Intertemporal choice

Daniel Read
London School of Economics and Political Science

Intertemporal Choice

Daniel Read
London School of Economics

Daniel Read
Department of Operational Research
London School of Economics and Political Science
Houghton Street
London, WC2A 2AE

Tel: 020 7955 7617
Email: d.read@lse.ac.uk

Intertemporal choice

'Intertemporal choice' is used to describe any decision that requires trade-offs among outcomes that will have their effects at different times. This definition is broad, as reflected by the range of choices and activities that are embraced by it. These include:

- Whether or not to have a flu shot;
- The choice between fruit salad or tiramisu;
- When to get down to work on a promised paper;
- Whether to invest in a pension plan or buy a widescreen TV; and
- (For a pigeon) one food pellet in one second, or two pellets in two seconds.

In every case, the agent is called on to decide between an earlier and usually smaller penalty or reward (e.g., a flu shot, a TV) and a later and usually larger one (the flu, comfortable retirement). The goal of research into intertemporal choice is to understand how these choices are made, and how they should be made.

This chapter introduces the major issues in intertemporal choice from the perspective of behavioural decision making. The first section describes some basic concepts, including both the terms used and the basic mathematical notation and calculations. This is followed by a discussion of normative issues which highlights the fact that we cannot easily determine whether the weight people put on future outcomes is due to delay discounting, or to rational utility maximisation. Next, I consider the assumptions that underpin (usually implicitly) the interpretation of experimental findings. After these foundations are in place, I survey the major findings – those usually classified as 'anomalies' – in intertemporal choice, and discuss some psychological theories that have been developed to explain them. I conclude the chapter by describing probably the most pressing practical problem in intertemporal choice – the problem of self-control. My primary goal is less to provide a broad survey of what is now a voluminous body of research[1], than to highlight major issues that have sometimes been given too little weight.

The terminology of intertemporal choice

The conventional analysis of intertemporal choice takes the view that the effect of delay on the subjective value of future outcomes can be captured by a *discount function*, which plays the same role in intertemporal choice as the probability weighting function does in risky choice. Just as outcomes should receive less weight the lower their probability of occurrence, so it is usually assumed they should receive less weight the longer they are delayed.

The discount function can be denoted as *F(d)*, where *d* is delay, the time intervening between the present and the time of consumption. According to *discounted utility theory*, the utility obtained from a series of future consumption occasions occurring at regular intervals by an agent with a given discount function is given by[2]:

$$U^t = \sum_{d=0}^{n} F(d)u(c(t+d)),\qquad\qquad(1)$$

Where *F(0) = 1*, *t* is the time of evaluation, *c(t+d)* refers to the resources consumed at time *t+d*, and *u(c(t+d)* refers to the instantaneous utility derived from that consumption. If *F(d)* is a declining function of delay, the decision maker is *impatient*, and the rate at which *F(d)* changes is the *pure rate of time preference*.

The discount function $F(d)$ is often given as a *discount rate (r),* which is the proportional change in value of $F(d)$ over a standard time period (usually one year), or as a *discount factor (δ),* which is the proportion of value that remains after delaying an outcome by that standard period. At a given delay $d \geq 1$ the discount rate is:

$$r(d) = -\frac{F(d) - F(d-1)}{F(d-1)}, \tag{2}$$

and the discount factor is:

$$\delta(d) = \frac{F(d)}{F(d-1)} = 1 - r(d). \tag{3}$$

In Table 1, $F(d)$, $r(d)$ and $\delta(d)$ are illustrated for four widely used discount functions. As will be explained shortly, the exponential function, in which the discount rate is the same for all future time periods, is the economic gold-standard. The other functions in Table 1 are often called 'hyperbolic,' meaning only that delay is in the *denominator* of the discount function so that the impact of an increase in delay of one period, from d to $d+1$, is greater the shorter the original delay d (Ainslie, 1975). Hyperbolic functions have been widely adopted in part because they appear to better reflect how people (and even animals) value future outcomes (Laibson, 1997; Rachlin & Green, 1972).

Table 1: Four frequently discussed discount functions and their corresponding discount rates (r) and discount factors (δ).

| | Exponential | Hyperbolic discounting | | |
		One-parameter (Mazur, 1984)	Generalized (Loewenstein & Prelec, 1992)	Proportional discounting (Harvey, 1994)
F(d)	$\left(\dfrac{1}{1+r}\right)^d$	$\dfrac{1}{1+kd}$	$(1+\alpha d)^{-\beta/\alpha}$	$\dfrac{h}{h+d}$
r(d)	$\dfrac{r}{1+r}$	$\dfrac{k}{1+kd}$	$\dfrac{\beta}{1+\alpha d}$	$\dfrac{1}{h+d}$
δ(d)	$\dfrac{1}{1+r}$	$\dfrac{1+k(d-1)}{1+kd}$	$\left(\dfrac{1+\alpha(d-1)}{1+\alpha d}\right)^{\beta/\alpha}$	$\dfrac{h+d-1}{h+d}$

* In financial terms, exponential discounting is equivalent to *compound* interest, while one-parameter hyperbolic discounting is *simple* interest.

Normative analysis of the discount rate

The modal strategy in behavioral decision research is to compare the predictions of a normative model with observed behavior. Usually, the predictions of the normative model are themselves indeterminate or controversial, but for intertemporal choice there *is* an uncontroversial normative principle, due to Irving Fisher (1930), that dictates the rate at which money should be discounted: Rational decision makers will borrow or lend so that their marginal rate of substitution between present and future money (i.e., the rate at which it can be exchanged while keeping utility constant) will equal the market interest rate[3].

One implication of this is that the pure rate of time preference will be independent of the willingness to trade off monetary amounts available at different times. To see why, I will discuss a '5% market' in which borrowing and lending can be done at 5%, so that $100 invested today will yield $105 in one year. Consider an impatient individual who is indifferent between consuming $100 now or $120 in a year (in the terminology of Eq. 1, r(d) = 0.25). Suppose he is offered a choice between $100 now and $110 in one year, what should he choose? The answer is he should wait for the $110. Then, he should borrow $100 on the capital market and pay back $105 in one year when he receives the $110, thereby earning $5 for his troubles. A similar argument holds for someone who is very patient and has r(d) = 0.025, meaning she is indifferent between consuming $100 now or $102 in one year. If given a choice between £100 now and $104 in a year she should take the $100 and lend it for a year, at which point she will be paid back at $105 including interest. In short, all rational decision makers will make identical choices between delayed money. This does not mean they will make the same *consumption* decisions: The impatient man will probably spend his $100 immediately, while the patient woman will lend hers. But their intertemporal trade-offs for money will not tell us how impatient they are (c.f., Mulligan, 1996).

More generally, discounting can have two bases. The first is that *opportunity costs* grow over time[4], and the positive discount rate compensates for this. The interest lost from delaying the receipt of money or paying it out too soon is one example of an opportunity cost: In the 5% world, for example, waiting one year for £104 instead of taking £100 immediately would cost the decision maker £1 in interest. The second basis is *impatience* or pure time preference, meaning that a given amount of utility is preferred the earlier it arrives. I will consider these two bases in turn.

Firstly, delays are costly in almost every productive activity, and so it is rational to prefer more now rather than later. For the farmer, seed grain (like all capital goods), can be transformed into more grain next year, and so the farmer prefers one bushel now to one next year. Likewise, a taxicab driver waiting for the delivery of his cab is missing fares, and the businessman waiting for a cab is losing quality time at home. In general, most people grow wealthier over time, and therefore a given quantity of almost any resource brings more benefit now than in the future. In addition, delays are desirable for losses, and for much the same reason as they are undesirable for gains. A given amount lost will be worth less in the future (as in the farmer who would rather lose one bushel next year than this year), and time enables us to prepare or prevent future losses, so we should want to defer them as long as possible. Discounting forthese reasons is rational, because delay is the medium in which value, risk and uncertainty grow.

A well-developed example of rational discounting is found in evolutionary theory (e.g., Rogers, 1994; Sozou & Seymour, 2003). Theoretically, intertemporal choices are made to maximize Darwinian fitness, meaning the expected reproductive capacity of organisms and their kin (weighted by the degree of kinship). The organism reaches an economic equilibrium when the marginal rate of substitution between present and future fitness is held constant. To give a simple example, imagine you could produce one child now, or more in ten years, but that the probability that you will be able to implement any future reproduction plan is only 0.5. For instance, you might be unable to find a partner, dead, or infertile. Then you should wait only if by doing so, and if you can implement your plan, you will get more than two offspring. The marginal rate of substitution

between now and then is 0.5, since you will trade one child in the future for half a child now. This evolutionary account rationalizes discounting, but only because a given amount of resource usually contributes more to Darwinian fitness the earlier it is consumed. Therefore, taking an opportunity now rather than later is an example of taking the option with the lowest opportunity cost.

