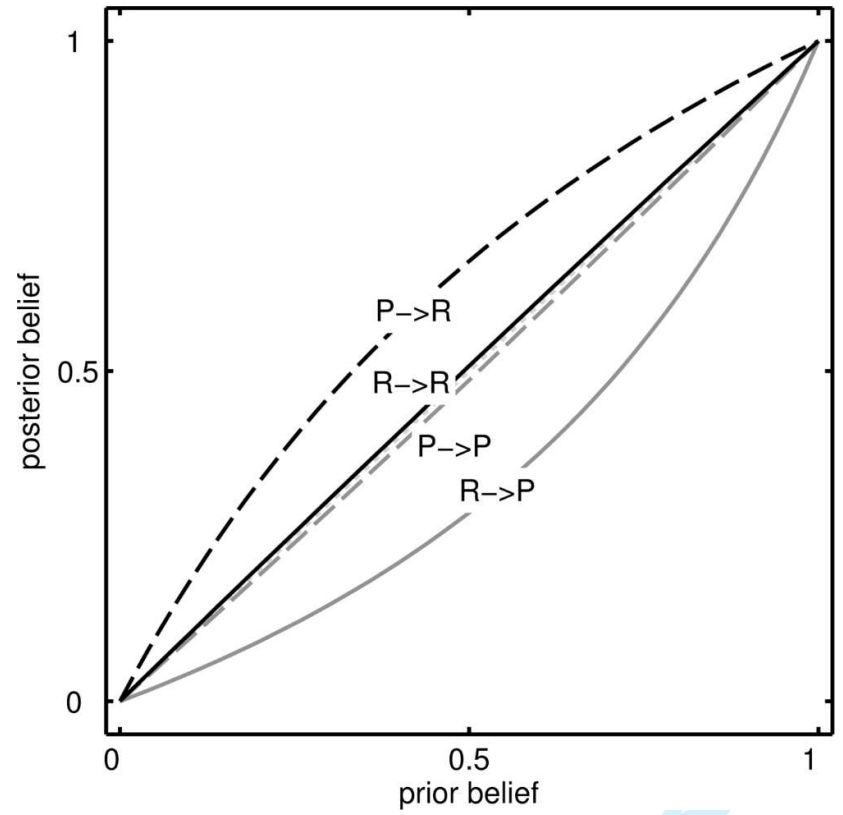
91   **Figure A3**: Effect of three 'treatments' compared to control conditions for a very long period (i.e. as Figure

92   2 but 2048 time steps). (a) Mean energy reserves x over time when conditions always Rich ('glut': g, dashed

93   line) or when conditions switch between Poor and Rich every 32 epochs ('slow dieting attempts periods': s,

94   solid grey line), or when conditions change between Poor and Good every 8 epochs ('quick dieting

95   attempts': q, solid black line), compared to the mean across Poor and Rich conditions in the Good world

96   ('control': n, dotted line). (b) Belief that the world is Good ρ for the same period and treatments. Under

97   normal conditions ρ settles down at a high level, whereas during a glut conditions are always rich so learning

98   is slower as $\lambda_{B,R} \approx \lambda_{G,R}$ but tends towards one, and in the long term energy storage in a glut tends towards that

99   under normal conditions. In dieting conditions ρ tends towards zero, and the long term mean is

100   approximately the same for quick and slow fluctuations. The arrow indicates the endpoint of Figure 2 and the

101   point at which the values shown in Figure 3 are calculated.
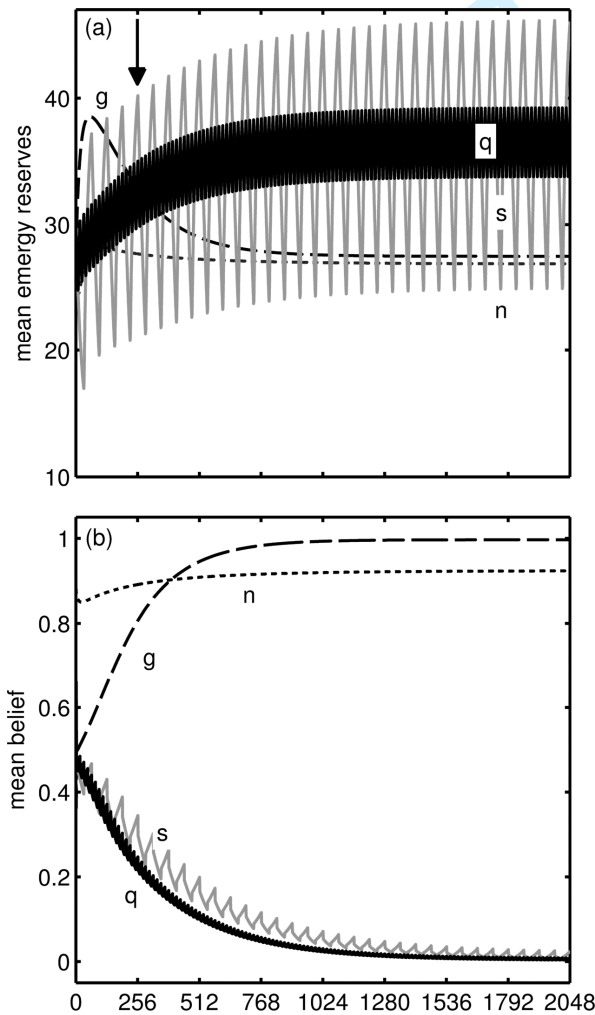
102

