

1 When is it adaptive to be patient? A general
2 framework for evaluating delayed rewards

3

4 Tim W. Fawcett^{a,*}, John M. McNamara^b, Alasdair I. Houston^a

5

6 ^a *School of Biological Sciences, University of Bristol, Woodland Road, Bristol BS8 1UG, UK*

7 (tim.fawcett@cantab.net, a.i.houston@bris.ac.uk)

8 ^b *School of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, UK*

9 (john.mcnamara@bris.ac.uk)

10

11

12 *Corresponding author. Tel.: +44 117 9287478; fax: +44 117 3317985.

13 *E-mail address:* tim.fawcett@cantab.net.

14 The tendency of animals to seek instant gratification instead of waiting for greater long-term
15 benefits has been described as impatient, impulsive or lacking in self-control. How can we
16 explain the evolution of such seemingly irrational behaviour? Here we analyse optimal
17 behaviour in a variety of simple choice situations involving delayed rewards. We show that
18 preferences for more immediate rewards should depend on a variety of factors, including
19 whether the choice is a one-off or is likely to be repeated, the information the animal has
20 about the continuing availability of the rewards and the opportunity to gain rewards through
21 alternative activities. In contrast to the common assertion that rational animals should devalue
22 delayed rewards exponentially, we find that this pattern of discounting is optimal only under
23 restricted circumstances. We predict preference reversal whenever waiting for delayed
24 rewards entails loss of opportunities elsewhere, but the direction of this reversal depends on
25 whether the animal will face the same choice repeatedly. Finally, we question the ecological
26 relevance of standard laboratory tests for impulsive behaviour, arguing that animals rarely
27 face situations analogous to the self-control paradigm in their natural environment. To
28 understand the evolution of impulsiveness, a more promising strategy would be to identify
29 decision rules that are adaptive in a realistic ecological setting, and examine how these rules
30 determine patterns of behaviour in simultaneous choice tests.

31

32 *Keywords:* delay discounting; ecological rationality; impulsiveness; intertemporal choice;
33 optimal foraging; self-control

34

35 **1. Introduction**

36

37 The way in which animals, including humans, value rewards that occur in the future is
38 of interest to a broad range of disciplines including economics (Frederick et al., 2002),
39 psychology (Mazur, 2007a,b), pharmacology (Bickel and Marsch, 2001; Reynolds, 2006),
40 neuroscience (Berns et al., 2007; Kalenscher and Pennartz, 2008; Roesch et al., 2007) and
41 behavioural ecology (Freidin et al., 2009; Kagel et al., 1986; Kacelnik, 1997, 2003; Stephens,
42 2002; Stephens and Dunlap, 2009, 2011; Stevens and Mühlhoff, in press; Stevens et al.,
43 2005a). Frequently, studies find that animals reject delayed rewards in favour of more
44 immediate gratification, even when they would gain greater long-term benefits by waiting
45 (Ainslie, 1974; Bateson and Kacelnik, 1996; Henly et al., 2008; Mazur, 1987; McDiarmid
46 and Rilling, 1965; Rachlin and Green, 1972; Stephens and Anderson, 2001). Such behaviour
47 has been described as impatient (Kacelnik, 2003), impulsive (Henly et al., 2008), short-
48 sighted (Stephens and Anderson, 2001) or lacking in self-control (Mazur and Logue, 1978).
49 Why is it that animals behave in this way?

50 Models of behaviour can be categorised as descriptive or normative (Kacelnik, 1997).
51 Descriptive models summarise what animals do whereas normative models specify what they
52 ought to do (Houston et al., 2007; Shapiro et al., 2008). In the context of how animals
53 evaluate delayed rewards, descriptive models focus on the quantitative details of preferences
54 measured in the laboratory and seek a mechanistic explanation for the precise patterns we
55 observe (e.g. Mazur, 2006). Typically, these models do not attempt to explain why particular
56 discounting mechanisms have evolved. Normative models, in contrast, adopt a functional
57 perspective and try to understand the evolutionary basis of decision making, asking how
58 natural selection will shape preferences under natural conditions (e.g. Stephens et al., 2004).
59 In these models the mechanistic underpinnings of the evolved preferences are usually not

60 considered. The two approaches are clearly closely related, because animals tested in the
61 laboratory are using rules that were shaped in their ancestral environment; but it does not
62 follow that all aspects of laboratory behaviour will be optimal (Houston and McNamara,
63 1989, 1999; McNamara, 1996; McNamara and Houston, 1980).

64 Our aim in this article is to present a simple and general framework for understanding
65 how natural selection shapes the evaluation of delayed rewards. Thus, our emphasis is on the
66 functional (normative) approach. We wish to shed light on the following problem: when
67 faced with a choice between options with differing delays, what should an optimal decision
68 maker do? What is the precise pattern of discounting it should use to devalue delayed
69 rewards?

70

71 *1.1. Costs of being patient*

72 There are two main reasons why it might be costly, in fitness terms, to wait for a
73 delayed reward. First, there is a risk that the anticipated reward may become unavailable
74 before it can be collected (*collection risk*; Houston et al., 1982). In this context, Stephens
75 (2002) distinguishes between an interruption risk—the chance of losing the next food item,
76 for example because a conspecific competitor eats it first—and a termination risk—the
77 chance that an entire sequence of foraging is cut short, for example because of the sudden
78 appearance of a predator. Second, even if collection is guaranteed, there may be lost
79 opportunities associated with the time spent waiting (*opportunity cost*; Stephens, 2002):
80 assuming the animal cannot perform other activities while it is waiting, it forgoes the
81 opportunity to gain rewards by other means (McNamara, 1982). As we will see below, both
82 the collection risk and the opportunity cost can strongly influence optimal behaviour.

83

84 *1.2. Empirical facts to explain*

85 To give a satisfying explanation of preferences for delayed rewards, there are some key
86 empirical results that functional models need to account for: hyperbolic discounting,
87 preference reversal and effects of reward magnitude.

88

89 *1.2.1. Hyperbolic discounting*

90 If delayed rewards are discounted at a constant rate per unit of time spent waiting, their
91 perceived value decays according to an exponential function. For example, if an immediate
92 reward loses half of its value when it is delayed by 5 minutes, one delayed by 10 minutes
93 should be half as valuable again (i.e. its value should drop to one quarter of its immediate
94 value). It is generally agreed that an exponential pattern of discounting should result when
95 interruptions occur randomly over time (Dasgupta and Maskin, 2005; Green and Myerson,
96 1996; Stevens, 2010). However, empirical data suggest that discounting is not exponential
97 but hyperbolic, the discounting rate gradually falling with added delay (Ainslie, 1974; Mazur,
98 1987, 2006). This implies that additional delays do not have much effect on reward valuation
99 if the delays are already long, in contrast to exponential discounting in which the discount
100 rate does not change.

101

102 *1.2.2. Preference reversal*

103 In the classic 'self-control' paradigm (Mazur and Logue, 1978; Fig. 1a), in which an
104 animal is given a choice between a small reward delivered after a short delay (SS) and a
105 larger reward after a long delay (LL), it typically shows an impulsive preference for the
106 former option (Ainslie, 1974; Bateson and Kacelnik, 1996; Henly et al., 2008; Mazur, 1987;
107 McDiarmid and Rilling, 1965; Rachlin and Green, 1972; Stephens and Anderson, 2001).
108 Increasing both delays by the same amount, however, can sometimes induce a switch to the
109 more delayed option. This preference reversal has been reported in pigeons (Ainslie and

110 Herrnstein, 1981; Green et al., 1981; Rachlin and Green, 1972), rats (Green and Estle, 2003)
111 and humans, the latter for both hypothetical (Green et al., 1994) and actual (Kirby and
112 Herrnstein, 1995) amounts of money. Exponential discounting does not predict preference
113 reversal, whereas hyperbolic discounting can (Kalenscher and Pennartz, 2008).

114 There are two possible forms of preference reversal that are regularly discussed in the
115 literature, but often not clearly distinguished. The first form occurs across two different
116 choice situations, involving the same reward magnitudes but with an added delay in one
117 situation; the animal prefers the more immediate option when the delays are short (Fig. 1a)
118 and the more delayed option when they are extended (Fig. 1b). The other occurs within the
119 same choice situation, as time runs forwards: having initially chosen the later reward, the
120 animal may switch its preference to the sooner reward as its collection point approaches (Fig.
121 1c). These forms are often treated equivalently (e.g. Casari, 2009; Kalenscher and Pennartz,
122 2008; Kirby and Herrnstein, 1995; Sozou, 1998) but, as we shall see below, whether it is
123 valid to do so depends on what we assume about the information that is available to the
124 animal.

