The overall discount rate

Alexander J.H. Houston[a], Alasdair I. Houston[b],*

[a] Department of Physics, University of Warwick, Coventry, West Midlands, CV4 7AL, UK
[b] School of Biological Sciences, Bristol Life Sciences Building, University of Bristol, Tyndall Avenue, Bristol, BS8 1TQ, UK

ARTICLE INFO

ABSTRACT

Discounting refers to the way in which the value of an outcome depends on the delay until it is obtained. If an organism's discount function is known, then its rate of discounting at any delay can be found. If the function is not known, the normalised area under an estimate of the discount function has been used as a measure that summarises the strength of discounting over a range of delays. We propose a new measure of the strength of discounting: the overall discount rate W, which is the drop in value from the start to the end of the value curve divided by the area under this curve. We show that our measure has various advantages over the normalised area, namely it can be linked to the instantaneous rate of discounting and it respects the special nature of exponential discounting. It does not give a unique value to each curve, but we prove that this is incompatible with the requirement that similar discount functions be assigned similar values, and so this is not a defect of the measure.

1. Introduction

The value of an outcome for an organism typically decreases as the delay before it is obtained increases. This phenomenon, known as discounting, has been demonstrated in a range of organisms including bees (Cheng et al., 2002), pigeons (Mazur et al., 1987; Mazur and Biondi, 2009; Green et al., 2010), rats (Mazur and Biondi, 2009) and humans (Green et al., 1994; Reynolds et al., 2004; Charlton et al., 2013; Holt et al., 2014). From its origin in economics (Samuelson, 1937), it has become a focus of research in psychology (Green and Myerson, 2004; Killeen, 2009; Vanderveldt et al., 2016) and behavioural ecology (Sozou, 1988; Stephens et al., 2004; Fawcett et al., 2012; Stevens, 2014). Discounting also plays a central role in the investigation of drug addiction, gambling and mental illness (Reynolds, 2006; Bickel et al., 2014; Horan et al., 2017). A complete characterisation of discounting can be given in two equivalent ways. One approach is to specify how an animal values a reward that is obtained after a delay D by the discount function $V(D)$. The alternative is to specify how the instantaneous discount rate $r(D)$ depends on D. The link between these approaches is that

$$r(D) = -\frac{1}{V(D)}\frac{dV(D)}{dD} \qquad (1)$$

e.g. Laibson (1997); Prelec (2004), see also Kagel et al. (1986), Takahashi et al. (2008). In studies of non-human subjects, the discount function can be estimated by giving the subject a choice between two options that differ in amount and delay and finding pairs of options that the subject regards as being of equal value, in that it has no preference for one as opposed to the other (Mazur et al., 1987). The delays used in such experiments are often less than one minute (Mazur et al., 1987). Although humans are sometimes studied using this procedure (Rodriguez and Logue, 1988), their value functions are often assessed by asking questions about preferences (Green et al., 1994; Reynolds et al., 2004; Charlton et al., 2013; Holt et al., 2014; Weatherly, 2014; Mitchell et al., 2015; Yoon et al., 2017); the delays investigated may be longer than a year (Green et al., 1994). A general goal is to go from such data to a measure that characterises the strength of discounting, with a common procedure being to assume a form of discount function, and find the parameters of the equation that give the best fit. Thus this procedure starts with a model of discounting and uses the parameters of the model as measures of discounting. In this paper we develop a general measure of discounting that can be applied to all data of the form that we have described. We also formalise an existing measure proposed by Myerson et al. (2001) and establish that not all of the desirable features of a measure can be achieved.

2. Models of discounting

The standard model in economics is exponential discounting:

$$V = A \exp(-\theta D), \qquad (2)$$

* Corresponding author.
E-mail addresses: a.houston@warwick.ac.uk (A.J.H. Houston), a.i.houston@bristol.ac.uk (A.I. Houston).

where A is the value of the outcome in the absence of any delay i.e. the amount. In this case the instantaneous discount rate is θ, as can be seen by substituting this form of V into Eq. (1). In fact, exponential discounting is the only form of discounting for which the instantaneous discount rate is constant; setting r to be constant in Eq. (1) gives $V' \propto V$ which defines an exponential function. Thus exponential discounting means that the rate of decrease in value is proportional to the current value and so the decision-maker is essentially always discounting in the same way, regardless of how long it has been waiting. This gives exponential discounting a special status that we elaborate upon in Section 3.

An alternative model to exponential discounting is the hyperbolic discounting function

$$V = A/(1 + kD), \qquad (3)$$

where k is a rate parameter that determines how value decreases with delay. At a delay of $1/k$ the value of an option is $A/2$. In many cases Eq. (3) gives a better fit to the data than exponential discounting (Rodriguez and Logue, 1988; Mazur and Biondi, 2009; Charlton et al., 2013). The generalised hyperbolic discounting function

$$V = A/(1 + kD)^s \qquad (4)$$

has also been considered (Loewenstein and Prelec, 1992; McKerchar et al., 2010; Holt et al., 2014), and has been justified by Myerson and Green (1995). We concentrate on these frequently studied cases; other forms of discounting are reviewed by Doyle (2013).