The second kind of discounting, properly denoted as pure time preference or impatience, involves putting more weight on expected utility (u(c) in Equation 1) the earlier it comes. The concept of utility is notoriously difficult to define, but it is best thought of as 'whatever we ultimately care about.' Perhaps it is Darwinian fitness, perhaps it is happiness or feelings of mastery, orperhaps a combination of these and other things besides. Whatever utility is, impatience means that someone who currently expects to experience equal utility at two future times, will want to increase the earlier utility by one unit, in exchange for a decrease in the later utility of more than one unit. The overall consequence is that, under the assumption of diminishing marginal utility to consumption, total lifetime utility will be reduced.

Because impatience will reduce lifetime utility, there have been few attempts to justify it, and it is often met with disapproval (Peart, 2000). It is, indeed, difficult to explain why we should care less about future than current utility, in the sense that you would take less utility now over more later. One who has tried to justify it is Parfit (1984), who has argued for a position earlier proposed (and rejected) by Sidgwick (1874). The argument begins with the very old idea that we can be described as a sequence of 'selves' distributed over time: you will not be (entirely) the same person tomorrow as you are today. Moreover, it is ethically justifiable to care less about the utility of other people than about ourselves. For the same reason, the argument concludes, we can justifiably value the utility of our future selves less than that of our current selves. It is clear that even if we agree with the premises, we need not agree with the conclusion since it requires that we see no relevant differences in our relationship to our future selves and other people[5]. Nonetheless, this argument may be the only way to justify true impatience.

Time consistency

Strotz (1955) proposed an additional normative principle for intertemporal choice: if we make plans for future consumption, we should stick to them unless we have agood reason to do otherwise. This is called *time consistency*, and can be illustrated with a commonplace example. If you currently prefer £1,010 in 13 months to £1,000 in 12 months, then – unless you have an unexpected need for immediate money – in 12 months you should prefer £1,010 in 1 month to £1,000 immediately. This may not occur. Indeed, we see *time inconsistency* everyday: many smokers *plan* to quit, but light up when the urge hits; many dieters *plan* to have an apple for dessert but end up with bread-and-butter pudding; and many academics *plan* to work on the promised chapter but find other things to do when the time comes.

Because exponential discounting is the only way to avoid time inconsistency, it is widely held to be the only rational way to discount. In contrast, the *hyperbolic* discountfunction, which can produce time inconsistency, has been widely adopted as a more-realistic way of describing how people (and animals) value future outcomes. The crucial properties of hyperbolic functions are shown in Figure 1. The choice is between

two alternatives, a smaller-sooner (x_1) and a larger-later (x_2) one. While the larger-later alternative is preferred when both are substantially delayed, when the smaller-sooner alternative becomes imminent it undergoes a rapid increase in value and is briefly preferred. The smaller-sooner reward might be the pleasure from a cigarette, the larger-later reward might be good health: one week in advance, you prefer the prospect of good health, but as time passes the desire for the cigarette grows faster than the desire for good health, until, for what may bea very brief period, the cigarette is preferred. Unfortunately, it is during this brief period that irreversible decisions can usually be made, and a lifetime's resolve can besometimes overcome by a moment's weakness. The language of multiple selves is often used in this context (e.g., Ainslie, 1992; Elster, 1979; Nozick, 1993; Read, 2001; Strotz, 1955). The farsighted decision maker who prefers the larger-later alternative prior to the crossover point is $Self_a$, while the myopic (self-interested) decision maker who wants the smaller-sooner option is $Self_b$[6]. In a later section on self-control I will discuss how $Self_a$ can ensure that $Self_b$ does its bidding.

Figure 1

Hyperbolic discount function, showing how preferences can briefly change from a larger-later (x_2) to a smaller-sooner (x_1) reward, and how a single individual can be divided into 'selves' in relation to the rewards.
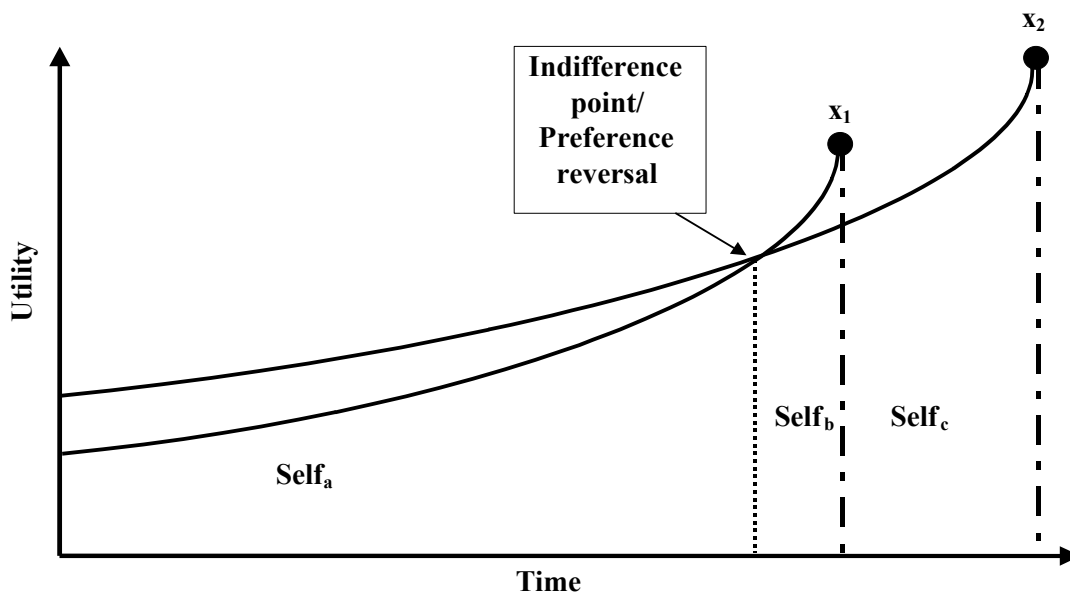
Exponential discounting means discounting at a *constant* rate. This constancy can be divided into a weak and strong form: both forms are delay-independent, but the strong form (called *stationarity*; Koopmans, 1960) is also *date-independent*, while the weak form is *date-dependent*. *Date-independence* means that not only can decision makers make consistent plans, but also that they can advance or delay a sequence without changing their preferences. This means, for instance, that if you prefer the sequence $A=\{20,10,22\}$ to $B=\{20,20,10\}$, where the numbers represent utility in successive time periods, you will also prefer $\hat{A}=\{10,22\}$ to $\hat{B}=\{20,10\}$. That is, dropping a common outcome in one time period and advancing the remainder of the sequence by one period will not change the preference order. I call this *cross-sectional* time consistency, since it means that preferences for different sequencesevaluated at the same time are consistent with one another.

The weaker version of constant time discounting allows date-dependence which means that you might, for example, have different discount rates for the different years of your life (as proposed by Rogers, 1994; Trostel & Taylor, 2001, many others), but that these remain unchanged as you get older. Date-dependent discounting will not necessarily lead to cross-sectional time consistency, but it does entail *longitudinal* time consistency, which means that once you have made a plan, you will not change it with the passage of time. To illustrate with the above sequences, imagine someone who has the following discount factors for three periods: $\delta(1)=0.9$, $\delta(2)=0.7$, $\delta(3)=0.9$. He prefers A to B:

$$u(A: 20,10,22) = (0.9*20)+(0.9*0.7*10)+(0.9*0.7*0.9*22) \cong 37$$
$$u(B: 20,20,10) = (0.9*20)+(0.9*0.7*20)+(0.9*0.7*0.9*10) \cong 36.$$

The passage of one year will not change this preference, because he will still evaluate the second and third period outcomes using $\delta(2)$ and $\delta(3)$:

$$u(\hat{A}:-,10,22) = (0.7*10)+(0.7*0.9*22) \cong 21$$
$$u(\hat{B}:-,20,10) = (0.7*20)+(0.7*0.9*10) \cong 20.$$

This is longitudinal dynamic consistency. On the other hand, if the sequences are advanced by one period and the first period outcome dropped, then $\delta(1)$ and $\delta(2)$ will be used:

$$u(\hat{A}:10,22) = (0.9*10)+(0.9*0.7*22) \cong 20$$
$$u(\hat{B}:20,10) = (0.9*20)+(0.9*0.7*10) \cong 21.$$

Now, B′ is preferred to A′, and cross-sectional time consistency is violated. I will shortly discuss the significance of this distinction for empirical studies of intertemporal choice.

Empirical research into time preference

There are two broad approaches tothe study of intertemporal choice. One is to observe consumption in different periods, and to infer from this consumption what the person's discount rate must be (e.g., Trostel & Taylor, 2001). For example, if someone with £1,000 spends £550 in the first period and £450 in the second, they are showing positive time preference.

In judgment and decision making research, the usual method is to estimate the discount rate r or discount factor δ by obtaining choices or matching judgments. This typically involves finding pairs of outcomes x_1 and x_2 that occur at delays d_1 and d_2 such that:

$$u(x_1) = u(x_2).$$

Then r and δ can be derived in the following way:

$$\delta = \left(\frac{u(x_1)}{u(x_2)} \right)^{\frac{1}{d_2 - d_1}} \tag{4}$$

$$r = \left(\frac{u(x_2)}{u(x_1)} \right)^{\frac{1}{d_2 - d_1}} - 1. \tag{5}$$

In most experiments the outcomes are monetary, although there aremany exceptions, including a growing literature on trading-off health outcomes (reviewed in Chapman, 2003). In these studies, the interpretation of results depends on a set of untested and sometimes implausible assumptions. These are described in the first section below. In the next section, a set of results that appear to be at variance with various normative principles – called anomalies – are summarized.