125

126 *1.2.3. Effect of reward magnitude*

127 Several accounts of choice assume that delayed options have a value given by $R/f(D)$,
128 where R is the reward and $f(D)$ is some positive increasing function of delay D . Such
129 accounts predict that choice will be unaffected when the magnitude of the rewards is
130 changed, provided their ratio (R_1/R_2) is kept constant. However, humans discount delayed
131 rewards less strongly when they are choosing between larger amounts of money (Green et al.,
132 1997, 1999; Myerson and Green, 1995). Similarly, capuchin monkeys (*Cebus apella*) tested
133 in two separate self-control studies (Addessi et al., 2011; Amici et al., 2008) were
134 significantly more tolerant for delay when the rewards were larger (2 vs. 6 food items, as

135 opposed to 1 vs. 3; for discussion, see Addessi et al., 2011). Rats show a lower rate of
136 discounting for less concentrated sucrose solutions, which they prefer (Farrar et al., 2003),
137 but a higher rate of discounting for larger amounts of food (Wogar et al., 1992). Green et al.
138 (2004) found no effect of reward magnitude in pigeons and rats.

139

140 1.3. Seeking an adaptive explanation

141 There have been several previous attempts to explain impulsiveness and the evolution
142 of hyperbolic discounting. Kagel et al. (1986) proposed that if animals are uncertain of the
143 rate of interruptions ('hazard' rate) and have to estimate this, they should gradually lower
144 their estimate as time passes while they are waiting for a reward. If a long time has elapsed
145 and an interruption has still not occurred, this indicates that the underlying hazard rate is
146 likely to be low and an animal should therefore be more willing to wait even longer for a
147 reward. Sozou (1998) developed this idea into a formal model and showed that estimation of
148 a constant but unknown hazard rate could produce hyperbolic discounting.

149 Stephens and colleagues (Stephens, 2002; Stephens et al., 2004) put forward an
150 alternative explanation for impulsive choice, based on constraints on discrimination.
151 Assuming that animals can detect a difference between two delays more easily when the
152 delays are short than when they are long (an example of Weber's Law; Gibbon, 1977),
153 decisions might be more accurate when made on the basis of short-term consequences. This
154 increased accuracy might favour a general tendency to evaluate options in terms of short-term
155 gains. An alternative approach argues that a hyperbolic decay function can be explained by
156 assuming that discounting is based on subjective time perception (Takahashi, 2005;
157 Zauberman et al., 2009).

158 While interesting and potentially important, these ideas rely on additional factors—
159 uncertainty over the interruption rate, or biases in discrimination—to explain impulsiveness,

160 on top of the basic economic considerations of energy (benefit) and time (cost). Here we take
161 a more fundamental approach. We seek to identify optimal decisions in a variety of simple
162 choice situations in which the available options differ only in the size of the reward and the
163 delay till that reward can be collected. The focal animal knows (i.e. is adapted to) the
164 interruption rate and can discriminate between the options accurately. Under these conditions,
165 it has repeatedly been claimed that a rational animal should discount delayed rewards
166 exponentially (e.g. Bickel and Marsch, 2001; Kalenscher and Pennartz, 2008; Kirby and
167 Herrnstein, 1995). We show that this view is unfounded. Optimal choice between delayed
168 rewards can cover a variety of different patterns of discounting, depending on whether a
169 given choice is likely to be repeated (Kacelnik, 1997, 2003; Stephens, 2002) and what
170 alternative options the animal may have for gaining energy outside the current choice
171 situation. Our aim is not to develop one definitive model of choice that accounts for all the
172 empirical observations mentioned above, but to construct a general framework for
173 investigating these kinds of problems and expose the logic of evaluating delayed rewards.

174

175

176 **2. A general model of choice between delayed rewards**

177

178 We consider an animal facing a choice between different foraging options, each of
179 which offers a reward after some delay. Our overarching assumption is that natural selection
180 acts on the total reward obtained by some final time T (Houston and McNamara 1999;
181 McNamara and Houston, 1986, 1987). For a given option i the net energetic gain from the
182 reward is R_i , but this is only collected after a delay D_i . Given that it has to wait for the
183 reward, there is a risk that the animal may lose it before it can be collected (e.g. because a
184 competitor consumes it first or a predator interrupts the foraging bout). If we assume a

185 constant interruption rate α during a delay of duration D_i , then the chance that the animal
 186 successfully collects the reward is given by the negative exponential function $e^{-\alpha D_i}$ and the
 187 discounted value of the food reward is therefore $R_i e^{-\alpha D_i}$. Following this delay D_i , we assume
 188 that the animal forages at some rate of gain γ for the remaining time $T - D_i$. Thus its expected
 189 total reward by the final time T is

$$190 \quad H_i = R_i e^{-\alpha D_i} + \gamma(T - D_i). \quad (1)$$

191 If the future gains do not depend on the current options the animal is facing, γT is common to
 192 all options and the best option maximises $R_i e^{-\alpha D_i} - \gamma D_i$. When facing a choice between
 193 several alternatives differing in the reward amount R_i and the delay D_i , we can identify the
 194 best option graphically by plotting $R_i e^{-\alpha D_i}$ against D_i , as shown in Fig. 2.

195 The expression $R_i e^{-\alpha D_i} - \gamma D_i$ neatly captures the essential trade-off between the
 196 energetic gain from the chosen option and the cost of waiting for it. The term γD_i is an
 197 opportunity cost (McNamara, 1982): it is the energetic gain that would have been achieved
 198 by seeking rewards elsewhere. Thus γ represents the opportunity cost per unit time. There are
 199 different possible interpretations of γ , depending on the situation we are modelling. If the
 200 animal faces a one-off choice, then γ is simply a ‘background’ rate of energetic gain that is
 201 independent of the options available in the choice situation; it is an externally imposed
 202 parameter. If, on the other hand, the animal faces the same choice situation repeatedly, then γ
 203 is the long-term rate of gain on the choice cycles and is determined by the rewards and delays
 204 of the options available (Kacelnik, 1997, 2003). The importance of this distinction will
 205 become clear in the detailed models presented below.

206 In keeping with most empirical work on time discounting, we focus on choice decisions
 207 between two options, as illustrated in Fig. 1. One option (smaller-sooner, SS) offers a

208 relatively small reward R_{SS} after a short delay D_{SS} , while the other option (larger–later, LL)
 209 offers a larger reward R_{LL} after a longer delay D_{LL} . By definition, $R_{LL} > R_{SS}$ and $D_{LL} > D_{SS}$.

210

211

212 3. One-off choice

213

214 The simplest situation is where the animal faces a one-off choice between SS and LL
 215 and then reverts to some background foraging rate γ . There are two basic cases we need to
 216 consider.

217

218 3.1. Zero opportunity cost ($\gamma = 0$)

219 If the background foraging rate is zero, the animal cannot gain energy through any
 220 other means outside the choice situation; its gains are restricted to the two options SS and LL.
 221 In this case there is no opportunity cost of waiting for a delayed reward, so $\gamma = 0$. Then the
 222 best option maximises $R_i e^{-\alpha D_i}$, which implies that choice should be based on standard
 223 exponential discounting. If collection is guaranteed ($\alpha = 0$) then the animal should simply
 224 wait for the option with the bigger reward (LL), whereas a high risk of interruption favours
 225 the more immediate option (SS). In general (i.e. for any value of α), the animal should choose
 226 the SS option whenever $R_{SS} e^{-\alpha D_{SS}} > R_{LL} e^{-\alpha D_{LL}}$, which after rearranging gives the condition

$$227 \quad D_{LL} - D_{SS} > \frac{1}{\alpha} \ln \left(\frac{R_{LL}}{R_{SS}} \right). \quad (2)$$

228 The difference $D_{LL} - D_{SS}$ is unchanged when a constant delay is added to both options,
 229 implying that there should be no preference reversal (Fig. 3a). There is also no effect of
 230 reward magnitude on choice provided the reward ratio R_{LL}/R_{SS} does not change.

231

232 3.2. *Non-zero opportunity cost ($\gamma > 0$)*

233 When there is some background rate of gain $\gamma > 0$, this will influence the animal's
 234 optimal decision. It should now maximise $R_i e^{-\alpha D_i} - \gamma D_i$, trading off the potential gains from
 235 the options available in the choice situation against the opportunity cost of not being able to
 236 forage at the background rate while it is waiting for a reward. For short delays, the animal
 237 should prefer the LL option if the reward R_{LL} is sufficiently large. If the delays are increased,
 238 however, there comes a point at which the expected rate of gain from the current choice
 239 situation drops below the background rate of gain. When this happens, the animal should exit
 240 the choice situation as soon as possible, which is achieved by choosing the option with the
 241 shorter delay. So as a constant delay is added to both options, this model predicts a reversal
 242 of preference from the LL to the SS option (Fig. 3b).

243 We can prove this mathematically. The animal should choose the SS option whenever
 244 $R_{SS} e^{-\alpha D_{SS}} - \gamma D_{SS} > R_{LL} e^{-\alpha D_{LL}} - \gamma D_{LL}$, which after rearranging gives the condition

$$245 \quad D_{LL} - D_{SS} > \frac{1}{\gamma} (R_{LL} e^{-\alpha D_{LL}} - R_{SS} e^{-\alpha D_{SS}}). \quad (3)$$

246 For relatively short delays and a sufficiently large value of R_{LL} , this inequality will not be
 247 satisfied and so the animal should choose the LL option. Now consider the effect of adding a
 248 constant delay to both options. The difference $D_{LL} - D_{SS}$ will not change whereas the right-
 249 hand side will get smaller, tending to zero as the amount of delay added goes to infinity. As
 250 soon as the right-hand side is smaller than $D_{LL} - D_{SS}$, the inequality is satisfied and the
 251 animal should switch its preference to the SS option.