3. Measuring discounting

When data have been obtained from more than one experiment, it can be difficult to use V or r to assess whether the strength of discounting differs between the experiments. It is not obvious what the forms of the equations should be, nor how to compare them across experiments even if they are known. What is desirable is a measure of discounting with the following properties:

i it assigns a single value to a curve
ii this value is unique to the curve
iii similar curves are assigned similar values
iv the special status of exponential discounting as having a constant instantaneous discount rate is reflected
v the measure can be applied to a set of data points without a particular form of discount function being assumed

We elaborate on the last two points. Point (iv) is perhaps most easily explained via an analogy between two fundamental forms of constant or memoryless process, those described by a constant change in absolute value and those with a constant fractional change. The first of these corresponds to a linear graph with constant gradient, while the second imposes the condition that the gradient divided by the value is constant. In the context of discounting this is the statement that the instantaneous discount rate is constant and specifies an exponential form of discounting. Let us now consider assigning a measure to characterise such processes. Take an object travelling at a constant speed v, so its graph of distance against time is a straight line with constant slope v. If a proposed measure assigned an overall speed that depended on the time for which the object travels, we would not regard this as satisfactory. An obvious measure is total distance divided by total time, which assigns a speed v to the object for all times travelled. In discounting, the value function can never be negative and so it is clearly the exponential which is of relevance here and it is equally unsatisfactory to have a measure which does not highlight the uniquely constant form of discounting that it describes. As we show below, the area under the discount curve (Myerson et al., 2001) assigns an overall discount rate to exponential discounting that depends on the time period over which discounting is measured (see Eq. (6) and Fig. 1). In contrast, our measure results in a
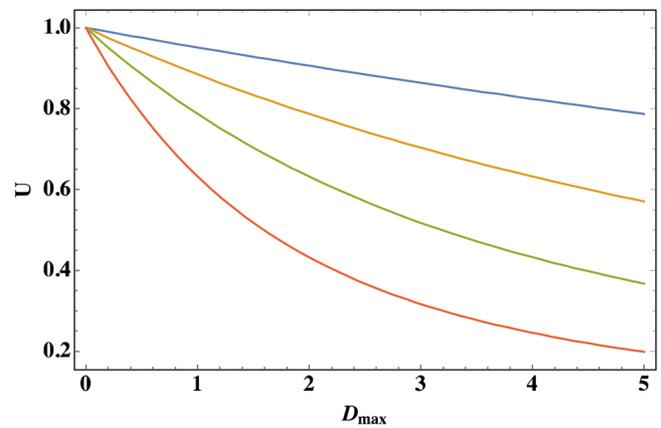

Fig. 1. The normalised area under the curve U as a function of D_{max} when discounting is exponential (Eq. (2)) for four values of the instantaneous discount rate θ. From top to bottom, $\theta = 0.1, 0.25, 0.5, 1.0$. Note that lower values of U correspond to stronger discounting.

value that is equal to the instantaneous discount rate θ regardless of the time period (see Eq. (10)). To be clear, we are not saying anything about the likelihood of animals discounting in an exponential fashion, rather that an adequate discount measure should capture the special status of exponential discounting.

As for point (v), we emphasise that although we can apply our method to data without assuming a particular form of discount function, we can also investigate the value assigned by the measure in particular cases. This is how we establish the result about exponential discounting mentioned above. The same approach can be applied to any form of discounting e.g. the q-exponential discount function (Cajueiro, 2006; Takahashi et al., 2008) or the functions in Doyle (2013), although an explicit equation may not be obtainable in all cases.

No current measures of discounting satisfy all the requirements that we have listed. Myerson et al. (2001) proposed that the area under the discount curve is a suitable measure. This measure satisfies three of the above points (points (i), (iii) and (v)). In this paper we build on the work of Myerson et al. to develop a new measure which satisfies four of the above points, with it failing to always provide a unique value to each curve. We also prove that it is impossible to satisfy uniqueness if similar curves are assigned similar values, i.e. we cannot have a measure that satisfies both (ii) and (iii). We call our measure the overall discount rate and compare it to the area-under-the-curve measure advocated by Myerson et al. (2001), which we formalise. We also show how our rate can be linked to the instantaneous discount rate and argue that it is as good as possible given our proof that no measure can have all the desirable features.