Assumptions

Most reported studies of intertemporal choice make at least a few, and sometimes all, of the assumptions described in this section.

1. *Outcomes will occur with certainty*.

The term $u(c)$ in Equation 1 represents the expected utility from consumption at a given time. If consumption cannot be assumed to occur with certainty, its value should be discounted for uncertainty as well as its delay. Since the future is always uncertain, with uncertainty increasing as a function (not necessarily linear) of delay, observed discounting should always involve a combination of the effects of risk and delay. Experimental results are usually interpreted on the assumption that participants assume future outcomes will occur with certainty. Yet any expected risk will increase apparent discounting and, moreover, non-linearities in the rate at which risk increases with time could well lead to behavior consistent with hyperbolic or other varieties of non-constant time discounting (Benzion, Rapoport, & Yagil, 1989; Sozou, 1998)

2. *Outcomes are consumed instantly*.

The results of studies are usually interpreted as if c, in the expression $u(c)$ from Equation 1, is identical to whatever is received at that moment. This assumption will rarely be true for any good that can be stored or traded. The consumption of money, for instance, will be spread over time, and may even have been consumed by borrowing *before* it has been received. Moreover, even spending money does not mean that its consumption is achieved at that moment – the utility from the purchase of durable goods extends far beyond the moment of purchase. A similar argument holds for any outcome that need not be consumed at once, or whose consumption has distributed consequences. Indeed, the utility of even the least fungible of outcomes – including specific experiences such as headaches or kisses – can have lasting consequences. Because the utility stream from most outcomes is unknown, this assumption is particularly problematic whenever comparisons are made between discounting in different domains.

3. *Target outcomes are not evaluated in the context of other possibilities*.

This means that participants choose without thinking about their other opportunities. As already discussed, because rational decisions involve minimizing opportunity costs, most observed decisions should reflect market forces rather than pure

time preference. It is only when there is no market (such as with health outcomes) that rational choices will reflect pure time preference. This assumption is a particularly tough nut to crack. If a person's observed δ deviates from what it *should* be based on their market opportunities, this means they are behaving 'irrationally.' But what form does this irrationality take? One suggestion is that if people do not behave as they would if they were economic maximizers, then their choices must reflect their true rate of impatience (Coller & Williams, 1999). But this is a problematic view, since if people are irrational, what justification can there be for assuming that their irrationality takes a specific, albeit analytically convenient, form?

4. *The utility of a good is related to the quantity of that good by a multiplicative constant.*

In most studies, δ or r is computed by substituting x_1 and x_2 for $u(x_1)$ and $u(x_2)$ in Equations 4 and 5, but this may not be valid. This is best illustrated with an example: Suppose I observe that you are indifferent between $100 today and $121 in one year, and I am willing to make Assumptions 1, 2 and 3. I still cannot conclude that this represents your pure rate of time preference because to do so I must still assume that $u(\$100)/u(\$121) = 100/121 = \delta$. If we take the usual view that marginal utility (or *value*) is a decreasing function of amount, treating utility as linear in amount will generally lead to overestimates of the true discount rate. For instance, if $u(c) = c^{1/2}$, then for this example δ would be 10/11, sothat if we make the linearity assumptionwe would overestimate discounting. In general, if we are not sure about the utility function, then any given value of δ for money is compatible with any value of δ for utility.

5. *The utility from outcomes is timing-independent.*

Just as ice is more valuable in the summer than the winter, so it is generally true that the utility from a given kind of consumption will depend on its timing. Therefore, if people make intertemporal trade-offs between outcomes that differ in value at different times, these differences will influence measured discount rates. If the good being evaluated has less per-unit value earlier rather than later, then we will underestimate discount rates, and if it is less valuable later we will overestimate them. To give an example undoubtedly relevant to research practice, college and university students perceive themselves as poor when in school, yet expect to be well off in the near future, after they graduate. Any moderate time-horizon is likely to involve both sides of this divide. These students will *appear* more impatient than they really are, because a given amount of money means more to them now than later.

Although these five assumptions are rarely discussed in the context of empirical reports, data are frequently interpreted in a way that is only warranted if they are all true. There have been few attempts to determine the role of uncertainty in time discounting, how people expect to consume what they receive, whether they think of opportunity costs, or how much the subjective utility of goods influences discount rates. I suggest that we, as researchers, should focus more attention on these assumptions. We should make sure they are justified and, if not, attempt to vitiate their effects.

Anomalies

The goal of much intertemporal choice research has been to test the economic model, and to demonstrate how it fails. Usually, the economic principle tested is not the first (least-controversial) one which states that money should be discounted at the prevailing market rate[7], although it is well known that people do not do this (Benzion et al., 1989; Frederick et al., 2002, Thaler, 1981). The focus has rather been on whether each person can be characterized by a stationary discount rate, regardless of what that rate is. The consensus is that people do not apply a single rate to all decisions. Rather, their implied discount rate is highly domain dependent (Chapman & Elstein, 1995; Madden, Bickel, & Jacobs, 1999), and even within domains it depends on the choice context. Consequently several 'anomalies,' or systematic deviations from constant discounting, have been identified. These are described and briefly discussed in this section.

Time inconsistency. The concept of time inconsistency was introduced earlier. *Cross-sectional* time inconsistency, involving variations of delay within a single session, has been demonstrated in several domains (Green, Fristoe, & Myerson, 1994; Kirby & Herrnstein, 1995). Kirby and Herrnstein (1995) offered people a choice between a small prize immediately and a large prize later. They first established that the smaller-sooner prize was preferred to the larger-later one, and then delayed both prizes by a constant amount. Participants almost invariably changed to the larger, later prize, usually after a very small increase in delay.

Longitudinal time inconsistency, or impulsivity, which means that the preference for smaller-sooner over larger-later options increases as they become closer in time, is the prototypical phenomenon of intertemporal choice. Yet there have been no studies showing results like those implied by Figure 1 – in which preferences change in the direction of a smaller-sooner x_1 after a period of favoring larger-later x_2. Rather, the few demonstrations of longitudinal time inconsistency have shown that when there are two non-monetary outcomes available at the same time, but one is more immediately beneficial than the other, preference will often switch in the direction of the more tempting alternative at the moment of consumption (Read & Van Leeuwen, 1998). This pattern *is* often explained using hyperbolic discounting along with assumptions about how the utility from different options is distributed over time (see the account in Read, 2003), but this may not be the best explanation (as we will discuss below). It remains a puzzle why such a widely discussed phenomenon as longitudinal dynamic inconsistency has not been properly tested.

Delay effect. If we elicit the present-value of a delayed outcome, or the future-value of an immediate outcome, then the obtained value of δ will be larger (and r will be smaller) the longer the delay. I will use the notation $p(x,d_i)$ to denote the present value of the outcome (x,d_i), and $f(x,d_i)$ to denote the future value of the outcome $(x,0)$. Then, if $d_2>d_1$:

$$\left[\frac{p(x,d_1)}{x}\right]^{\frac{1}{d_1}} < \left[\frac{p(x,d_2)}{x}\right]^{\frac{1}{d_2}}, \ and$$

$$\left[\frac{x}{f(x,d_1)}\right]^{\frac{1}{d_1}} < \left[\frac{x}{f(x,d_2)}\right]^{\frac{1}{d_2}}$$

To illustrate, Thaler (1981) obtained the future value of amounts of money available now, if they were delayed by times varying from 3 months to 10 years. If the amount now was $15, the discount rate varied from 277% for a three-month delay to 19% for a ten-year delay. Similar results have been reported in dozens of subsequent studies (e.g., Benzion et al., 1989; Green & Myerson, 1996, Raineri & Rachlin, 1993; Kirby, 1997). As in the example, most demonstrations of the delay effect concern money, although it has been shown for other domains, including health and holidays (Chapman et al., 2001; Chapman & Elstein, 1995). The delay effect is generally attributed to some form of hyperbolic discounting, although it could equally be due to the interval effect.

Interval effect. Thedifference between the delays to two outcomes is the interval between them (d_j-d_i). . Discounting depends strongly on the length of this interval, with longer intervals leading to smaller values of r or larger values of δ. To illustrate, consider two intervals of equal length: one from d_1 to d_2; the other from d_2 to d_3. A decision maker equates pairs of outcomes available at the beginning and end of each interval:
$u(x_1) = u(x_2)$,
$u(x_2) = u(x_3)$.
He also equates the outcome available at the beginning of the first interval with one at the end of the second interval:
$u(x_1) = u(\hat{x}_3)$.
The interval effect means that, in general, $x_3 > \hat{x}_3$. Or, in other words, shorter intervals lead to more discounting per-time-unit (Read, 2001, Read & Roelofsma, 2003).