252 In this choice situation there is also an effect of reward magnitude: for a given reward
 253 ratio R_{LL}/R_{SS} , larger rewards will be discounted less strongly than smaller rewards.

254

255

256 4. Repeated choice

257

258 We now consider cases where the animal faces the same choice repeatedly. Once the
 259 delay for its chosen option has elapsed and the animal has collected its reward, the cycle
 260 begins anew with the same two options (SS and LL) available. This changes the economics of
 261 the situation because instead of a fixed background rate, γ now depends on the rewards and
 262 delays of the options chosen on the choice cycles (Kacelnik, 1997, 2003).

263 To start with, we look at the general case in which the animal adopts behaviour pattern
 264 u over the repeated cycles of choice (u can represent any aspect of behaviour, but we avoid
 265 being specific about this here). Its rate of gain in this situation is the reward obtained per unit
 266 of time spent waiting, or $R(u)/D(u)$. The behaviour u^* that maximises this rate is found by
 267 differentiating $R(u)/D(u)$ with respect to u and setting it equal to zero, which after some
 268 rearrangement gives

$$269 \quad R'(u^*) \cdot D(u^*) = R(u^*) \cdot D'(u^*) \quad (4)$$

270 (where primes denote the first derivative with respect to u). If we denote the maximum
 271 possible rate of gain as $\gamma^* = R(u^*)/D(u^*)$ and substitute this into equation (4), we get

$$272 \quad R'(u^*) - \gamma^* D'(u^*) = 0. \quad (5)$$

273 Note that this is equivalent to maximising $R(u^*) - \gamma^* D(u^*)$, which has exactly the same
 274 form as the general model outlined in section 2 (for an alternative derivation, see McNamara,
 275 1982). Thus the optimal behaviour in a situation of repeated choice is just a special case of
 276 this general model.

277

278 4.1. Infinite number of cycles

279 If the sequence of cycles continues indefinitely, γ is entirely determined by the options
 280 chosen. The long-term rate of gain from repeatedly choosing option i is $\gamma = R_i e^{-\alpha D_i} / D_i$,
 281 which when substituted into equation (1) gives an expected pay-off of

$$282 \quad H_i = \frac{R_i e^{-\alpha D_i}}{D_i} T. \quad (6)$$

283 The animal should choose whichever option maximises this pay-off, i.e. the option that gives
 284 the higher rate of gain $R_i e^{-\alpha D_i} / D_i$. Thus when there is an infinite sequence of cycles, the
 285 animal should follow a strategy of rate maximisation. This predicts a preference reversal from
 286 the SS to the LL option as the delays for both options are increased by a fixed amount (Fig.
 287 4). Choice is unaffected by reward magnitude provided the reward ratio R_{LL}/R_{SS} does not
 288 change.

289

290 4.2. Uncertain number of cycles

291 Lastly, we consider what happens if there are repeated cycles of choice, but it is
 292 uncertain how long the sequence will continue. We now assume that interruptions, when they
 293 occur, terminate the entire sequence of cycles. This could represent the arrival of a predator,
 294 for example, or of a dominant competitor who displaces the focal animal from the foraging
 295 patch, forcing it to seek gains elsewhere (Houston et al., 1982; Kagel et al., 1986). Such
 296 events happen stochastically at an average rate λ and immediately afterwards the animal
 297 switches to some background rate of gain γ .

298 Let the random variable Y denote the time elapsed before the sequence is terminated,
 299 and N_i denote the number of cycles completed in this period given that the animal repeatedly
 300 chooses option i . We can write the expected values of these variables as $E(Y)$ and $E(N_i)$,
 301 respectively. The animal gains reward amount R_i for each completed cycle and then forages

302 at the background gain rate γ for the remaining time $T - E(Y)$ after the termination has
 303 occurred, so its expected pay-off H_i is

$$304 \quad H_i = R_i \cdot E(N_i) + \gamma(T - E(Y)). \quad (7)$$

305 The expected time before the sequence is terminated is simply the reciprocal of the
 306 termination rate, that is $E(Y) = 1/\lambda$. For any given cycle of duration D_i the chance that
 307 termination does not occur is $e^{-\lambda D_i}$, so the chance that the sequence is terminated after n
 308 cycles is $P(N_i = n) = (1 - e^{-\lambda D_i}) \cdot e^{-n\lambda D_i}$. If T is sufficiently large we can treat the possible
 309 values of N_i as an infinite sequence, which gives the expected number of completed cycles as

$$310 \quad E(N_i) = (1 - e^{-\lambda D_i}) \sum_{n=0}^{\infty} n e^{-n\lambda D_i}. \quad (8)$$

311 Since $e^{-\lambda D_i} < 1$, the infinite series in this equation converges to $\sum_{n=0}^{\infty} n e^{-n\lambda D_i} = e^{-\lambda D_i} / (1 - e^{-\lambda D_i})^2$,

312 which leaves us with $E(N_i) = e^{-\lambda D_i} / (1 - e^{-\lambda D_i}) = (e^{\lambda D_i} - 1)^{-1}$. Substituting the expressions for
 313 $E(Y)$ and $E(N_i)$ back into the pay-off equation (7), we get

$$314 \quad H_i = \frac{R_i}{e^{\lambda D_i} - 1} + \gamma \left(T - \frac{1}{\lambda} \right). \quad (9)$$

315 Since $\gamma(T - 1/\lambda)$ is common to all options, the animal should choose whichever option
 316 maximises $R_i / (e^{\lambda D_i} - 1)$. So it should choose the SS option whenever

$$317 \quad \frac{R_{SS}}{R_{LL}} > \frac{e^{\lambda D_{SS}} - 1}{e^{\lambda D_{LL}} - 1}. \quad (10)$$

318 This inequality is satisfied when the SS option gives an immediate reward ($D_{SS} = 0$), since the
 319 right-hand side is zero. As both delays are increased by the same amount, however, the right-
 320 hand side increases and eventually converges to a value of $e^{-\lambda k}$, where $k = D_{LL} - D_{SS}$. This
 321 implies that preference reversal will occur if the rewards and delays of the two options are

322 such that $R_{SS}/R_{LL} < \exp[-\lambda(D_{LL} - D_{SS})]$, with preference switching to the LL option as both
323 delays are increased. An example of this is shown in Fig. 5.

324 Note that in this situation there is no effect of reward magnitude on choice provided the
325 reward ratio R_{SS}/R_{LL} does not change.

326

327

328 5. Preference reversals over time

329

330 We have seen that an optimality approach can predict preference reversal when the
331 delays associated with two options are increased by the same amount. Experimentally, this
332 scenario corresponds to a comparison between two separate choice situations: in one, the
333 animal is given a choice between two rewards after delays D_{SS} and D_{LL} (Fig. 1a); in the other,
334 it is given a choice between the same two rewards after delays $D_{SS} + \delta$ and $D_{LL} + \delta$ (Fig. 1b).
335 Preference reversal between these two situations has been documented by several studies
336 (e.g. Ainslie and Herrnstein, 1981; Green et al., 1981; Green and Estle, 2003; Rachlin and
337 Green, 1972), with animals preferring a smaller, sooner reward in the former case but a
338 larger, later reward when the delays are extended to $D_{SS} + \delta$ and $D_{LL} + \delta$.

339 There is another form of preference reversal that we have not yet considered in detail.
340 After making an initial choice between two options, an animal may have the opportunity to
341 reverse its decision at a later time point, when the delays to both options have decreased (Fig.
342 1c). Again the comparison is between a choice when the delays are $D_{SS} + \delta$ and $D_{LL} + \delta$
343 (initial choice) and a choice when the delays are D_{SS} and D_{LL} (later choice), but now we are
344 dealing with a preference reversal within the same choice situation, as time runs forwards
345 from t to $t + \delta$. If the animal switches its choice at the later time point $t + \delta$, it is not obvious

346 why it would not choose this option in the first place. Is it ever adaptive for an animal to
347 reverse its choice in this way?