3.1. Normalised area under the curve

Myerson et al. (2001) present a measure of discounting that does not assume a particular form of value function. They start from data that provide an estimate of the value of an outcome at various delays, i.e. an estimate of the value function. Their procedure is to normalise each axis by dividing by its maximum value, either A or D_{max}, so that each axis runs from zero to one. They then find the area underneath the value function, giving a measure of discounting that lies between zero and one. For discussion of the calculation of the area see Borges et al. (2016) and Gilroy and Hantula (2018). The result of this procedure is the following measure:

$$U = \frac{\int_0^{D_{max}} V(D)dD}{AD_{max}}. \qquad (5)$$

Calculating the area under the value curve without normalising the

axes would simply give the integral in the numerator. Normalising the axes by A and D_{max} has the effect of reducing the answer by a factor of AD_{max}. This can be clearly seen by considering the un-normalised curve in a rectangle with sides of A and D_{max}. When the axes are normalised the fraction of the rectangle which lies under the curve will not change, but the area of the whole rectangle is divided by AD_{max}. Indeed, it is the fraction of the area of the rectangle lying under the curve which is given by this measure, meaning the answer is a dimensionless number between zero and one, with a value of one corresponding to no discounting and a value of zero corresponding to total and immediate discounting. Normalising means that the measure does not depend on A, but it does still depend on D_{max}, because D_{max} does not just provide a scaling but also affects the shape of the value curve. Putting the appropriate forms of V into Eq. (5) allows us to evaluate the measure for the three forms of discounting outlined previously. For exponential discounting this gives

$$U = \frac{1 - e^{-\theta D_{max}}}{\theta D_{max}} \tag{6}$$

which can be expressed in terms of the dimensionless parameter $x = \theta D_{max}$ as $U = \frac{1 - e^{-x}}{x}$. A representation such as this is always possible because U is dimensionless.

In Fig. 1 we give some examples of Eq. (6) for different values of the exponential discounting parameter θ. It can be seen that U drops from an initial value of one in all cases. This is a general feature that results from normalising by D_{max}. Whilst the area under the curve vanishes in the limit as D_{max} tends to zero, the value curve becomes essentially horizontal and the fraction of the available area that lies under the curve tends to one. In this figure and in Figs. 2 and 4 we have not plotted the results in terms of a dimensionless parameter because we feel that it is helpful to retain the usual discounting parameters.

As Killeen (2015), p12 shows, for the hyperbolic equation and the generalised hyperbolic equation the area-under-the-curve measure gives

$$U = \frac{\ln(1 + kD_{max})}{kD_{max}} = \frac{\ln(1 + x)}{x} \tag{7}$$

and

$$U = \frac{1 - (1 + kD_{max})^{1-s}}{kD_{max}(s - 1)} = \frac{1 - (1 + x)^{1-s}}{x(s - 1)} \tag{8}$$

respectively, where now $x = kD_{max}$. Fig. 2a gives some examples of Eq. (7), illustrating the effect of k in the region that characterises the discounting behaviour of rats and pigeons obtaining food ($k = 0.14 \text{ s}^{-1}$ for rats, $k = 0.64 \text{ s}^{-1}$ for pigeons, Mazur and Biondi, 2009). Fig. 2b gives some examples of Eq. (8), illustrating the effect of s when $k = 1$.

Mitchell et al. (2015) fit the hyperbolic model (Eq. (3)) to the discounting behaviour of each of their subjects and then plot the area under the curve U against k (see their Fig. 2). Eq. (7) gives the mathematical relationship between U and k if Eq. (3) holds exactly and hence approximates it if Eq. (3) gives a good account of the data. When this is the case, Eq. (7) also determines the relationship between U and $\ln k$ (Mitchell et al., 2015) or $\log k$ (Yoon et al., 2017). As is noted above, k only enters Eq. (7) via the dimensionless variable $x = kD_{max}$. The dependence of U on x and $\ln x$ is shown in Fig. 3 and curves from different experiments should lie along these general curves when U is plotted against $x = kD_{max}$ or $\ln x$ rather than k or $\ln k$. In line with this, the difference between the three curves in Fig. 2 of Mitchell et al. (2015) seems to be the result of different values of D_{max}. Mitchell et al. and Yoon et al. give the impression that the functional dependence of U on $\ln k$ or $\log k$ provides additional information to that given by the dependence of U on k. Our purpose here is to underline the fact that plots of the form of Fig. 3 give us nothing not already available from Fig. 2 since both are fully determined by Eq. (7). They are simply different representations of the same relationship.