The interval effect provides an alternative account for the delay effect. If d_1=0, meaning x_1 is received immediately, then the delay and the interval are confounded, and the two effects can be confused. This was suggested by Read (2001b), who reported equal discount rates for intervals of equal size, regardless of when they occurred (that is, $x_1/x_2 = x_2/x_3$). More recent work (Read & Roelofsma, 2003), however, suggests that there is both a delay and interval effect, although the interval plays the largest role. A systematic analysis of the relative contributions of delay and interval to discounting is yet to be done.

Magnitude effect. This means that the discount rate is higher for smaller amounts. Imagine a small (x_1^S) and large (x_1^L) outcome. If these are equated with outcomes available at different times, such as:
$u(x_1^S) = u(x_2^S)$
$u(x_1^L) = u(x_2^L)$'

then $x_1^L / x_2^L > x_1^S / x_2^S$. The magnitude effect is the most robust of the 'classic' anomalies, and is obtained whenever discounting is measured for differing amounts under otherwise identical conditions. The effect is greatest for smaller amounts and short delays: Kirby (1995) obtained strong magnitude effects when comparing discounting of $10 and $20 over 30 days; Shelley (1993) observed magnitude effects for values between $40 and $200, and almost no effect for larger amounts; Green, Myerson and McFadden (1997) found negligible effects for amounts above $200. The magnitude effect has also been generalized to other domains (Chapman, 1996b) and, moreover, Chapman and Winquist (1998) has shown that something similar happens in other contexts – people give smaller (proportional) tips the larger the restaurant bill.

Direction effect. The discount rate obtained by increasing the delay to an outcome is greater than that from reducing the delay. This is demonstrated by first endowing the decision maker with an outcome (x_1) or (x_2), and then asking them to specify either the *delayed* x_2^D or the *expedited* x_1^E that would make them indifferent between staying with the outcome they have or changing to the new one:

$$u(x_1) = u(x_2^D)$$
$$u(x_1^E) = u(x_2)$$

The result is that $x_1 / x_2^D < x_1^E / x_2$, indicating discounting is greater for delaying than expediting. This asymmetry was first reported by Loewenstein (1988). In one study, he gave participants a gift certificate worth $7 for a local record shop, available in either 1, 4 or 8 weeks. They then chose between the gift certificate they had, and differently-valued gift certificates available earlier or later. For example, someone who received an 8-week gift certificate could choose between that and (say) a $6 certificate in 4 weeks. The value put on a given delay was much greater when the certificate was delayed than when it was expedited. Loewenstein (1988) interpreted this using prospect theory. Expediting or delaying an outcome means losing something at one time and gaining it at another. Because of loss aversion, the substitute outcome must be disproportionately large to compensate for the loss. For delay, the substitute outcome (x_2^D) is the later amount and so this effect increases the discount rate, while for expediting the substitute (x_2^E) comes earlier, thus decreasing the discount rate.

Sign effect. The discount rate is lower for losses than gains. That is, imagine that negative (x_i^-) or positive (x_i^+) outcomes are equated with either earlier or later outcomes, so that:

$$u(x_1^+) = u(x_2^+)$$
$$u(x_1^-) = u(x_2^-)$$

The sign effect means that $x_1^+ / x_2^+ > x_1^- / x_2^-$. For single outcomes, this has proved a relatively robust effect (Antonides & Wunderink, 2001; Thaler, 1981). Indeed, people often show zero or even negative discount rates for losses – in many studies people will want to take even monetary losses immediately rather than defer them (Yates & Watts,

1975) and are very likely to want to experience bad health outcomes immediately (Van der Pol & Cairns, 2000).

The sign effect is not, however, uncontroversial. Shelley (1994) found more discounting for losses and gains, as well as an interaction between the sign effect and the direction effect: The discount rate from delaying a loss is greater than that from delaying a gain, while that from expediting a gain is greater than that from expediting a loss (Shelley, 1993). Moreover, the tendency to discount losses less contradicts studies of approach-avoidance conflicts, which found that the avoidance gradient is steeper than the approach gradient – implying that losses will be discounted more rapidly than gains (Miller, 1959). This idea was developed into a theory of intertemporal choice by Mowen and Mowen (1991), but evidence of a sign effect appeared to contradict it. More recently, however, studies of the discounting of mixed (positive and negative) outcomes suggest that their negative parts receive more weight with delay. For example, Soman (1998) showed that when people evaluate rebates, which have a positive component (amount to be received) and a negative one (effort required to redeem them), the negative component receives much less weight the longer the outcome is delayed.

Sequence effects. A sequence is a set of dated outcomes all of which are expected to occur, such as one's salary or mortgage payments. In a wide variety of choice contexts, people usually (although not always) prefer constant or increasing sequences to decreasing ones, even when the total amount in the sequence is held constant[8]. To illustrate, given a choice between three ways to distribute $300 over three months — increasing = {90, 100, 110}, constant = {100,100,100}, or decreasing = {110,100,90} — most will choose constant over increasing, and increasing over decreasing (Barsky, Juster, Kimball, & Shapiro, 1997; Chapman, 1996a; Gigliotti & Sopher, 1997; Loewenstein & Prelec, 1993;Loewenstein & Sicherman, 1991; Read & Powell, 2002). Such a preference is puzzling for two reasons. First, it shows poor economic judgment because under any positive interest rate a decreasing monetary sequence has a higher present value than an increasing one[9]. Indeed, just as the personal level of impatience should have no relationship to observed discounting, neither should one's preference for monetary sequences (Deaton, 1987). Second, it is difficult to reconcile with other studies of time preference, since anyone with a positive discount rate will prefer a decreasing income distribution.

Theories of time discounting

There are two main theoretical approaches to the anomalies in intertemporal choice. The first is to *rationalize* the anomalies by showing that they are what we would expect rational actors to do in the environment in which they find themselves. Such explanations are usually behavioral or economic, based on finding a fit between actions and choices. Many cases of such rationalizing were described above: a positive event in the future might be valued less than it is now because it is more uncertain, because it will lead to less evolutionary fitness, or because we will be richer in the future and so want it less than we do now. The second approach is to *mechanize* phenomena by describing processes that will produce the anomalies, without committing ourselves to whether the output of that process is rational or not. The current section offers a brief review of some proposed mechanisms. As will be seen, no theory accounts for all observed phenomena,

and the theories are not necessarily even rivals. Intertemporal choice is a complex
phenomenon, and probably determined by many mechanisms.

Value function approaches.
 The earlier discussion of Assumption 4 showed that the discount rate could be over or
under-estimated by treating outcome magnitude as proportional to outcome utility. In
fact, some intertemporal-choice anomalies can be a non-linear utilityor value function
with special properties.
        One such approach was reported by Loewenstein and Prelec (1992) who showed
that many anomalies could be attributed to a value function, like the one in Figure 2,
having three properties: Steeper for losses than gains; more elastic for losses than gains;
and more elastic the larger the absolute value of the option[10]. This function will produce
magnitude, sign and direction effects. For example, the sign effect occurs because the
proportional change in value from -$5 to -$10 is smaller than the change from $5 to $10
(more elastic for losses than gains), and the magnitude effect occurs because the
proportional change in value from $50 to $100 is greater than that from $5 to $10 (more
elastic the larger the amount).

Figure 2.
Value function having the three properties described by Loewenstein &
Prelec (1992) to account for magnitude, sign and direction effects.

Loewenstein and Prelec value function
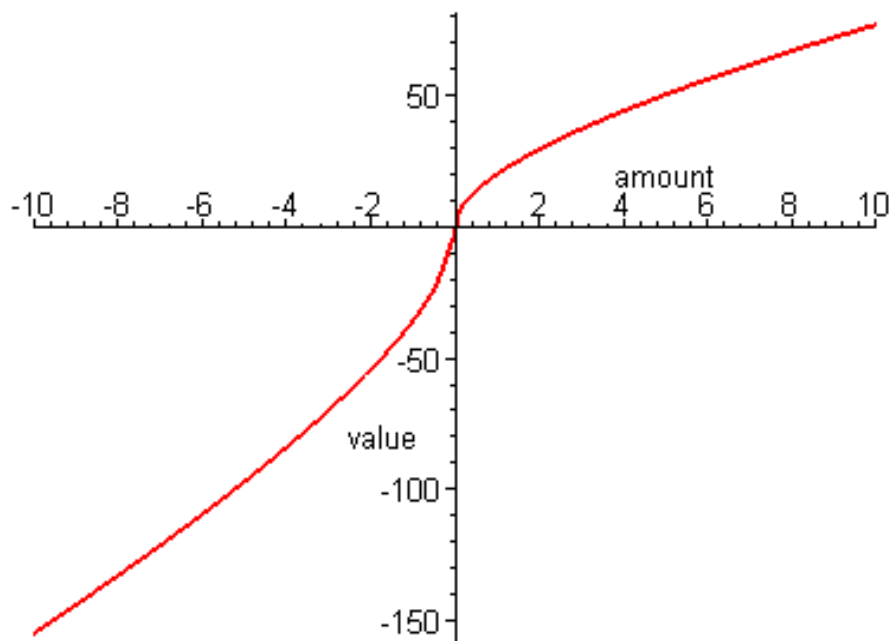
        The value function cannot be the only explanation for intertemporal choice
anomalies. One reason is that it cannot account for time-specific effects, such as the
delay effect and time inconsistency, as shown by the fact that Loewenstein and Prelec's
(1992) complete model also includes the generalized hyperbolic discount function from

Table 1. A more general issue is that the value-function approach assumes alternative-based choice, which is incompatible with many intertemporal choice phenomena.