348 A number of authors (e.g. Casari, 2009; Kalenscher and Pennartz, 2008; Kirby and
349 Herrnstein, 1995; Sozou, 1998) have treated preference reversal over time as the reverse case
350 of preference reversal when a constant delay is added to both options: instead of both delays
351 being extended by the same amount, both delays are shortened by the same amount. But in
352 fact these two cases are distinct, and the failure to distinguish between them can lead to
353 misunderstandings. For example, Sozou's (1998) model of hyperbolic discounting, in which
354 individuals estimate the underlying hazard rate, predicts greater patience (increased
355 preference for the LL option) when a fixed delay is added to both options, a pattern supported
356 by empirical studies comparing two separate choice situations (e.g. Rachlin and Green,
357 1972). Dasgupta and Maskin (2005) later used Sozou's logic to predict what would happen
358 when individuals estimate the underlying hazard rate within a single choice situation, and
359 argued that it incorrectly predicts increasing patience as time runs forwards. However, to
360 analyse this type of situation rigorously, an explicit account of the process is needed. Whether
361 we should expect preference reversal over time depends on how the passage of time affects
362 the economics of the choice situation. Specifically, the pattern of choice depends on the
363 information the animal has about the continuing availability of the rewards. Dasgupta and
364 Maskin (2005) alluded to this point, but they focused on a more complicated choice situation
365 in which the delays to the two rewards are uncertain. Here we state the distinction in more
366 general terms. There are two possible scenarios:

367 *Case 1: the passage of time changes the estimated probability of collecting a given*
368 *reward.* When making its initial choice between the SS option and the LL option, both of
369 these options are available to the animal; but assuming a certain risk that its chosen option
370 will be lost during the delay ($D_{SS} + \delta$ or $D_{LL} + \delta$) it has to wait before it can collect the

371 reward, the reward value should be discounted accordingly. Now imagine that when time has
372 run forwards to $t + \delta$, both options are still available; neither has been lost during the
373 preceding period. If the animal can update its assessment of the collection risk to take account
374 of this fact, the economics of the choice situation have changed. The preceding period can be
375 ignored and the animal should discount only over the remaining delay, D_{SS} or D_{LL} . This is the
376 inverse of the situations considered earlier (where a constant delay was added to both
377 options), and can therefore support preference reversal over time as the optimal behaviour
378 under some conditions. In Fig. 6 we illustrate this for a one-off choice with a non-zero
379 opportunity cost (Fig. 6a) and for an infinite sequence of repeated choices (Fig. 6b).
380 Preference reversal occurs in opposite directions in these two situations, as was the case when
381 a constant delay was added to both options (sections 3.2 and 4.1). Most empirical data
382 support the pattern shown in Fig. 6b, in which the animal becomes increasingly impatient as
383 time passes.

384 *Case 2: the passage of time has no effect on the estimated probability of collecting a*
385 *given reward.* Alternatively, the animal may not know at the later time point whether either
386 of the options is still available. Although there is less time remaining before it can collect its
387 chosen reward, it should also take into account the chance the reward was lost during the
388 preceding period from t to $t + \delta$. Thus it should still discount the rewards over the original
389 delays $D_{SS} + \delta$ and $D_{LL} + \delta$. No preference reversal is expected in this case: the animal faces
390 exactly the same economic situation as before, so it should stick by its original decision.

391 So, whether we expect an animal to show preference reversal over time depends
392 critically on the information it has about the continuing availability of the options as time
393 passes (Dasgupta and Maskin, 2005). Many previous studies have overlooked this key
394 consideration or have otherwise conflated two distinct types of preference reversal.

395

396

397 **6. Summary of predictions**

398

399 In Table 1 we summarise the main features of optimal behaviour in the various
400 different choice situations we have considered. Starting from some relatively modest
401 assumptions, our general model yields a surprisingly rich array of predictions. Preference
402 reversals may occur whenever there is a non-zero opportunity cost, in other words whenever
403 the animal loses opportunities to forage elsewhere while it is waiting for delayed rewards.
404 However, the expected direction of preference reversal depends on whether the choice
405 situation is a one-off or is repeated. When choice is repeated, optimal behaviour leads to the
406 form of preference reversals documented in the empirical literature, with greater patience for
407 more delayed rewards. However, under these same conditions we predict no effect of the
408 reward ratio R_{LL}/R_{SS} . Conversely, in a one-off choice with a non-zero opportunity cost we
409 predict lower rates of discounting for larger rewards, but preference reversals in the opposite
410 direction to that typically seen in experiments. Thus, although our model successfully
411 predicts isolated features of intertemporal choice behaviour, no single version of the model
412 can account for all of the empirically observed patterns. To understand how animals evaluate
413 delayed rewards, it seems that we need to take into account additional factors besides the
414 ones we have focused on here, collection risk and opportunity cost. For example, we might
415 incorporate certain constraints on decision making, such as discrimination biases (Stephens,
416 2002; Stephens et al., 2004) or uncertain interruption rates (Sozou, 1998). At the same time,
417 however, it is important to question whether animals are likely to be adapted to the
418 intertemporal choice situations they encounter in the laboratory. In the next section we
419 critically evaluate the ecological relevance of choice between delayed rewards.

420

421

422 **7. Ecological relevance**

423

424 We have examined a series of simple choice scenarios in which a foraging animal is
425 fully adapted to the rewards (R_i) and delays (D_i) of alternative options, as well as to the
426 frequency (α or λ) with which interruptions occur. This has been a useful exercise for
427 identifying what choices the animal should make to maximise its long-term energy gain in
428 these specific situations. But why would we expect animals to have evolved an ability to
429 choose between rewards with different delays? What kinds of natural situations would entail
430 such a choice, in which the animal has access to a given option but the reward cannot be
431 harvested until a later point in time? Under what circumstances might patience be
432 ‘ecologically rational’ (Gigerenzer et al., 1999)? Several possibilities have been proposed.

433

434 *6.1. Fruit ripening (Dasgupta and Maskin, 2005; Stevens and Stephens, 2008)*

435 When a frugivore encounters unripe fruit, it is faced with a choice between eating it
436 immediately or waiting until it has ripened, in which case the energetic reward it gains will be
437 greater. If, as seems likely, the animal is free to forage elsewhere while it is waiting for the
438 fruit to ripen, this situation might reasonably be modelled as a one-off choice with zero
439 opportunity cost. Assuming a constant collection risk (e.g. a risk that competing frugivores
440 consume the fruit in the meantime), this predicts standard exponential discounting. However,
441 since fruit is likely to become increasingly attractive to foragers as it ripens, the collection
442 risk actually rises as time passes and so the assumption of a constant α is invalid. In addition,
443 the timescale of fruit ripening is far greater than the delays used in self-control experiments,
444 which typically last a few seconds or minutes. It is not yet known how animals devalue food
445 items that they cannot eat until days or weeks later.

446

447 *6.2. Extractive foraging (Stevens and Stephens, 2008)*

448 Some foods (e.g. nuts, shellfish) have to be extracted from hard, inedible cases before
449 they can be consumed. Although the handling time imposes a fixed delay to the reward, it is
450 unclear how well this corresponds to the self-control paradigm studied experimentally. First,
451 it is not obvious what the alternative, more immediate option is, unless the animal has a range
452 of different food types it can exploit in the same habitat. Second, it seems likely that in most
453 cases of extractive foraging the collection risk will be negligible. Finally, from a mechanistic
454 rather than functional perspective, the animal may already gain some psychological
455 reinforcement from handling a food item before it has extracted the food (Shettleworth and
456 Jordan, 1986).

457

458 *6.3. Caching for the winter (Stevens and Stephens, 2008, 2009; Stevens, 2010)*

459 A variety of birds and mammals cache food for later use, and this has been interpreted
460 as a preference for a delayed reward over immediate consumption. On closer inspection,
461 however, this type of behaviour differs in important ways from the kind of situation studied
462 in impulsiveness tests in the laboratory. When an animal faces a choice between caching a
463 given food item or eating it now, it could be argued that the immediate and delayed options
464 have the same reward magnitude, whereas in the self-control paradigm the delayed reward is
465 bigger (generating a conflict between reward size and time cost). If anything, cached food
466 will provide a smaller net energetic gain when it is eventually consumed, because of decay
467 and the energetic cost of recovering it (e.g. digging it up). The fitness value of the food item
468 may nevertheless be greater when it is recovered and eaten in midwinter than if it was eaten
469 when found earlier in the year, because the background rate of gain from foraging has
470 declined dramatically and the animal may be closer to starvation. But this situation is rather

471 different from the self-control paradigm, in which rewards are delayed by seconds or minutes
472 and changes in the animal's state can be disregarded. Instead, the decision to cache seems to
473 be driven by other factors that are missing from the self-control set-up. Most probably, the
474 animal has been selected to cache for the winter to guard against the risk of energetic shortfall
475 during a predictable period of poor foraging success. At the same time, caching may allow it
476 to use additional resources when it is already satiated, as suggested by the fact that caching
477 typically occurs at a time when excess food is available (Smith and Reichman, 1984).

478

479 *6.4. Patch leaving (Stephens and Dunlap, 2009, 2011; Stephens et al., 2004; Stevens and*
480 *Stephens, 2009; Stevens, 2010)*

481 Stephens and colleagues have framed the classic patch-leaving problem of behavioural
482 ecology in terms of the self-control paradigm. In this view, the decision to remain in a given
483 foraging patch and keep searching for additional food items represents choice for a more
484 immediate reward, whereas leaving the patch and travelling to a new one represents choice
485 for a delayed, but potentially larger, reward (Stephens et al., 2004). When the distribution of
486 food is highly clustered there is a clear opportunity cost, since while travelling between
487 patches the animal cannot continue to feed. However, as Stephens and Dunlap (2009) point
488 out, the patch-leaving problem differs from the self-control problem in having a nested
489 decision structure: choice of one option ('stay') requires the animal eventually to choose the
490 other option ('leave') before the same choice is repeated. Interestingly, blue jays (*Cyanocitta*
491 *cristata*) make better long-term decisions in the self-control situation (Stephens and Dunlap,
492 2009).