103   **Figure A4**: Optimal foraging intensity f^* as a function of energy reserves x for three durations of rich and

104   poor conditions in the bad world ($t_{B,C}$) and three durations of poor conditions in the bad world ($t_{G,P}$) as shown

105   on panels, to aid interpretation of Figure 3. Lines are shown for Poor (grey) and Rich (black) conditions, and

106   for $\rho=0$ (dashed) and $\rho=1$ (solid). Other parameter values as shown in Table 1.
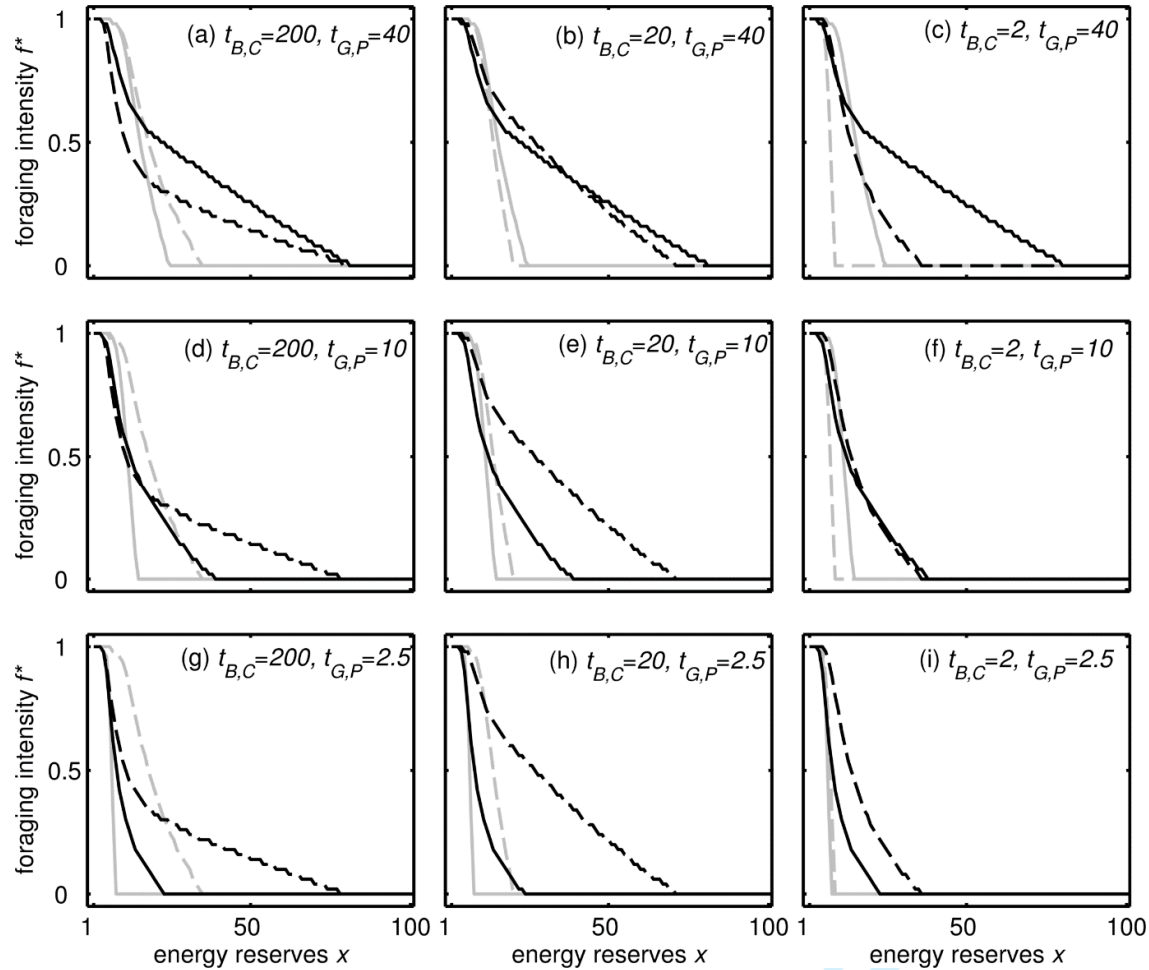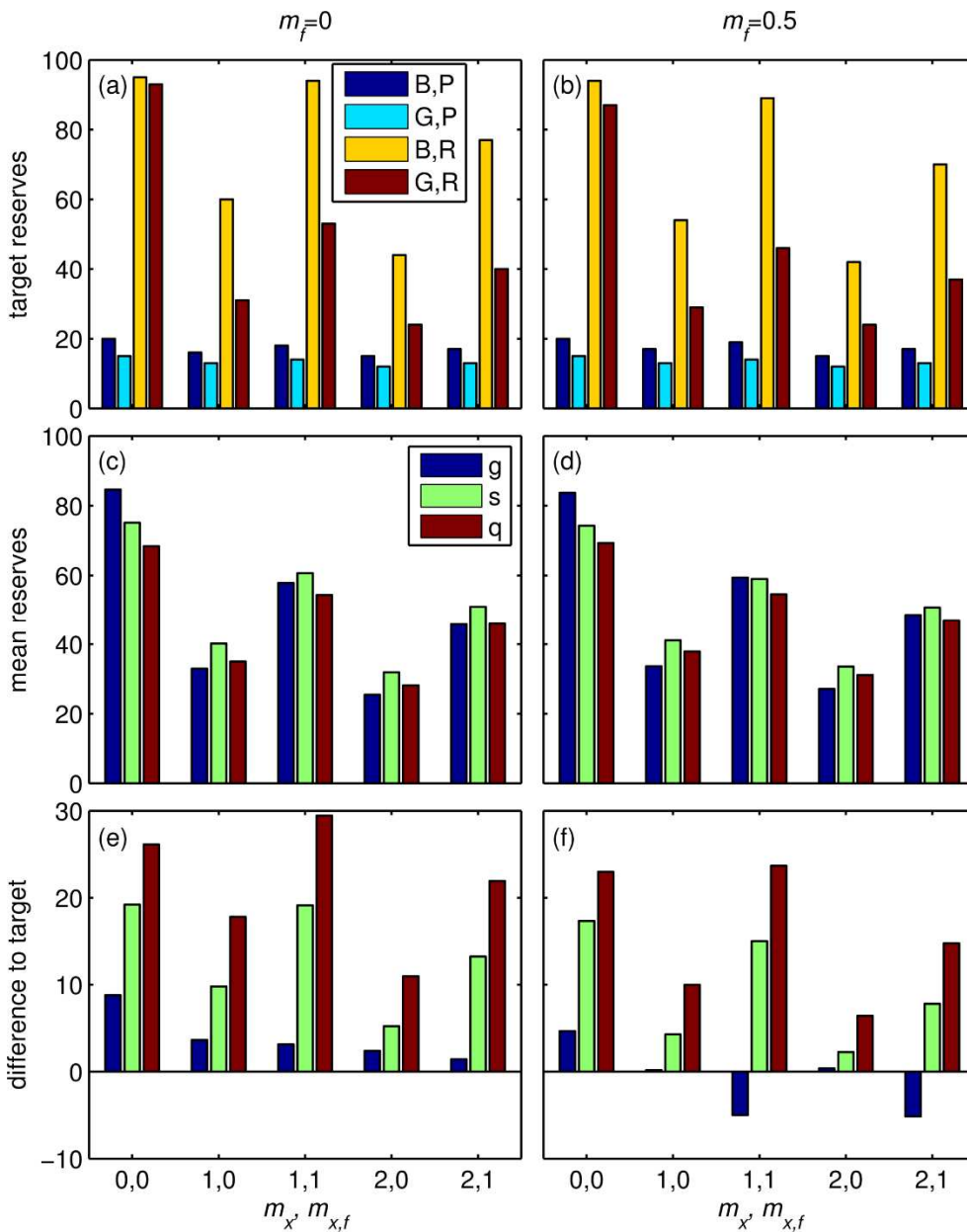