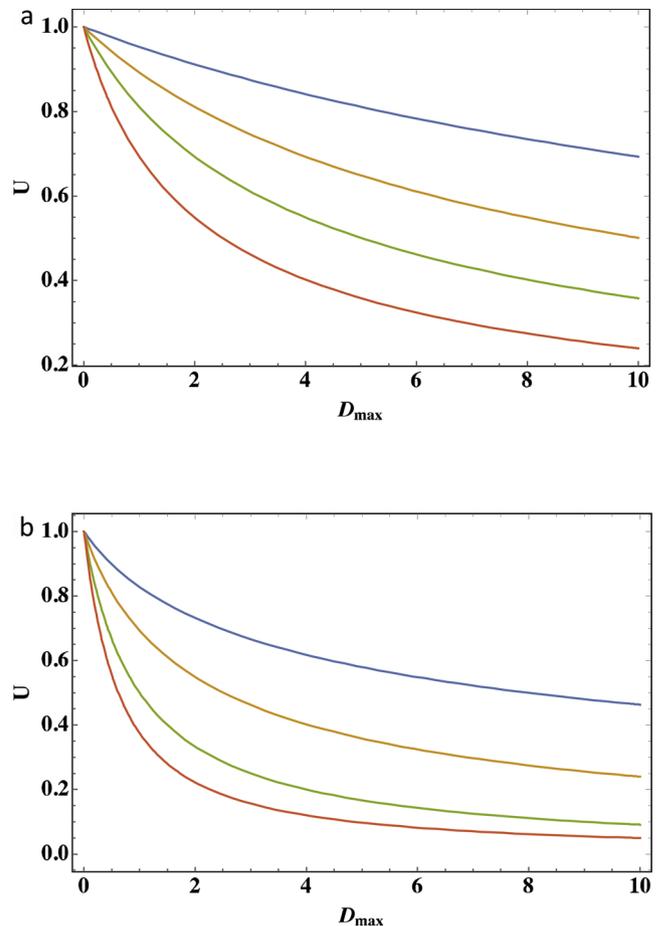There are some problems with the area under the curve U as a

Fig. 2. The normalised area under the curve U as a function of D_{max} for hyperbolic and generalised hyperbolic discounting. (a) discounting is given by Eq. (3), U is given by Eq. (7). From top to bottom, $k = 0.1, 0.25, 0.5, 1.0$. (b) discounting is given by Eq. (4) with $k = 1$. U is given by Eq. (8). From top to bottom, $s = 0.5, 1.0, 2.0, 3.0$. (The curve for $s = 1$ is calculated using Eq. (7)).

measure of discounting. Myerson et al. (2001) discuss the fact that two different curves can have the same area (with fixed D_{max}), suggesting that in such cases it might be useful to compare the curves over more than one range of delays. Another problem is that it is not clear how U relates to a measure such as the instantaneous discount rate.

In the next section we present our measure and show how it can be linked to the instantaneous discount rate. We also give the form of our measure for the three cases we have discussed.

3.2. The overall discount rate

There are two ways to obtain an overall discount rate from the instantaneous discount rate r. One way is integrate r to get

$$\int_0^{D_{max}} r \, dD = -\int_0^{D_{max}} \frac{V'}{V} dD = -[\ln V]_0^{D_{max}} = \ln\left(\frac{V(0)}{V(D_{max})}\right).$$

This measure depends only on the endpoints; the pattern of discounting between these values is irrelevant, which is not a desirable feature of an overall measure of discounting. An alternative is to integrate the numerator and denominator separately and define

$$W = -\frac{\int_0^{D_{max}} V' dD}{\int_0^{D_{max}} V \, dD} = \frac{V(0) - V(D_{max})}{\int_0^{D_{max}} V \, dD} \tag{9}$$
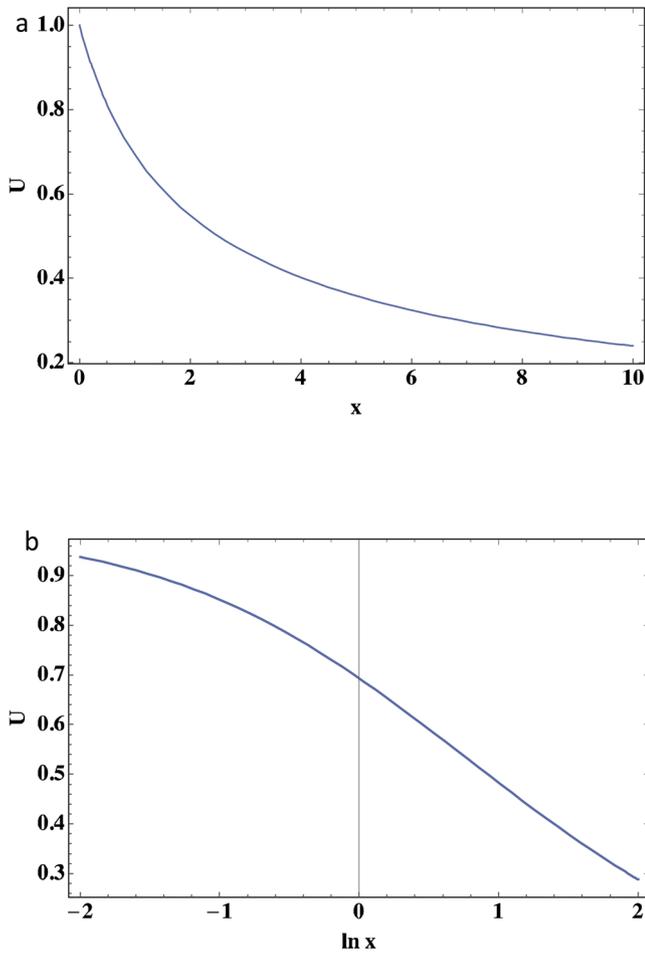
Fig. 3. The relationship between U and k or $\ln k$ when discounting is hyperbolic (see Eq. (3)). Eq. (7) shows that k only appears via the dimensionless variable $x = kD_{max}$. (a) U as a function of x. $U = 1$ when $x = 0$ and tends to 0 as x tends to ∞. (b) U as a function of $\ln x$. U tends to 1 as $\ln x$ tends to $-\infty$ and tends to 0 as $\ln x$ tends to ∞.