493

494 *6.5. Sequential mate search (Stevens, 2010)*

495 Discussions of intertemporal choice typically revolve around foraging decisions, but
496 similar issues may apply in other domains. In a mate-choice context, a female's decision to
497 reject a low-quality male in the hope she will later find a superior mate could be viewed as
498 choice for a larger, later reward. This is not exactly equivalent to the standard self-control
499 paradigm, in particular because the delay to the later option and the size of the associated
500 reward (the exact quality of the superior mate) are both uncertain, but there are some
501 intriguing parallels. This kind of situation might also generate some interesting
502 complications, such as changes in the degree of impulsiveness over time; for example,
503 unpaired females are likely to become increasingly impatient as the end of the mating season
504 approaches. An analogous effect has been noted for diving animals, which should become
505 less selective in their foraging as they near the time at which they have to return to the surface
506 for air (Houston and McNamara, 1985).

507
508 In summary, despite some superficial similarities, there appear to be few—if any—biological
509 situations that correspond directly to the self-control paradigm used in laboratory tests of
510 impulsive behaviour. On this basis, we question whether it is reasonable to expect that natural
511 selection has furnished animals with the decision rules for behaving rationally (i.e. in a way
512 that maximises their gains) in these particular experimental set-ups (Houston, 2009; Houston
513 et al., 2007; McNamara and Houston, 1980, 2009). Some of the apparently short-sighted
514 behaviours observed in the laboratory may be the product of rules that work well in more
515 naturalistic situations such as patch exploitation (Stephens 2002). Future work on the
516 adaptive basis of impulsiveness should identify what kinds of rules perform well in
517 ecologically relevant scenarios and use these rules to predict behavioural patterns in
518 laboratory experiments like the self-control paradigm. The current trend for *post hoc*
519 ecological explanations of why certain taxa exhibit higher or lower discounting rates in

520 laboratory experiments (e.g. Addessi et al., 2011; Cheng et al., 2002; Rosati et al., 2007;
521 Stevens et al., 2005a,b) needs to be paired with a predictive, model-based approach to
522 understanding animal behaviour.

523

524

525 **8. Key points**

526

527 We finish by summarising some key points from our analysis.

- 528 1. Depending on the precise situation considered, optimality models of choice between
529 delayed rewards can predict a range of different types of behaviour—including no
530 preference reversal, preference reversal in either direction, lower discounting rates for
531 bigger rewards or no effect of reward magnitude. Exponential discounting is expected only
532 under certain circumstances.
- 533 2. To predict how an animal should respond, we need to know more than just the rewards
534 and delays of the available options. We also need to know what information is available to
535 the animal and what it perceives about the current situation. Of critical importance is
536 whether the animal has evolved to expect one-off choices, repeated choices or can adjust
537 its behaviour flexibly depending on the persistence of the current situation.
- 538 3. Preference reversals over time are not equivalent to preference reversals across separate
539 choice situations. Whether preferences should reverse over time depends on the
540 information the animal has about the continuing availability of the rewards.
- 541 4. The structure of the self-control paradigm used in laboratory experiments does not fit most
542 intertemporal choice situations in the natural environment. Expecting animals to behave
543 rationally in self-control tests might therefore be unrealistic.

544 Acknowledgements

545

546 A.I.H. and J.M.M. came up with the idea for this paper. All three authors developed the idea
547 and T.W.F. wrote the paper. We thank the European Research Council for financial support
548 (Advanced Grant 250209 to A.I.H.), Peter Sozou for discussion and Andy Higginson, Fabio
549 Paglieri, Jeff Stevens, Pete Trimmer, Marco Vasconcelos and an anonymous referee for
550 comments on an earlier version of the manuscript.

551

552

553 References

554

555 Addessi, E., Paglieri, F., Focaroli, V., 2011. The ecological rationality of delay tolerance:
556 insights from capuchin monkeys. *Cognition* 119, 142–147.

557 Ainslie, G., 1974. Impulse control in pigeons. *J. Exp. Anal. Behav.* 21, 485–489.

558 Ainslie, G., Herrnstein, R.J., 1981. Preference reversal and delayed reinforcement. *Anim.*
559 *Learn. Behav.* 9, 476–482.

560 Amici, F., Aureli, F., Call, J., 2008. Fission–fusion dynamics, behavioral flexibility, and
561 inhibitory control in primates. *Curr. Biol.* 18, 1415–1419.

562 Bateson, M., Kacelnik, A., 1996. Rate currencies and the foraging starling: the fallacy of the
563 averages revisited. *Behav. Ecol.* 7, 341–352.

564 Berns, G.S., Laibson, D., Loewenstein, G., 2007. Intertemporal choice—toward an
565 integrative framework. *Trends Cogn. Sci.* 11, 482–488.

566 Bickel, W.K., Marsh, L.A., 2001. Toward a behavioral economic understanding of drug
567 dependence: delay discounting processes. *Addiction* 96, 73–86.

- 568 Casari, M., 2009. Pre-commitment and flexibility in a time decision experiment. *J. Risk.*
569 *Uncertain.* 38, 117–141.
- 570 Cheng, K., Peña, J., Porter, M.A., Irwin, J.D., 2002. Self-control in honeybees. *Psychon.*
571 *Bull. Rev.* 9, 259–263.
- 572 Dasgupta, P., Maskin, E., 2005. Uncertainty and hyperbolic discounting. *Am. Econ. Rev.* 95,
573 1290–1299.
- 574 Farrar, A.M., Kieres, A.K., Hausknecht, K.A., de Wit, H., Richards, J.B., 2003. Effects of
575 reinforcer magnitude on an animal model of impulsive behavior. *Behav. Processes* 64,
576 261–271.
- 577 Frederick, S., Loewenstein, G., O'Donoghue, T., 2002. Time discounting and time
578 preference: a critical review. *J. Econ. Lit.* 40, 351–401.
- 579 Freidin, E., Aw, J., Kacelnik, A., 2009. Sequential and simultaneous choices: testing the diet
580 selection and sequential choice models. *Behav. Processes* 80, 218–223.
- 581 Gibbon, J., 1977. Scalar expectancy theory and Weber's law in animal timing. *Psychol. Rev.*
582 84, 279–325.
- 583 Gigerenzer, G., Todd, P.M., ABC Research Group (Eds), 1999. *Simple Heuristics that Make*
584 *Us Smart.* Oxford University Press, New York.
- 585 Green, L., Estle, S.J., 2003. Preference reversals with food and water reinforcers in rats. *J.*
586 *Exp. Anal. Behav.* 79, 233–242.
- 587 Green, L., Myerson, J., 1996. Exponential versus hyperbolic discounting of delayed
588 outcomes: risk and waiting time. *Am. Zool.* 36, 496–505.
- 589 Green, L., Fisher, E.B., Perlow, S., Sherman, L., 1981. Preference reversal and self-control:
590 choice as a function of reward amount and delay. *Behav. Anal. Lett.* 1, 43–51.
- 591 Green, L., Fristoe, N., Myerson, J., 1994. Temporal discounting and preference reversals in
592 choice between delayed outcomes. *Psychon. Bull. Rev.* 1, 383–389.

- 593 Green, L., Myerson, J., McFadden, E., 1997. Rate of temporal discounting decreases with
594 amount of reward. *Mem. Cognit.* 25, 715–723.
- 595 Green, L., Myerson, J., O'Donoghue, P., 1999. Amount of reward has opposite effects on the
596 discounting of delayed and probabilistic outcomes. *J. Exp. Psychol. Learn. Mem. Cogn.*
597 25, 418–427.
- 598 Green, L., Myerson, J., Holt, D.D., Slevin, J.R., Estle, S.J., 2004. Discounting of delayed
599 food rewards in pigeons and rats: is there a magnitude effect? *J. Exp. Anal. Behav.* 81,
600 39–50.
- 601 Henly, S.E., Ostdiek, A., Blackwell, E., Knutie, S., Dunlap, A.S., Stephens, D.W., 2008. The
602 discounting-by-interruptions hypothesis: model and experiment. *Behav. Ecol.* 19, 154–
603 162.
- 604 Houston, A.I., 2009. Flying in the face of nature. *Behav. Processes* 80, 295–305.
- 605 Houston, A.I., McNamara, J.M., 1985. A general theory of central-place foraging for single-
606 prey loaders. *Theor. Popul. Biol.* 28, 233–262.
- 607 Houston, A.I., McNamara, J.M., 1989. The value of food: effects of open and closed
608 economies. *Anim. Behav.* 37, 546–562.
- 609 Houston, A.I., McNamara, J.M., 1999. *Models of Adaptive Behaviour*. Cambridge University
610 Press, Cambridge.
- 611 Houston, A.I., Kacelnik, A., McNamara, J.M., 1982. Some learning rules for acquiring
612 information. In: McFarland, D. (Ed.), *Functional Ontogeny*. Pitman, London, pp. 140–
613 191.
- 614 Houston, A.I., McNamara, J.M., Steer, M.D., 2007. Do we expect natural selection to
615 produce rational behaviour? *Phil. Trans. R. Soc. B* 362, 1531–1543.
- 616 Kacelnik, A., 1997. Normative and descriptive models of decision making: time discounting
617 and risk sensitivity. In: Bock, G.R., Cardew, G. (Eds.), *Characterizing Human*