Attribute-comparison models.
 In *alternative-based choice* each option is first evaluated, as in Equation 1, by summing up the discounted utility of each unit of consumption, and then the option with the highest utility is chosen. In *attribute-based choice*, on the other hand, options are compared attribute-by-attribute, and choice is based on some transformation of attribute differences. Such *attribute-based* choice leads to anomalous intertemporal choices whenever the decision weight put on delay or outcome magnitude depends on the choice context. This can occur when the weight assigned to attribute values depends on the magnitudes being compared (as with subadditive discounting, Read, 2001b), or when some differences are non-compensatory, as in lexicographic or threshold-based choice. Many studies have shown that attribute-based models are both sufficient to account for many familiar anomalies, and necessary to explain some patterns of intertemporal choices.

Intransitive choice cannot come from alternative-based choice processes, so evidence for systematic intransitivity is also evidence for attribute-based choice. Roelofsma and Read (2000) showed that intransitive intertemporal choice is easy to produce, and attributed this to the use of a lexicographic-like procedure with interval evaluated first and amount evaluated second. According to their model, and within the range of choices they studied, when the interval is large enough then the smaller-sooner option is taken, but otherwise the larger-later one is. The interval effect, described above, also reveals alternative-based decision processes – because the interval is a difference between delays, the fact that intervals and not delays determine choice means that the delays are being compared *prior* to choice (Read, 2001b).

Rubinstein (2000) and Leland (2002) have both argued that many classic anomaliesarise from an attribute-based process based on what they call *similarity* relations between option attributes. According to Rubinstein, when the levels for an attribute are very similar, that attribute is given little or no weight in choice. So, for instance, 101 and 111 days are very similar, while 1 and 11 days are not. This can explain the delay effect and time inconsistency, because a given interval is given little weight when the delays are long, but a lot of weight when the delays are short. Rubinstein showed that when the similarity and hyperbolic discounting hypotheses conflict, the similarity hypothesis wins out. One of his many examples used the following pair of choices:

    A. In 60 days you expect delivery of a stereo costing $960. You must pay on
        receipt. Will you accept a delivery delay of 1 day for a discount of $2?
    B. Tomorrow you expect delivery of a stereo costing $1,080. You must pay on
        receipt. Will you accept a delivery delay of 60 days for a discount of $120?
Hyperbolic discounting predicts that nobody who accepts A will refuse B[11], yet more people agree to B than to A. Rubinstein argues that in Option A the two payments are very similar ($958; $960), and so they receive almost no decision weight. On the other hand, the two payments of Option B are very dissimilar ($960; $1,080) and so receive much more weight.

Attribute-based models account for delay and interval effects in an intuitively appealing manner, and also predict new phenomena such as intransitivity and Rubinstein's anti-hyperbolic effects. Their shortcoming, however, is that they are primarily applicable to rudimentary choices involving delayed amounts of money – there is no clear way, for instance, to use the theory of similarity to explain intertemporal choices between a sports car and a sedan, or between smoking and a stick of gum. This is the goal of the richer psychological theories reported next.

Cognitive/representation theories. These theories explain intertemporal choice in terms of how we represent future outcomes as a function of delay. Böhm-Bawerk suggested that one reason for discounting was that 'we limn a more or less incomplete picture of our future wants and especially of the remote distant ones' (cited in Loewenstein, 1992, p. 14). This general approach – viewing the problem of discounting as being due to how we represent and think about future outcomes – has wide currency. One model has been proposed by Becker and Mulligan (1997), who argue that the discount rate is a function of the resources invested in imagining the future. In their model, decision makers maximize lifetime utility subject to difficulties in envisioning exactly how rewarding the future will be. Hence, they will expend resources to make their image of the future vivid and clear. To cite one example, we might spend time with our parents to remind ourselves of what our needs will be when we are their age.

One influential cognitive approach isTrope and Liberman's (2000, 2003)(2000, 2003) *temporal construal theory*. They begin by asserting that option attributes vary in their centrality, with 'high-level' attributes being more central to the option than 'low-level' ones. The definition of high- and low-level is broad, but typically low-level attributes are more concrete and mundane than high-level ones. Trope and Liberman argue that as options are delayed, their representation becomes increasingly dominated byhigh-level attributes, and so their present-value also becomes more influenced by those attributes. For example, the act of 'writing a book chapter' has high-level attributes might be 'obtaining a publication' or 'informing the field,' while its low-level attributes include 'several difficult days of writing.' When the decision to write it is made (and the deadline is remote) the low-level attributes are much less important than when the deadline is near. This can lead to time inconsistencybetween options differing in the value of their high- and low-level attributes. An option that has relatively unattractivelow level attributes but relatively attractive high-level ones can go from being desirable when still delayed (and the high-level attributes are most important), to undesirable when available immediately (when the low-level attributes dominate). This can explain Soman's (1998) rebate study, mentioned earlier. He showed that a mail-in rebate is a more effective selling strategy if both the effort from applying for the rebate and the payoff aredelayed than if they both occur immediately. To put this in the terms of temporal construal theory, the high-level attribute is the reward amount, while the low-level attribute is the effortinvolved, and the importance of the high-level attribute relative to the low-level one is greater when both are delayed than when they are imminent.

Emotion-based theories. Becker and Mulligan suggest that imagining the future more vividly can decrease our discount rate for that future. This depends, however,on what is being imagined. Imagining something can increase our desire for it along with

our impatience (Mischel, Ayduk & Mendoza-Denton, 2003). For example, if we focus our attention on the wonderful meal waiting for us when we get home we might become hungrier, and therefore   more likely to snack between meals. Our desire for food, triggered by the vivid image, becomes a strong impetus for action.

This idea forms the basis for a major re-examination of time preference by Loewenstein (1996)(1996). He argues that the temporal and *physical* proximity of options that can reduce aversive arousal states (hunger, sexual arousal,withdrawal symptoms) leads to a disproportionate but transient increase in the attractiveness of those options. Indeed, they can even create a feeling of deprivation– we get hungry when we see a steak, and sexually aroused when we see a potential partner.

While visceral factors can explain impulsive choices, the arousal that is produced is not caused by delay, but rather by the presence of the aggravating stimulus. If you cannot see or smell food the temporal proximity of possible consumption might not lead to impulsive preferences; while if you can see and smell it, even if you cannot consume it, the strong and overwhelming desire can still arise: The bacon in your own fridge (which you can eat right now) has much less effect than the smell of bacon from your neighbour's skillet (which you cannot). Usually the presence of the visceral good is confounded with delay-to-consumption, but when visceral arousal is unaccompanied by temporal proximity it nonetheless has the expected effect. To illustrate,Read and Van Leeuwen (1998) reported that people who chose a snackthey would not get for a week were more likely to choose an unhealthy snack if they were hungry when choosing than if they were not (c.f., Loewenstein, Nagin & Paternoster, 1997).

Although impulsivity from visceral states cannot account for all findings in intertemporal choice, it might account for most of those that interest us most – the weakness of resolvein the face of temptation. Indeed, the ultimate goal of the study of intertemporal choice is probably to help us overcome this weakness of will, the lack of which contributes to a vast number of major social problems, including under-saving, obesity, and addictions of all sorts.

Impulsivity and self-control

As we have discussed above, discounting can be due to rational devaluation of the future or to pure impatience. Although pure impatience leads to a reduction in total lifetime utility, by itself it does not lead to impulsivity. Impulsivity occurs when we become more impatient when opportunities are available immediately than when they are delayed. We often believe it is bad to give in to our impulses, and so strive to overcome them. To use the terminology of Figure 1, we say that our $Self_a$ attempts to inhibit the bad impulses of $Self_b$. We call this *self-control*.

Much has been written about the problem of self-control. A long philosophical tradition, going back at least to early Greek philosophers, concerns the problem of what the concept of *self*-control means (Price, 1995). It is from this tradition that the concept of multiple-selves and the conflict between them originates. Behavioral scientists have focused on two closely related aspects of self-control: its strategies and tactics, and the concept of willpower. Both concepts are central to Thaler and Shefrin's (1981) economic theory of self-control, which divides the self into a planner (roughly, $Self_a$) and a doer ($Self_b$), with the doer deploying resources to get the planner to stay in line.