Fig. 4. The overall discount rate W as a function of D_{max}. (a) discounting is given by Eq. (3), W is given by Eq. (11). From top to bottom, $k = 1.0$, 0.5, 0.25, 0.1. (b) discounting is given by Eq. (4) with $k = 1$. W is given by Eq. (12). From top to bottom, $s = 3.0$, 2.0, 1.0, 0.5. (The curve for $s = 1$ is calculated using Eq. (11)).

$$= \frac{1 - V(D_{max})}{\int_{0}^{D_{max}} V dD}$$

$$V(0)$$

In this measure the numerator is determined solely by the endpoints but the denominator takes into account the entire pattern of discounting. We therefore prefer this second procedure and take W as our measure, that is the drop in value divided by the area under the curve. We refer to this as the overall discount rate. The numerator is a value while the denominator has dimensions of value multiplied by time, making W a rate, in contrast to the dimensionless U.

Like the area-under-the-curve measure U, our discount rate W does not assume a particular form of value function. It is, however, instructive to evaluate W in various specific cases, allowing us to make quantitative comparisons with the earlier analysis of U and delineate the relative merits of each measure. One key difference between W and U is that our measure preserves the special status of exponential discounting. The instantaneous discount rate r typically gives a measure which is dependent on D as opposed to merely being dependent on the D_{max} used in a particular experiment. However, when the discounting is exponential r is a constant, independent of both D and D_{max}. It therefore seems that a good measure of discounting should judge an exponential form of V to correspond to constant discounting. As shown in the previous section, this feature is lost when the area under the value
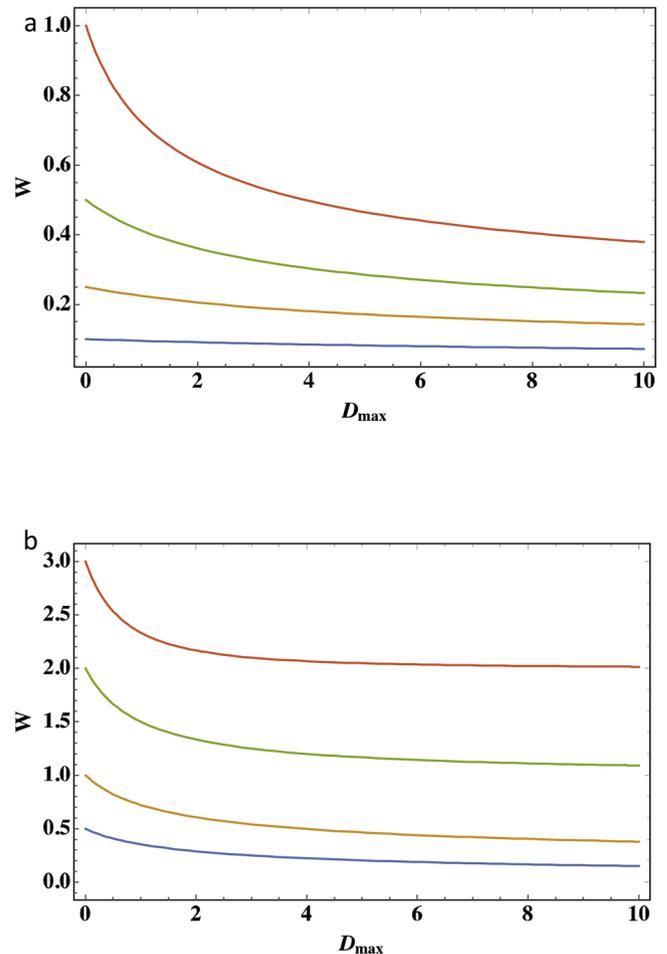
curve U is used as a measure of discounting. In contrast, our measure W retains this behaviour, giving

$$W = \frac{A(1 - e^{\theta D_{max}})}{\frac{A(1 - e^{\theta D_{max}})}{\theta}} = \theta. \tag{10}$$

Our discounting measure therefore gives the same result for the exponential case as the instantaneous discount rate. More generally the two rates are related as shown in Eq. (9), from which it can be seen that as D_{max} tends to zero W tends to $r(0)$. This means that when the length of the experiment becomes very short the overall discount rate, W, tends towards the initial instantaneous discount rate. W therefore allows useful information about the form of discounting to be attained from short experiments.