- 618 Psychological Adaptations (Ciba Foundation Symposium, vol. 208). Wiley, Chichester,
619 pp. 51–70.
- 620 Kacelnik, A., 2003. The evolution of patience. In: Loewenstein, G., Read, D., Baumeister,
621 R.F. (Eds.), *Time and Decision: Economic and Psychological Perspectives on*
622 *Intertemporal Choice*. Russell Sage Foundation, New York, pp. 115–138.
- 623 Kagel, J.H., Green, L., Caraco, T., 1986. When foragers discount the future: constraint or
624 adaptation? *Anim. Behav.* 34, 271–283.
- 625 Kalenscher, T., Pennartz, C.M.A., 2008. Is a bird in the hand worth two in the future? The
626 neuroeconomics of intertemporal decision-making. *Prog. Neurobiol.* 84, 284–315.
- 627 Kirby, K.N., Herrnstein, R.J., 1995. Preference reversals due to myopic discounting of
628 delayed reward. *Psychol. Sci.* 6, 83–89.
- 629 Mazur, J.E., 1987. An adjusting procedure for studying delayed reinforcement. In: Rachlin,
630 H. (Ed.), *Quantitative Analyses of Behavior: the Effect of Delay and of Intervening*
631 *Events on Reinforcement Value*. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 55–
632 73.
- 633 Mazur, J.E., 2006. Mathematical models and the experimental analysis of behavior. *J. Exp.*
634 *Anal. Behav.* 85, 275–291.
- 635 Mazur, J.E., 2007a. Rats' choices between one and two delayed reinforcers. *Learn. Behav.*
636 35, 169–176.
- 637 Mazur, J.E., 2007b. Species differences between rats and pigeons in choices with
638 probabilistic and delayed reinforcers. *Behav. Processes* 75, 220–224.
- 639 Mazur, J.E., Logue, A.W., 1978. Choice in a “self-control” paradigm: effects of a fading
640 procedure. *J. Exp. Anal. Behav.* 30, 11–17.
- 641 McDiarmid, C.F., Rilling, M.E., 1965. Reinforcement delay and reinforcement rate as
642 determinants of schedule preference. *Psychon. Sci.* 2, 195–196.

- 643 McNamara, J.M., 1982. Optimal patch use in a stochastic environment. *Theor. Popul. Biol.*
644 21, 269–288.
- 645 McNamara, J.M., 1996. Risk-prone behaviour under rules which have evolved in a changing
646 environment. *Am. Zool.* 36, 484–495.
- 647 McNamara, J.M., Houston, A.I., 1980. The application of statistical decision theory to animal
648 behaviour. *J. Theor. Biol.* 85, 673–690.
- 649 McNamara, J.M., Houston, A.I., 1986. The common currency for behavioral decisions. *Am.*
650 *Nat.* 127, 358–378.
- 651 McNamara, J.M., Houston, A.I., 1987. A general framework for understanding the effects of
652 variability and interruptions on foraging behaviour. *Acta Biotheor.* 36, 3–22.
- 653 McNamara, J.M., Houston, A.I., 2009. Integrating function and mechanism. *Trends Ecol.*
654 *Evol.* 24, 670–675.
- 655 Myerson, J., Green, L., 1995. Discounting of delayed rewards: models of individual choice. *J.*
656 *Exp. Anal. Behav.* 64, 263–276.
- 657 Rachlin, H., Green, L., 1972. Commitment, choice and self-control. *J. Exp. Anal. Behav.* 17,
658 15–22.
- 659 Reynolds, B., 2006. A review of delay-discounting research with humans: relations to drug
660 use and gambling. *Behav. Pharmacol.* 17, 651–667.
- 661 Roesch, M.R., Calu, D.J., Schoenbaum, G., 2007. Dopamine neurons encode the better option
662 in rats deciding between differently delayed or sized rewards. *Nat. Neurosci.* 10, 1615–
663 1624.
- 664 Rosati, A.G., Stevens, J.R., Hare, B., Hauser, M.D., 2007. The evolutionary origins of human
665 patience: temporal preferences in chimpanzees, bonobos, and human adults. *Curr. Biol.*
666 17, 1663–1668.

- 667 Shapiro, M.S., Siller, S., Kacelnik, A., 2008. Simultaneous and sequential choice as a
668 function of reward delay and magnitude: normative, descriptive and process-based
669 models tested in the European starling (*Sturnus vulgaris*). *J. Exp. Psychol. Anim.*
670 *Behav. Process.* 34, 75–93.
- 671 Shettleworth, S.J., Jordan, V., 1986. Rats prefer handling food to waiting for it. *Anim. Behav.*
672 34, 925–927.
- 673 Smith, C.C., Reichman, O.J., 1984. The evolution of food caching by birds and mammals.
674 *Ann. Rev. Ecol. Syst.* 15, 329–351.
- 675 Sozou, P.D., 1998. On hyperbolic discounting and uncertain hazard rates. *Proc. R. Soc. B*
676 265, 2015–2020.
- 677 Stephens, D.W., 2002. Discrimination, discounting and impulsivity: a role for an
678 informational constraint. *Phil. Trans. R. Soc. B* 357, 1527–1537.
- 679 Stephens, D.W., Anderson, D., 2001. The adaptive value of preference for immediacy: when
680 shortsighted rules have farsighted consequences. *Behav. Ecol.* 12, 330–339.
- 681 Stephens, D.W., Dunlap, A.S., 2009. Why do animals make better choices in patch-leaving
682 problems? *Behav. Processes* 80, 252–260.
- 683 Stephens, D.W., Dunlap, A.S., 2011. Patch exploitation as choice: symmetric choice in an
684 asymmetric situation? *Anim. Behav.* 81, 683–689.
- 685 Stephens, D.W., Kerr, B., Fernández-Juricic, E., 2004. Impulsiveness without discounting:
686 the ecological rationality hypothesis. *Proc. R Soc. B* 271, 2459–2465.
- 687 Stevens, J.R., 2010. Intertemporal choice. In: Breed, M., Moore, J. (Eds.), *Encyclopedia of*
688 *Animal Behavior*. Academic Press, Oxford, pp. 203–208.
- 689 Stevens, J.R., Stephens, D.W., 2008. Patience. *Curr. Biol.* 18, R11–R12.

- 690 Stevens, J.R., Stephens, D.W., 2009. The adaptive nature of impulsivity. In: Madden, G.J.,
691 Bickel, W.K. (Eds.), *Impulsivity: the Behavioral and Neurological Science of*
692 *Discounting*. APA Press, Washington, DC, pp. 361–387.
- 693 Stevens, J.R., Mühlhoff, N., in press. Intertemporal choice in lemurs. *Behav. Processes* (this
694 issue).
- 695 Stevens, J.R., Hallinan, E.V., Hauser, M.D., 2005a. The ecology and evolution of patience in
696 two New World monkeys. *Biol. Lett.* 1, 223–226.
- 697 Stevens, J.R., Rosati, A.G., Ross, K.R., Hauser, M.D., 2005b. Will travel for food: spatial
698 discounting in two New World monkeys. *Curr. Biol.* 15, 1855–1860.
- 699 Takahashi, T., 2005. Loss of self-control in intertemporal choice may be attributable to
700 logarithmic time-perception. *Med. Hypotheses* 65, 691–693.
- 701 Wogar, M.A., Bradshaw, C.M., Szabadi, E., 1992. Choice between delayed reinforcers in an
702 adjusting-delay schedule: the effects of absolute reinforcer size and deprivation level.
703 *Q. J. Exp. Psychol. B* 45, 1–13.
- 704 Zauberman, G., Kim, B.K., Malkoc, S., Bettman, J.R., 2009. Discounting time and time
705 discounting: subjective time perception and intertemporal preferences. *J. Mark. Res.* 46,
706 543–556.
- 707

708 **Table 1.** Summary of optimal behaviour in a number of simple situations involving a choice
 709 between a smaller, sooner (SS) option offering a relatively small reward (R_{SS}) after a short
 710 delay (D_{SS}) and a larger, later option (LL) offering a larger reward (R_{LL}) after a longer delay
 711 (D_{LL}). While waiting for a delayed reward there is an opportunity cost γ per unit time, plus a
 712 risk either that random interruptions eliminate the chosen reward (average interruption rate α)
 713 or that they terminate the entire foraging sequence (average termination rate λ).
 714

Choice situation (n = number of cycles)	Quantity maximised	Preference reversal		Large rewards discounted less/more/same?
		with added delay	over time*	
One-off choice ($n = 1$) <i>no opportunity cost</i>	$R_i e^{-\alpha D_i}$	none	none	same
<i>with opportunity cost</i>	$R_i e^{-\alpha D_i} - \gamma D_i$	LL \rightarrow SS	SS \rightarrow LL	less
Repeated choice, uncertain n	$R_i / (e^{\lambda D_i} - 1)$	SS \rightarrow LL	LL \rightarrow SS	same
Repeated choice, infinite n	$(R_i e^{-\alpha D_i}) / D_i$	SS \rightarrow LL	LL \rightarrow SS	same

715

716 * Assuming that the animal always knows that both options are still available. (In the absence of this

717 information, no preference reversal is predicted.)