Successful self-control must obey two principles. First, a (pre)commitment must come from Self$_a$, because Self$_b$ prefers to take the wicked thing, and will take it unless constrained. Second, the commitment must either prevent Self$_b$ from acting on his or her bad impulses, or prevent Self$_b$ from coming into existence at all. To use the imagery of Figure 1, self-control means either choosing irrevocably before the crossover point, or else shifting the relative positions of the x_1 and x_2 lines so that x_2 is always higher.

Self-control is well-illustrated by Mischel's (1974, Mischel, Ayduk & Mendoza-Denton, 2003, Sethi, 2000) studies of its development. In the *delay of gratification* paradigm, children are shown two rewards, one better than the other, such as two marshmallows versus one. They can get the larger reward if they wait, but that any time before the waiting period is over they can ring a bell to get the smaller reward immediately. Very young children won't wait, but as they get older they learn to resist temptation either by trying not to think about the reward, by looking away (when it is present), or by otherwise distracting themselves. As they get older, they develop ways of thinking about the tempting reward in an abstract way, such as imagining a marshmallow as a cloud. That is, they attempt to reduce the visceral arousal evoked by the marshmallow. These tactics all involve Self$_a$ (who can resist the marshmallow) trying to reduce the duration of Self$_b$'s existence to zero by reducing the desire for the smaller-sooner alternative.

One important difference between different kinds of self-control tactics is how difficult it is to implement them. Children trying to resist eating a single marshmallow placed in front of them have a difficult time; it would be much easier if the marshmallow were not there in the first place. Ideal self-control tactics, often illustrated with the example of Odysseus having himself bound to the mast, give us little opportunity to decide whether to control ourselves or not. When we have more discretion we have to *resist* temptation, and the mechanical metaphor implies that some energy source is required to do the resisting. This energy source is often called willpower. Shefrin & Thaler (1988), for instance, describe willpower as the 'real psychic costs of resisting temptation,' and model it as a limited resource that shows diminishing marginal returns to investment.

Much the same idea is found in the work of Baumeister & Vohs (2003), who have actually likened willpower to a muscle, meaning that it has limited strength, gets tired with use, and gets stronger with repeated use. They have shown that people do indeed become increasingly less willing or able to exert cognitive effort after some time spent controlling themselves. In a series of mildly sadistic experiments, participants were placed in front of a bowl of freshly baked chocolate chip cookies and enjoined not to eat them. After they left the cookie filled room they attempted to solve unsolvable anagrams. These 'tempted' subjects persevered for much less time than did an untempted control group.

Conclusion

This paper is not a comprehensive account of all research into intertemporal choice. The field is now so large, and growing so rapidly, that it will undoubtedly soon warrant a handbook of its own. To give an idea of how true this is, consider the following *important* topics that are largely missing from this chapter: the physiology of intertemporal choice (Manuck et al., 2003); application of non-constant discounting models in economics (Laibson, 1998); formal models of self-control (Benabou & Pycia, 2002; O'Donoghue & Rabin, 2000, 2001); individual differences in discounting (Green, Myerson, & Ostaszewski, 1999; Trostel & Taylor, 2001); pathologies of discounting such as addiction and criminality (Bickel & Marsch, 2001, Wilson & Herrnstein, 1985); intergenerational discounting (Chapman, 2001; Frederick, 2003; Schelling, 1995); discounting of health outcomes (Bleichrodt & Gafni, 1996, Chapman, 2003); and applications to marketing (Wertenbroch, 2003) and public policy (Caplin, 2003). The list could continue. My goal has been to provide enough background to permit readers to engage in a critical exploration of this new field, as well as to prepare them to make contributions of their own.

Appendix: A note on hyperbolic discounting in animals.

Much theoretical discussion of intertemporal choice has been based on studies done with non-human animals, largely because both humans and animals display behaviors that can be described using hyperbolic discounting (see, Ainslie, 1992; Ainslie & Herrnstein, 1981; Rachlin, 1989). Indeed, the most-cited version of hyperbolic discounting (Mazur's one-parameter model from Table 1) was formulated to summarize the results of studies with pigeons. The purpose of this appendix is to suggest that we should be cautious in making the connection between what we observe in humans and animals: the parallels are more likely to be analogous, rather than homologous. While humans may well routinely sacrifice future pleasure for current pleasure, animals probably do something quite different (Kacelnik, 2003). In this appendix I will focus on only one issue – does hyperbolic discounting in animals reflect the application of an irrational decision rule?

A typical demonstration of hyperbolic discounting in animals is illustrated in Figure 3. A pigeon chooses to press one of two keys. Pressing one key leads to a smaller-sooner (SS available at D_{SS}) reward, the other leads to a larger-later reward (LL at D_{LL}). The reward is typically a period of access to food (2s and 4s in the figure, R_{SS} and R_{LL}), following which there is a post-reward waiting period during which the pigeon cannot make any further choices (W_{SS} and W_{LL}). The total time (T) from choice to choice is the same in both conditions: $D_{SS}+R_{SS}+W_{SS} = D_{LL}+R_{LL}+W_{LL} = T$. The pigeon will often choose SS, and the typical pigeon can be modeled with the Mazur discount function from Table 1. A choice of an earlier reward is irrational because it receives an average return of SS/T rather than LL/T.

Figure 3
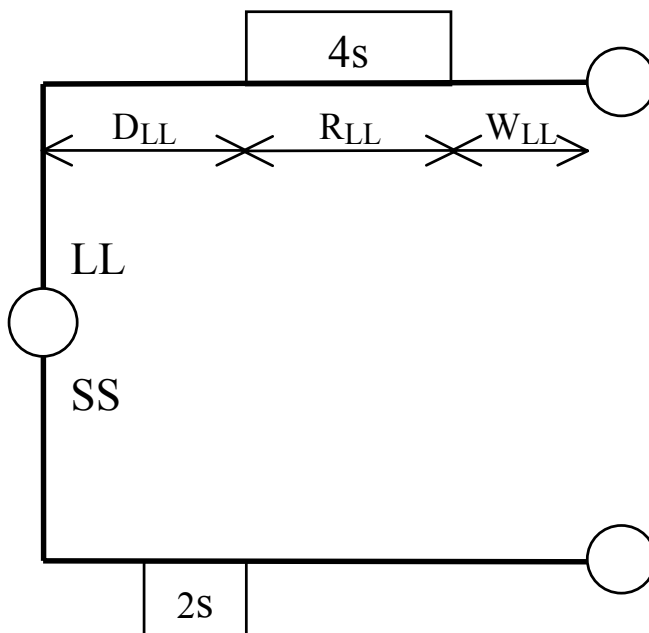
The pigeon, however, is actually operating in a most unusual environment. Since T is constant, the post-reward period W is *negatively* correlated with the size of the

reward and the pre-reward delay D. The sooner it receives its reward, the longer it has to wait for the next one. It is difficult to imagine any natural environment that has such a reward structure. It would mean for instance, that a tiger who chooses to hunt rabbits gets the first one very fast, but then has to wait a long time for the second; while a tiger hunting antelopes will get the first one after a long wait, but will then get the second one almost immediately. It is worthwhile considering what the optimal decision procedure would be in a more realistic environment.

First, let us consider the standard laboratory task as represented by the following time lines:

```
SS:    _____↓__2_____↓__2_____↓__2_____↓__2
LL:    _↓_____4_↓_____4_↓_____4_↓_____4
```
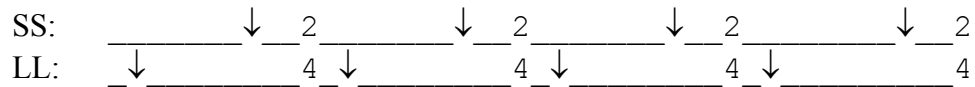
The numbers represent reward magnitudes, and the arrows represent the moment of choice. The SS sequence represents a world in which the animal has to wait longer after receiving a small reward before it can go hunting again. In this world, the animal choosing SS gets 1/3 of the return of the animal choosing LL.

Consider, however, a different time line that represents more realistic reward/choice sequences. These are identical to those above, except that the post-reward interval is removed:

```
SS:   ↓__2↓__2↓__2↓__2↓__2↓__2↓__2↓__2↓__2
LL:   ↓_____4↓_____4↓_____4↓_____4
```
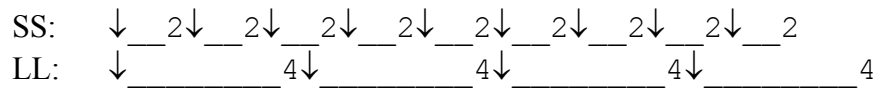
To put this another way, in this environment the animal gets to start searching for food as soon as it has consumed the last one, and the amount of time it takes to get the next animal is correlated with the time it took to catch the last one. In this very plausible world, the animal choosing SS will get about 10% more per-unit-time than the animal choosing LL.