We now apply our measure to hyperbolic discounting functions. When discounting is given by Eq. (3) we have

$$W = \frac{k^2 D_{max}}{(1 + kD_{max})\ln(1 + kD_{max})} \tag{11}$$

while for discounting described by Eq. (4)

$$W = \frac{k(s - 1)((1 + kD_{max})^s - 1)}{(1 + kD_{max})^s - (1 + kD_{max})}. \tag{12}$$

Fig. 4a and b give some examples, using the same values of k and s that are used in Fig. 2a and b. It follows from the general result discussed above that as D_{max} tends to zero, W tends to k and sk for

hyperbolic and generalised hyperbolic discounting respectively.

We note that whereas U always tends to zero as $D_{\max}$ tends to infinity, W can tend to informative values. For the generalised hyperbolic model with $s > 1$ W asymptotes to $k(s-1)$, with the rate of approach being faster for larger s, as can be observed from the top two curves in Fig. 4b. We have already highlighted how W is constant for all $D_{\max}$ when the discounting function is exponential.

4. The problem of uniqueness

We have identified five desirable properties for a measure of discounting. Property (iii) requires that similar curves are assigned similar values. This is captured by the mathematical concept of continuity. The requirement of uniqueness (property (ii)) means that the measure constitutes an injective map. Both these terms are explained in the appendix, where we show that it is not possible for a measure to satisfy both properties (ii) and (iii). The argument does not rely on the concept of discounting, and indeed all of the criteria given above except (iv) are reasonable requirements of any measure. The above limitation therefore applies to all attempts to assign a single measure to a function.

5. Discussion

In this paper we have proposed a new measure of the strength of discounting: the overall discount rate W. As part of our investigation we have extended the formalisation of the normalised area under the value curve, U, begun by Killeen (2015). Our analysis shows that if discounting is hyperbolic (i.e. is given by Eq. (3)), then Eq. (7) determines the relationship between U and k, $\ln k$ or $\log k$. Furthermore, the only relevant variable is $x = kD_{\max}$. The resulting general curve in the case of U as a function of $\ln x$ is shown in Fig. 3b, which has the same form as the data of Mitchell et al. (2015).

Our measure W is as easy to compute as U and, like U, W is theoretically neutral in that it does not assume a particular form of value function. We have shown that our measure has various advantages over that of Myerson et al. (2001). It respects the special nature of exponential discounting and its dimensions correspond to those of the standard measure of discounting, the instantaneous discount rate r, from which W can be derived. Another noteworthy feature is the behaviour of the two measures when $D_{\max}$ is small. We have shown that U always tends to one as $D_{\max}$ tends to zero. In contrast, W tends to the initial instantaneous discount rate $r(0)$. As a consequence, values of U are close together when $D_{\max}$ is small, but values of W can be noticeably different. This difference can be seen by comparing Figs. 2 and 4. The case of exponential discounting is even more striking. Fig. 1 shows how U decreases from one in this case. The behaviour of W is very different; W is equal to θ for all values of $D_{\max}$. Thus, even in experiments with short delay times, the overall discount rate W can be used to distinguish between different patterns of discounting. An apparent problem that our measure shares with that of Myerson et al. (2001) is that it cannot guarantee that any two discount curves can be distinguished for all values of $D_{\max}$. If we consider a pair of intersecting curves, there exists a value of $D_{\max}$ for which the areas under the two curves are equal and the Myerson measure cannot distinguish them. Similarly, there will exist a generically different value of $D_{\max}$ at which our measure returns the same number for both curves. Rather than being a defect of the measure of discounting, this can be seen to make intuitive sense, since an intersection can only occur by one of the two curves displaying greater discounting than the other initially and then less later on. It is therefore reasonable that the two curves would be judged as corresponding to equal discounting when considered over a period covering both these regimes.

The curves can always be distinguished by bringing in another piece of information. Myerson et al. (2001) suggest calculating the areas under each curve at two different values for $D_{\max}$; in the context of our approach, both the ratio of the drop to area (i.e. W) and the absolute drop $V(0) - V(D_{\max})$ used at a single value of $D_{\max}$ would perhaps be more natural. Although the curves can now be differentiated, e.g. by saying that one has stronger discounting at low D than the other, there is no longer a single number describing the discounting of the curve. Once such additional bits of information are used there is then no consistent way to combine them to give a single value that is unique for every curve. This highlights a general tension that exists between having a measure that gives a single value to a curve and one that is unique for every curve. The instantaneous discount rate lies at one extreme; since it is defined the whole way along the curve this infinite set of values determines a particular function and is therefore unique to that curve. On the other hand the overall discount rate W and the normalised area under the curve U both give a single value to a curve, but for any given curve and value of $D_{\max}$, there exists another curve which the measure cannot distinguish. We have proved that this limitation is inevitable if the measure assigns similar values to similar discount curves. We think that this is a fundamental property of a good measure, which suggests that our measure has as many desirable features as possible.