718

719 **Fig. 1.** A diagrammatic illustration of the self-control paradigm. Time runs from left to right,
 720 and the choices of a hypothetical animal are indicated by thick lines. **(a)** At the point
 721 indicated by the question mark, the animal chooses between a relatively small reward (R_{SS})
 722 delivered after a relatively short delay (D_{SS}) and a larger reward (R_{LL}) delivered after a longer
 723 delay (D_{LL}). The animal is said to choose impulsively if it prefers the more immediate option
 724 (SS) even when the more delayed option (LL) offers a higher rate of gain. **(b)** Preference
 725 reversal across separate choice situations: in a similar choice situation in which the delays
 726 have been extended to $D_{SS} + \delta$ and $D_{LL} + \delta$, the animal may instead prefer the more delayed
 727 option (LL). **(c)** Preference reversal over time: if the animal has the opportunity to alter its
 728 initial decision after time δ has elapsed, it may switch to the more immediate option (SS).
 729

730 **Fig. 2.** A graphical method for identifying the best option from a discrete set of alternatives
 731 (each represented by a circle) differing in their reward amount R_i and delay D_i , where
 732 interruptions to foraging occur at rate α and the opportunity cost per unit time of waiting for
 733 delayed rewards is γ . When there is no opportunity cost ($\gamma = 0$), the best option maximises the
 734 expected energetic gain $R_i e^{-\alpha D_i}$. Lines of constant fitness (dashed line) are horizontal and the
 735 best option (labelled A) is the one that reaches the highest point along the vertical axis. When
 736 there is a non-zero opportunity cost ($\gamma > 0$), the total opportunity cost γD_i increases with the
 737 time spent waiting, so lines of constant fitness (dotted line) slope upwards with increasing
 738 delay (since a greater expected energetic gain is needed to compensate for a longer delay).
 739 The best option (labelled B) maximises the expected energetic gain minus the total
 740 opportunity cost γD_i .
 741

742 **Fig. 3.** Change in the expected pay-offs from choosing a smaller-sooner (SS) reward (grey)
 743 or a larger-later (LL) reward (black) in a one-off choice situation, when an extra delay is

744 added to both options. The rewards offered are $R_{SS} = 1$ for the SS option and $R_{LL} = 11$ for the
745 LL option, after delays of $D_{SS} = 1$ and $D_{LL} = 10$ plus the added delay. The dashed line
746 indicates that the LL reward is devalued even when there is no added delay, since $D_{LL} > D_{SS}$.
747 The rate of interruptions is $\alpha = 0.1$ per unit time and the total time available is $T = 500$. **(a)**
748 When there is no opportunity cost of waiting for a reward ($\gamma = 0$), preference reversal does
749 not occur. **(b)** When there is an opportunity cost of waiting for a reward ($\gamma = 0.1$ per unit
750 time), preference reversal can occur, with the optimal choice switching from LL to SS as the
751 added delay increases.

752

753 **Fig. 4.** Change in the expected pay-offs from choosing a smaller–sooner (SS) reward (grey)
754 or a larger–later (LL) reward (black) in continually repeated cycles of the same choice
755 situation, when an extra delay is added to both options. The rewards offered are $R_{SS} = 1$ for
756 the SS option and $R_{LL} = 11$ for the LL option, after delays of $D_{SS} = 1$ and $D_{LL} = 10$ plus the
757 added delay. The dashed line indicates that the LL reward is devalued even when there is no
758 added delay, since $D_{LL} > D_{SS}$. The rate of interruptions is $\alpha = 0.1$ per unit time and the total
759 time available is $T = 500$. Preference reversal can occur, with the optimal choice switching
760 from SS to LL as the added delay increases.

761

762 **Fig. 5.** Change in the expected pay-offs from choosing a smaller–sooner (SS) reward (grey)
763 or a larger–later (LL) reward (black) in an uncertain number of cycles of the same choice
764 situation, when an extra delay is added to both options. The rewards offered are $R_{SS} = 1$ for
765 the SS option and $R_{LL} = 11$ for the LL option, after delays of $D_{SS} = 1$ and $D_{LL} = 10$ plus the
766 added delay. The dashed line indicates that the LL reward is devalued even when there is no
767 added delay, since $D_{LL} > D_{SS}$. The rate at which random events terminate the entire choice

768 sequence is $\lambda = 0.1$ per unit time and the total time available is $T = 500$. Preference reversal
769 can occur, with the optimal choice switching from SS to LL as the added delay increases.

770

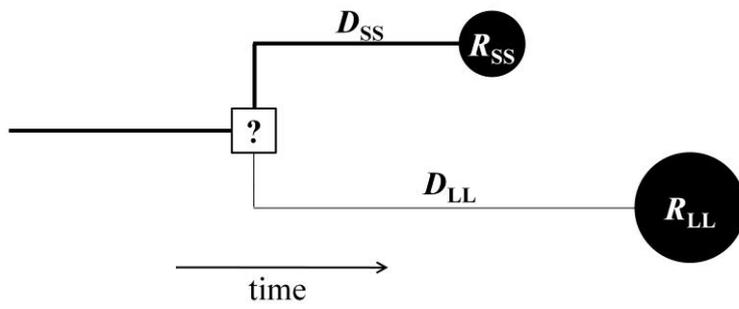
771 **Fig. 6.** The expected pay-offs from choosing a smaller–sooner (SS) reward (grey) or a larger–
772 later (LL) reward (black) in a reversible choice situation, in which the animal has the
773 opportunity to switch to a previously rejected option before the associated delay expires and
774 the reward can be collected (indicated by the dashed lines). The rewards offered are $R_{SS} = 1$
775 for the SS option and $R_{LL} = 11$ for the LL option, after initial delays of $D_{SS} = 36$ and $D_{LL} =$
776 45. The rate of interruptions is $\alpha = 0.1$ per unit time and the total time available is $T = 500$. As
777 time passes, the remaining delays for both options decrease. This can cause preference
778 reversal if the animal has information that both rewards are still available. **(a)** In a one-off
779 choice situation with an opportunity cost of waiting for a reward ($\gamma = 0.1$ per unit time), the
780 best option changes from SS to LL as time passes. This is the reverse case of Fig. 3b. **(b)**
781 When the choice situation is continually repeated, the best option changes from LL to SS as
782 time passes. This is the reverse case of Fig. 4. Note that when the time to collection falls to
783 zero, the expected pay-off (dashed lines) is infinite.

784

785 **Fig. 1**

786

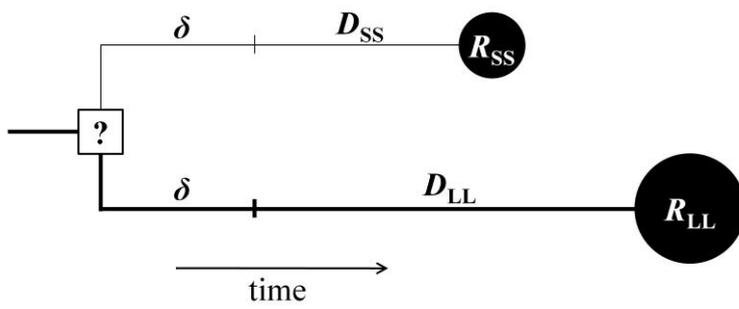
787 (a)



788

789

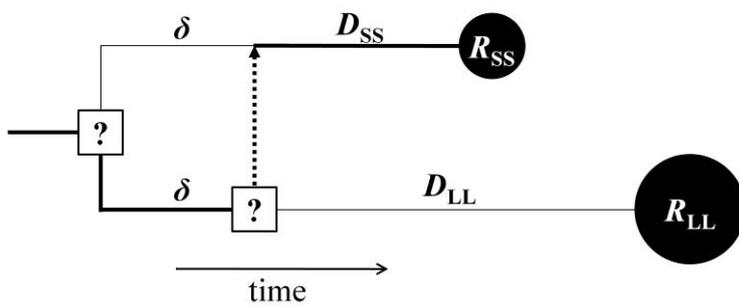
790 (b)