Suppose the animal is trying to optimise its total food consumption, and believes that it is facing the realistic Environment 2 above? Then it will be indifferent between options when $SS/(D_{SS}+R_{SS}) = LL/(D_{LL}+R_{LL})$. Because the variable time interval is in the denominator this implies hyperbolic discounting. To see this, standardize the time intervals by making $D_{SS}+R_{SS}=1$ and $D_{LL}+R_{LL}=T_{LL}$: the resultant discount function is given by $SS = LL/T_{LL}$, which is, in fact, the earliest variant of hyperbolic discounting ever proposed (Ainslie, 1975). Thus, the classic hyperbolic discount function, rightly considered to be irrational in single choice situations, would be rational in the real world. We should not be surprised at this result. It is rare to find that animals are irrational in their natural environment, and apparently irrational behaviour is almost always understandable as an adaptation to their environment (Dawkins, 1982).

The question remains whether human discounting is attributable to an overgeneralization of the animal's sensible decision rule. The analogy between human and animal choice cannot warrant this conclusion, but it is possible that similar processes do operate in humans. I once suggested, perhaps tongue-in-cheek, that hyperbolic discounting might be learned in the following way:

Consider how children are taught the virtue of patience. A child wants a reward (perhaps an ice cream) right now, but his parents prefer that he wait until tomorrow. Whatever delay is agreed, the child cannot open negotiations for a second ice cream until after the consumption of the first. If the ice cream comes tomorrow, that means that there is an entire day in which no ice cream negotiations can occur. Moreover, the agreed delay sets a precedent – the child knows very well that if he has to wait until tomorrow for the first ice cream, he will usually have to wait until tomorrow for all subsequent ice creams… Such a child would rationally adopt some version of [hyperbolic] discounting when dealing with his parents. He will kick up a real fuss to prevent even a few moments delay from the present, but will be relatively blasé about substantial additions to already long delays. (Read, 2000, p. 111)

References

Ainslie, G. (1975). 'Specious reward: A behavioral theory of impulsiveness and impulse control.' *Psychological Bulletin, 82,* 463-469.

Ainslie, G. (1992). *Picoeconomics: The Strategic Interaction of Successive Motivational States within the Person.* New York: Cambridge University Press.

Ainslie, G. & Herrnstein, R. J. (1981). Preference reversal and delayed reinforcement. *Animal Learning and Behavior*, 9, 476-482.

Antonides, G., & Wunderink, S. R. (2001). Subjective time preference and willingness to pay for an energy-saving durable good. *Zeitschrift Fur Sozialpsychologie, 32*(3), 133-141.

Barsky, R. B., Juster, F. T., Kimball, M. S., & Shapiro, M. D. (1997). Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *Quarterly Journal of Economics, 112*(2), 537-579.

Becker, G. S., & Mulligan, C. B. (1997). The endogenous determination of time preference. *Quarterly Journal of Economics, 112*(3), 729-758.

Benabou, R., & Pycia, M. (2002). Dynamic inconsistency and self-control: a planner-doer interpretation. *Economics Letters, 77*(3), 419-424.

Benzion, U., Rapoport, A., & Yagil, J. (1989). Discount Rates Inferred from Decisions - an Experimental Study. *Management Science, 35*(3), 270-284.

Bickel, W. K., & Marsch, L. A. (2001). Toward a behavioral economic understanding of drug dependence: Delay discounting processes. *Addiction, 96*(1), 73-86.

Bleichrodt, H., & Gafni, A. (1996). Time preference, the discounted utility model and health. *Journal of Health Economics, 15*(1), 49-66.

Caplin, A. (2003). Fear as a policy instrument. In Loewenstein, G., Read, D., & Baumeister, R. F. (Eds). *Time and Decision: Economic and psychological perspectives on intertemporal choice.* New York: Russell Sage Foundation.

Chapman, G. B. (1996a). Expectations and preferences for sequences of health and money. *Organizational Behavior and Human Decision Processes, 67*(1), 59-75.

Chapman, G. B. (1996b). Temporal discounting and utility for health and money. *Journal of Experimental Psychology-Learning Memory and Cognition, 22*(3), 771-791.

Chapman, G. B. (2001). Time preferences for the very long term. *Acta Psychologica, 108*(2), 95-116.

Chapman, G. B. (2003). Time discounting of health outcomes. . In Loewenstein, G., Read, D., & Baumeister, R. F. (Eds). *Time and Decision: Economic and psychological perspectives on intertemporal choice.* New York: Russell Sage Foundation.

Chapman, G. B., Brewer, N. T., Coups, E. J., Brownlee, S., Leventhal, H., & Leventhal, E. A. (2001). Value for the future and preventive health behavior. *Journal of Experimental Psychology-Applied, 7*(3), 235-250.

Chapman, G. B., & Elstein, A. S. (1995). Valuing the Future - Temporal Discounting of Health and Money. *Medical Decision Making, 15*(4), 373-386.

Chapman, G. B., & Winquist, J. R. (1998). The magnitude effect: Temporal discount rates and restaurant tips. *Psychonomic Bulletin & Review, 5*(1), 119-123.

Coller, M., & Williams, M. B. (1999). Eliciting individual discount rates. *Experimental Economics*, 2, 107-127.

Dawkins, R. (1982). *The extended phenotype*. Oxford: Oxford University Press.

Deaton, A. (1988). Consumers' expenditure. In Eatwell, G., Milgate, M. & Newman, P. (Eds.) *The New Palgrave: A dictionary of economics*. Houndmills, UK: Palgrave.

Elster, J. (1979). *Ulysses and the sirens*. Cambridge: Cambridge University Press.

Fisher, I. (1930). *The theory of interest*. New York: Macmillan.

Frederick, S. (2003). Time preference and personal identity. In Loewenstein, G., Read, D., & Baumeister, R. F. (Eds). *Time and Decision: Economic and psychological perspectives on intertemporal choice*. New York: Russell Sage Foundation.

Frederick, S. (2003). Measuring intergenerational time preference: Are future lives valued less? *Journal of Risk and Uncertainty, 26*(1), 39-53.

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature, 40*(2), 351-401.

Gigliotti, G., & Sopher, B. (1997). Violations of present-value maximization in income choice. *Theory and Decision, 43*(1), 45-69.

Green, L., Fristoe, N., & Myerson, J. (1994). Temporal Discounting and Preference Reversals in Choice between Delayed Outcomes. *Psychonomic Bulletin & Review, 1*(3), 383-389.

Green, L., & Myerson, J. (1996). Exponential versus hyperbolic discounting of delayed outcomes: Risk and waiting time. *American Zoologist, 36*(4), 496-505.

Green, L., Myerson, J., & McFadden, E. (1997). Rate of temporal discounting decreases with amount of reward. *Memory & Cognition, 25*(5), 715-723.

Green, L., Myerson, J., & Ostaszewski, P. (1999). Discounting of delayed rewards across the life span: age differences in individual discounting functions. *Behavioural Processes, 46*(1), 89-96.

Harvey, C. M. (1994). The reasonableness of nonconstant discounting. *Journal of Public Economics, 53*(1), 31-51.

Hoch, S. J., & Loewenstein, G. F. (1991). Time-Inconsistent Preferences and Consumer Self-Control. *Journal of Consumer Research, 17*(4), 492-507.

Kacelnik, A. (2003). The evolution of patience. In Loewenstein, G., Read, D., & Baumeister, R. F. (Eds). *Time and Decision: Economic and psychological perspectives on intertemporal choice*. New York: Russell Sage Foundation.

Kagel, J. H., Battalio, R. C., & Green, L. (1995). *Economic choice theory: An experimental analysis of animal behavior*. Cambridge: Cambridge University Press.

Kirby, K. N. (1997). Bidding on the future: Evidence against normative discounting of delayed rewards. *Journal of Experimental Psychology-General, 126*(1), 54-70.

Kirby, K. N., & Herrnstein, R. J. (1995). Preference Reversals Due to Myopic Discounting of Delayed Reward. *Psychological Science, 6*(2), 83-89.

Koopmans, T. C. (1960). Stationary ordinal utility and impatience. *Econometrica, 28*, 287-309.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, 112, 443-477.

Laibson, D. (1998). Life-cycle consumption and hyperbolic discount functions. *European Economic Review, 42*(3-5), 861-871.

Leland, J. W. (2002). Similarity judgments and anomalies in intertemporal choice. *Economic Inquiry, 40*(4), 574-581.

Loewenstein, G. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes, 65*(3), 272-292.

Loewenstein, G., Nagin, D., & Paternoster, R. (1997). The effect of sexual arousal on expectations of sexual forcefulness. *Journal of Research in Crime and Delinquency, 34*(4), 443-473.

Loewenstein, G., & Prelec, D. (1992). Anomalies in Intertemporal Choice - Evidence and an Interpretation. *Quarterly Journal of Economics, 107*(2), 573-597.

Loewenstein, G., & Sicherman, N. (1991). Do Workers Prefer Increasing Wage Profiles. *Journal of Labor Economics, 9*(1), 67-84.

Loewenstein, G. F. (1988). Frames of Mind in Intertemporal Choice. *Management Science, 34*(2), 200-214.

Loewenstein, G. F., & Prelec, D. (1993). Preferences for Sequences of Outcomes. *Psychological Review, 100*(1), 91-108.

Madden, G. J., Bickel, W. K., & Jacobs, E. A. (1999). Discounting of delayed rewards in opioid-dependent outpatients: Exponential or hyperbolic discounting functions? *Experimental and Clinical Psychopharmacology, 7*(3), 284-293.