Competing interests

We have no competing interests.

Appendix A

Here we give a formal argument for the statement made in Section 4. We begin by setting up the problem. Any measure of discounting will be a function which assigns a number to each possible data set. If the data set contains n points $(x_1, y_1), (x_2, y_2), . . ., (x_n, y_n)$ it can be thought of as the single point $(y_1, y_2, ..., y_n)$ in n-dimensional space (assuming that the x-values of the points are fixed and the data can only differ by the n y-values, otherwise the data set could be associated with a point in $2n$-dimensional space). The measure will then be a map $f : \mathbb{R}^n \to \mathbb{R}$ and we would like this map to be both continuous and injective. Being continuous means that making a small change to the data set results in a small change to the discounting measure, or formally that for all $\varepsilon > 0$ there exists a $\delta > 0$ such that whenever $|a - b| < \delta$ then $|f(a) - f(b)| < \varepsilon$. An injective map is one-to-one, meaning that distinct data sets, i.e. distinct points in $\mathbb{R}^n$, are mapped to distinct real numbers. However, it is impossible to have a measure which satisfies both of these criteria. We shall show this by imagining a continuous and injective map $f : \mathbb{R}^n \to \mathbb{R}$ which will lead to a contradiction. The argument is then as follows. We use the fact that $\mathbb{R}^n$ is path connected, meaning that for any two points a, b in $\mathbb{R}^n$ there is a continuous path between them. (A path is a map $\gamma : [0, 1] \to \mathbb{R}^n$ such that $\gamma(0) = a$ and $\gamma(1) = b$. The values in the interval $[0, 1]$ can be thought of as the time at which you are at a certain point along the path. The continuity of the path γ should not be confused with the continuity of f along this path.) If f is continuous, then along a path from a to b it takes all values between $f(a)$ and $f(b)$. However there is not only one such path; if we choose a point c not on our original path, then the property of path connectedness means there exists a path from a to c and from c to b, so we have generated a new path from a to b. Again, continuity means that f must take all the values between $f(a)$ and $f(b)$ on this new path.

Therefore more than one point in $\mathbb{R}^n$ is mapped to each of the values in $\mathbb{R}$ between $f(a)$ and $f(b)$ and this contradicts f being injective.

References

Bickel, W.K., Koffarnus, M.N., Moody, L., Wilson, A.G., 2014. The behavioral- and neuro-economic process of temporal discounting: a candidate behavioral marker of addiction. Neuropharmacology 76, 518–527. https://doi.org/10.1016/j.neuropharm.2013.06.013.

Borges, A.M., Kuang, J., Milhorn, H., Yi, R., 2016. An alternative approach to calculating Area-Under-the-Curve (AUC) in delay discounting research. J. Exp. Anal. Behav. 106, 145–155. https://doi.org/10.1002/jeab.219.

Cajueiro, D.O., 2006. A note on the relevance of the q-exponential function in the context of intertemporal choices. Phys. A Stat. Mech. Appl. 364, 385–388. https://doi.org/10.1016/j.physa.2005.08.056.

Charlton, S.R., Fantino, E., Gossett, B.D., 2013. Hyperbolic discounting of delayed social interaction. Learn. Behav. 41 (2), 159–167. https://doi.org/10.3758/s13420-012-0093-3. 2013.

Cheng, K., Peña, J., Porter, M.A., Irwin, J.D., 2002. Self-control in honeybees. Psychol. Bull. Rev. 9 (2), 259–263. https://doi.org/10.3758/BF03196280.

Doyle, J.R., 2013. Survey of time preference, delay discounting models. Judgm. Decis. Mak. 8 (2), 116–135.

Fawcett, T.W., McNamara, J.M., Houston, A.I., 2012. When is it adaptive to be patient? A general framework for evaluating delayed rewards. Behav. Proc. 89 (2), 128–136. https://doi.org/10.1016/j.beproc.2011.08.015.

Gilroy, S.P., Hantula, A., 2018. Discounting model selection with area-based measures: a case for numerical integration. J. Exp. Anal. Behav. 109, 433–449. https://doi.org/10.1002/jeab.318.

Green, L., Myerson, J., 2004. A discounting framework for choice with delayed and probabilistic rewards. Psychol. Bull. 130 (5), 769–792. https://doi.org/10.1037/0033-2909.130.5.769.