791

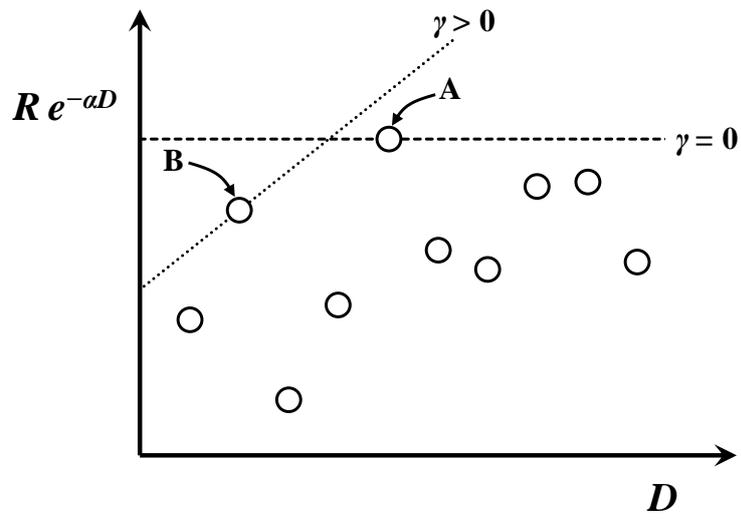
792

793 (c)



794

795

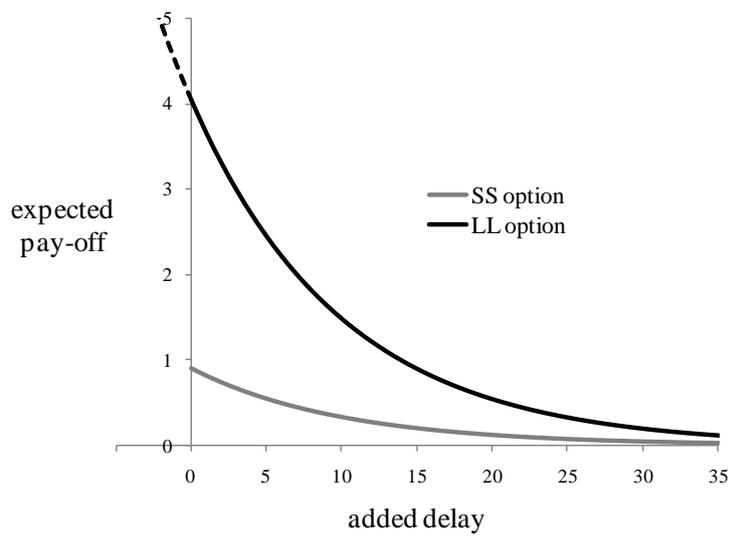
796 **Fig. 2**

797

798

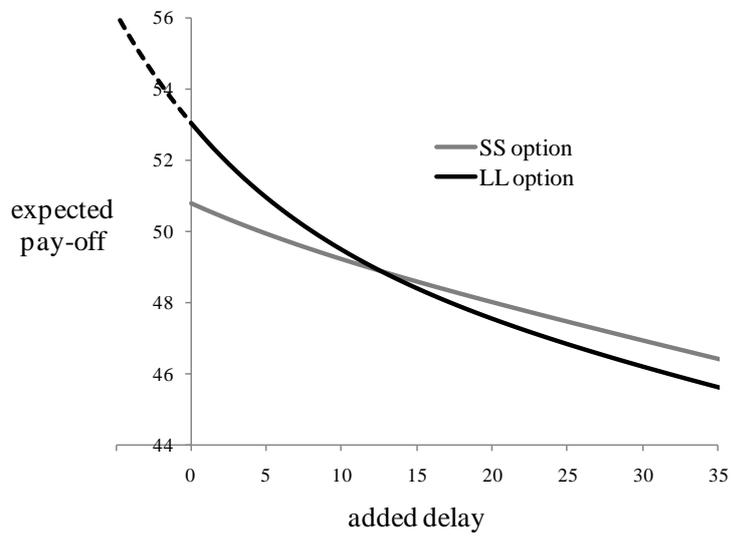
799 **Fig. 3**

800

801 **(a)**

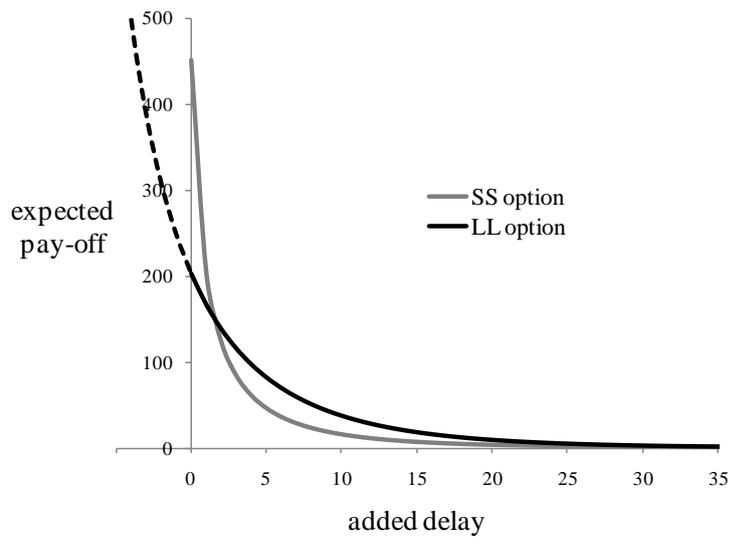
802

803

804 **(b)**

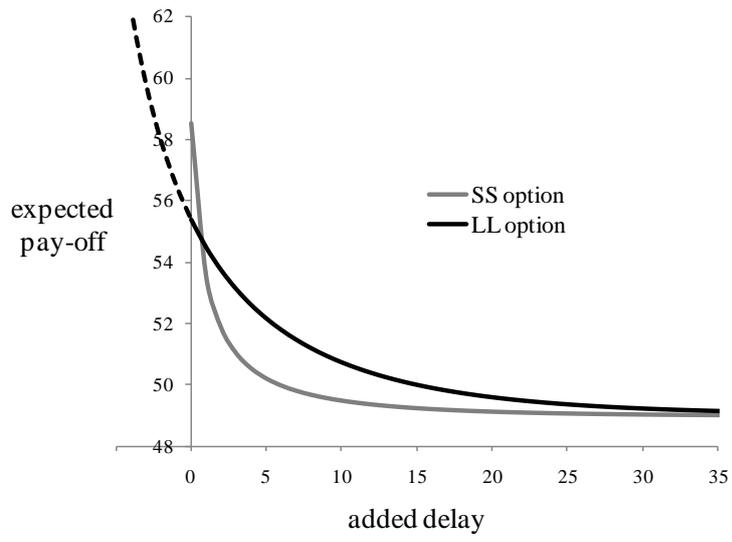
805

806

807 **Fig. 4**

808

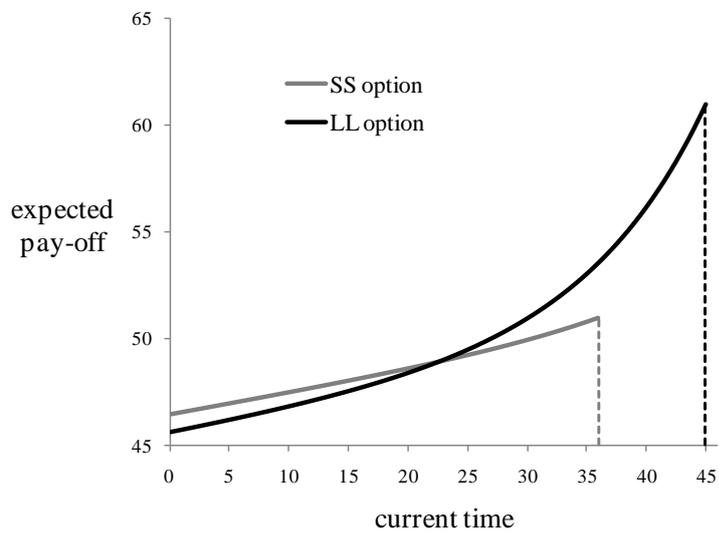
809

810 **Fig. 5**

811

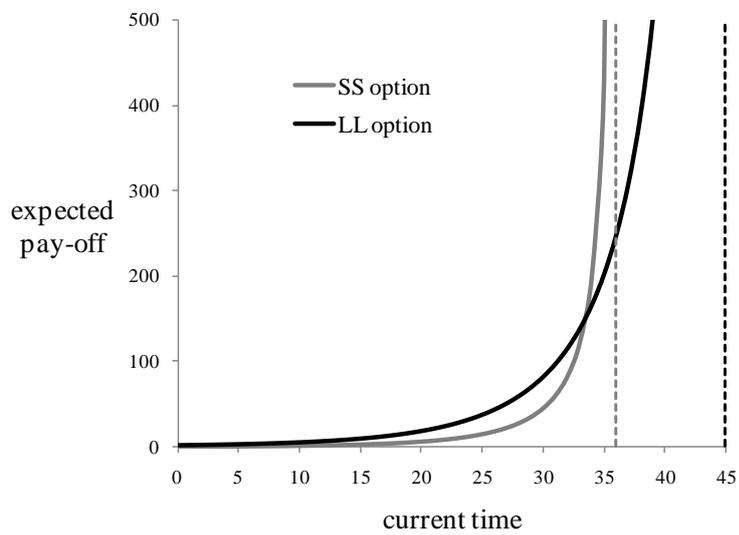
812 **Fig. 6**

813

814 **(a)**

815

816

817 **(b)**

818

819