Manuck, S.B., Flory, J.D., Muldoon, M. F., & Ferrell, R. E. (2003). A neurobiology of intertemporal choice. In Loewenstein, G., Read, D., & Baumeister, R. F. (Eds). *Time and Decision: Economic and psychological perspectives on intertemporal choice.* New York: Russell Sage Foundation.

Mazur, J. E. (1984). Tests of an equivalence rule for fixed and variable delays. *Journal of Experimental Psychology: Animal Behavior Processes*, *10*, 426-436.

Miller, N. E. (1959). Liberalization of basic S-R concepts: Extension to conflict behavior, motivation and social learning. In S. Koch (Ed.) *Psychology: A study of a science*. New York, McGraw-Hill.

Mischel, W. (1974). Processes in delay of gratification. *Advances in Experimental Psychology*, 7, 249-292.

Mischel, W., Ayduk, O., & Mendoza-Denton, R. (2003). Sustaining delay of gratification over time: A hot-cool systems perspective. In Loewenstein, G., Read, D., & Baumeister, R. F. (Eds). *Time and Decision: Economic and psychological perspectives on intertemporal choice.* New York: Russell Sage Foundation.

Mowen, J. C., & Mowen, M. M. (1991). Time and Outcome Valuation - Implications for Marketing Decision-Making. *Journal of Marketing, 55*(4), 54-62.

Mulligan, Casey. B. (1996). 'A logical economist's argument against hyperbolic discounting.' Working paper, University of Chicago.

Nozick, R. (1993) *The Nature of Rationality*. Princeton: Princeton University Press.

O'Donoghue, T., & Rabin, M. (2000). The economics of immediate gratification. *Journal of Behavioral Decision Making, 13*(2), 233-250.

O'Donoghue, T., & Rabin, M. (2001). Choice and procrastination. *Quarterly Journal of Economics, 116*(1), 121-160.

Parfit, D. (1984). *Reasons and Persons*. Oxford University Press.

Peart, S. J. (2000). Irrationality and intertemporal choice in early neoclassical thought. *Canadian Journal of Economics-Revue Canadienne D Economique, 33*(1), 175-189.

Price, A. W. (1995). *Mental conflict*. London: Routledge.

Pigou, A. C. (1920) *Economics of Welfare*. London: Macmillan.

Rachlin, H., and Green, L. (1972). Commitment, choice and self-control. *Journal of the Experimental Analysis of Behavior, 17*, 15-22.

Rachlin, H. (1989). *Judgment, decision and choice*. New York: Freeman.

Raineri, A. & Rachlin, H. (1993). The effect of temporal constraints on the value of money and other commodities. *Journal of Behavioral Decision Making, 6*, 77-94.

Rubinstein, A. (2000). *Is it "economics and psychology"? The case of hyperbolic discountin*g. Tel Aviv: Tel Aviv University and Princeton University.

Read, D. (2000). Can the concept of behavioural mass help explain nonconstant time discounting? *Behavioral and Brain Sciences, 23*(1), 111-+.

Read, D. (2001a). Intrapersonal Dilemmas. *Human Relations*, 54, 1093-1117.

Read, D. (2001b). Is time-discounting hyperbolic or subadditive? *Journal of Risk and Uncertainty, 23*(1), 5-32.

Read, D. (2003). Time and the marketplace. Working paper, London School of Economics.

Read, D., & Powell, M. (2002). Reasons for sequence preferences. *Journal of Behavioral Decision Making, 15*(5), 433-460.

Read, D. & Roelofsma, P.H.M.P. (2003). Subadditive versus hyperbolic discounting: A comparison of choice and matching. *Organizational Behavior and Human Decision Processes*, *91*, 140-153.

Read, D., & Van Leeuwen, B. (1998). Predicting hunger: The effects of appetite and delay on choice. *Organizational Behavior and Human Decision Processes, 76*(2), 189-205.

Roelofsma, P., & Read, D. (2000). Intransitive intertemporal choice. *Journal of Behavioral Decision Making, 13*(2), 161-177.

Rogers, A. R. (1994). Evolution of Time Preference by Natural-Selection. *American Economic Review, 84*(3), 460-481.

Schelling, T. C. (1995). Intergenerational Discounting. *Energy Policy, 23*(4-5), 395-401.

Shelley, M. K. (1993). Outcome Signs, Question Frames and Discount Rates. *Management Science, 39*(7), 806-815.

Shelley, M. K. (1994). Gain Loss Asymmetry in Risky Intertemporal Choice. *Organizational Behavior and Human Decision Processes, 59*(1), 124-159.

Sidgwick, H. (1874). *Methods of Ethics*. Indianapolis: Hackett.

Soman, D. (1998). The illusion of delayed incentives: Evaluating future effort- money transactions. *Journal of Marketing Research, 35*(4), 427-437.

Sozou, P. D. (1998). On hyperbolic discounting and uncertain hazard rates. *Proceedings of the Royal Society of London Series B-Biological Sciences, 265*(1409), 2015-2020.

Sozou P.D., & Seymour R.M. (2003). Augmented discounting: interaction between ageing and time-preference behaviour. *Proceedings of the Royal Society of London Series B-Biological Sciences*, *265*, 1047-1053

Strotz, R H. (1955). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies*, *23*, 165-180.

Thaler, R. (1981). Some empirical evidence of dynamic inconsistency. *Economics Letters*, *8*, 201-207.

Thaler, R. H., & Shefrin, H. M. (1981). An Economic-Theory of Self-Control. *Journal of Political Economy, 89*(2), 392-406.

Trope, Y., & Liberman, N. (2000). Temporal construal and time-dependent changes in preference. *Journal of Personality and Social Psychology, 79*(6), 876-889.

Trope, Y., & Liberman, N. (2003). Construal level theory of intertemporal judgment and decision. In Loewenstein, G., Read, D., & Baumeister, R. F. (Eds). *Time and Decision: Economic and psychological perspectives on intertemporal choice*. New York: Russell Sage Foundation.

Trostel, P. A., & Taylor, G. A. (2001). A theory of time preference. *Economic Inquiry, 39*(3), 379-395.

Van der Pol, M. M., & Cairns, J. A. (2000). Negative and zero time preference for health. *Health Economics, 9*(2), 171-175.

Wertenbroch, K. (2003). Self-rationing: Self-control in consumer choice. In Loewenstein, G., Read, D., & Baumeister, R. F. (Eds). *Time and Decision: Economic and psychological perspectives on intertemporal choice*. New York: Russell Sage Foundation.

Wilson, J. Q., & Herrnstein, R. J. (1985). *Crime and Human Nature*. New York: Simon & Schuster.

Yates, J. F., & Watts, R. A. (1975). Preferences for deferred losses. *Organizational Behavior and Human Performance, 13*, 294-306

Endnotes

[1] Readers who wish to find out more are directed to two volumes of chapters, Loewenstein & Elster (1992) and Loewenstein, Read & Baumeister (2003), and to a longer survey article by Frederick, Loewenstein & O'Donoghue (2002, reprinted in Loewenstein et al., 2003).

[2] We gain utility from consumption of resources. A consumption stream is a series of dated consumption episodes.

[3] Fisher's analysis assumes a perfect capital market, in which borrowing and lending can freely be carried out at the same rate, and that one can borrow as much or as little as desired subject to the constraint that total income is held constant. In a real market, people will differ in their borrowing opportunities. Some will have credit cards offering 0% interest rates, some will not be able to borrow at all. To anticipate the future discussion, if they are fully rational it is these opportunities that will constrain their intertemporal decision making about money.

[4] Opportunity cost is what can be earned from the best alternative use of some resources. To minimize opportunity costs means to choose the best, since the best is not amongst the foregone alternatives.

[5] Frederick (2003) has tested whether people's intuitions are Parfitian. He obtained the correlation between measures of time discounting and identification with one's future self. He found no relationship.

[6] The 'selves' in the hyperbolic discounting analysis are useful analytic fictions, unlike the real selves discussed by Parfit (1984) and his predecessors.

[7] Thaler (1981) who first described many of the anomalies, and who provided a model for most subsequent studies, *did* frame his work as a test of Fisher's normative model.

[8] There are many circumstances, however, when decreasing sequences are preferred. For very long sequences (such as lifetime health and income), people are very likely to choose decreasing sequences over increasing ones, and this tendency is stronger for health than money (Chapman, 1996a; Read & Powell, 2002).

[9] When the outcomes are non-fungible, such as ways of spending a weekend or experiencing health, this normative argument does not apply.

[10] Elasticity is the percentage change in amount divided by the percentage change in value. For instance, if $v=A^2$, then when A is 10 v=100. If A is increased by 10% to 11, V would increase in 121 – a 21% change. The elasticity would be 21/10=2.1.

[11] Hyperbolic discounting will say that the penalty for each day of delay will be less than its successor. If the 1st to 61st day delay can be compensated for by \$120, then the 61st day alone can certainly be compensated for by \$2.