107

108

109   **Figure A5:** Optimal foraging intensity f^* as a function of energy reserves x for the types of energy costs (m_x,

110   m_f, $m_{x,f}$) shown on panels, to aid interpretation of Figure 4. Lines are shown for Poor (grey) and Rich (black)

111   conditions, and for $\rho=0$ (dashed) and $\rho=1$ (solid). Other parameter values as shown in Table 1.
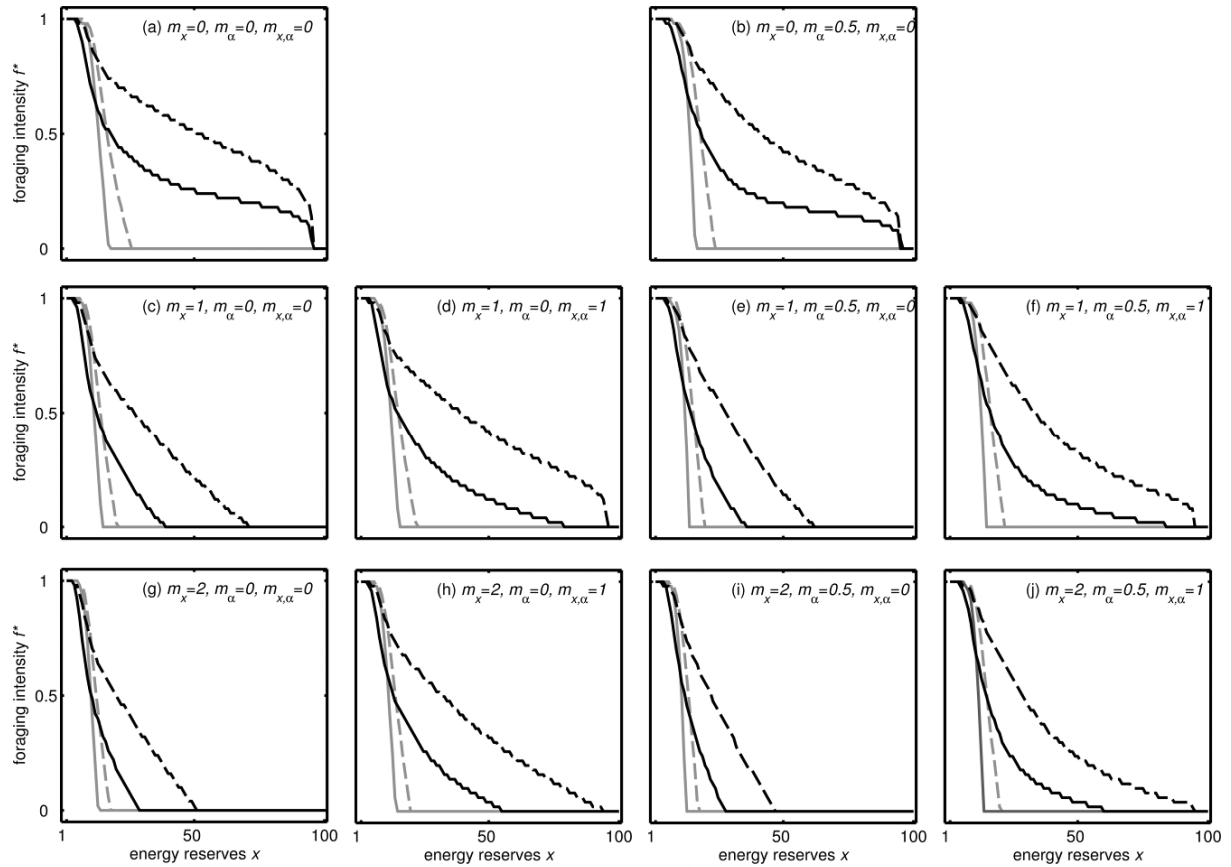
113   **Figure A6**: Target and realised reserves for various costs (a, c, e: m_f=0; b, d, f: m_f=0.5; m_x and $m_{x,f}$ as shown

114   on x-axis). (a, b) Target: the level of reserves at which the optimal foraging rate $f^*(x,\rho,C)$ is expected to

115   result in no change in reserves (x^*). Legend indicates P: poor conditions, R: rich conditions, B: certain world

116   is bad ρ=0, G: certain world is good ρ=1. (c, d) Realised mean reserves after 256 time steps for constant glut

117   (g), slow diet (s), quick diet (q). (e, f) Difference between the realised mean reserves and the target, when the

118   target is the weighted average of the target in the bad and good world, weighted by the probability the world

119   is good after 256 time steps.