Green, L., Fry, A.F., Myerson, J., 1994. Discounting of delayed rewards: a life-span comparison. Psychol. Sci. 5 (1), 33–36. https://doi.org/10.1111/j.1467-9280.1994.tb00610.x.

Green, L., Myerson, J., Calvert, A.L., 2010. Pigeons' discounting of probabilistic and delayed reinforcers. J. Exp. Anal. Behav. 94 (2), 113–123. https://doi.org/10.1901/jeab.2010.94-113.

Holt, D.D., Newquist, M.H., Smits, R.R., Tiry, A.M., 2014. Discounting of food, sex, and money. Psychon. Bull. Rev. 21, 794–802. https://doi.org/10.3758/s13423-013-0557-2.

Kagel, J.H., Green, L., Caraco, T., 1986. When foragers discount the future: constraint or adaptation? Anim. Behav. 34, 271–283. https://doi.org/10.1016/0003-3472(86)90032-1.

Killeen, P.R., 2009. An additive-utility model of delay discounting. Psychol. Rev. 116 (3), 602–610. https://doi.org/10.1037/a0016414.

Killeen, P.R., 2015. Models of ADHD: five ways smaller sooner is better. J. Neurosci. Methods 252 (S1), 2–13. https://doi.org/10.1016/j.jneumeth.2015.01.011.

Laibson, D., 1997. Golden eggs and hyperbolic discounting. Q. J. Econ. 112, 443–477. https://doi.org/10.1162/003355397555253.

Loewenstein, G., Prelec, D., 1992. Anomalies in intertemporal choice: evidence and an interpretation. Q. J. Econ. 107, 573–597. https://doi.org/10.2307/2118482.

Mazur, J.E., Biondi, D.R., 2009. Delay-amount tradeoffs in choices by pigeons and rats: hyperbolic versus exponential discounting. J. Exp. Anal. Behav. 91 (2), 197–211. https://doi.org/10.1901/jeab.2009.91-197.

Mazur, J.E., 1987. An adjusting procedure for studying delayed reinforcement. In: In: Commons, M.L., Mazur, J.E., Nevin, J.A. (Eds.), Quantitative Analyses of Behaviour Vol. 5. Erlbaum, Hillsdale, NJ, pp. 55–73.

McKerchar, T.L., Green, L., Myerson, J., 2010. On the scaling interpretation of exponents in hyperboloid models of delay and probability discounting. Behav. Proc. 84 (1), 440–444. https://doi.org/10.1016/j.beproc.2010.01.003.

Mitchell, S.H., Wilson, V.B., Karalunas, S.L., 2015. Comparing hyperbolic, delay-amount sensitivity and present-bias models of delay discounting. Behav. Proc. 114 (SI), 52–62. https://doi.org/10.1016/j.beproc.2015.03.006.

Myerson, J., Green, L., 1995. Discounting of delayed rewards: models of individual choice. J. Exp. Anal. Behav. 64 (3), 263–276. https://doi.org/10.1901/jeab.1995.64-263.

Myerson, J., Green, L., Warusawitharana, M., 2001. Area under the curve as a measure of discounting. J. Exp. Anal. Behav. 76 (2), 235–243. https://doi.org/10.1901/jeab.2001.76-235.

Prelec, D., 2004. Decreasing impatience: a criterion for non-stationary time preference and "Hyperbolic" discounting. Scand. J. Econ. 106, 511–532. https://doi.org/10.1111/j.1467-9442.2004.00380.x.

Reynolds, B., 2006. A review of delay-discounting research with humans: relations to drug use and gambling. Behav. Pharm. 17 (8), 651–667. https://doi.org/10.1097/FBP.0b013e3280115f99.

Reynolds, B., Richards, J.B., Horn, K., Karraker, K., 2004. Delay discounting and probability discounting as related to cigarette smoking status in adults. Behav. Proc. 65 (1), 35–42. https://doi.org/10.1016/S0376-6357(03)00109-8.

Rodriguez, M.L., Logue, A.W., 1988. Adjusting delay to reinforcement: comparing choice in pigeons and humans. J. Exp. Psychol.: Anim. Behav. Proc. 14 (1), 105–117. https://doi.org/10.1037/0097-7403.14.1.105.

Samuelson, P.A., 1937. A note on measurement of utility. Virgil Madgearu Rev. Econ. Stud. Res. 4 (2), 155–161 1937.

Sozou, P.D., 1988. On hyperbolic discounting and uncertain hazard rates. Proc. R. Soc. B: Biol. Sci. 265 (1409), 2015–2020. https://doi.org/10.1098/rspb.1998.0534.

Stephens, D.W., Kerr, B., Fernández-Juricic, E., 2004. Impulsiveness without discounting: the ecological rationality hypothesis. Proc. R. Soc. B: Biol. Sci. 271 (1556), 2459–2465. https://doi.org/10.1098/rspb.2004.2871.

Stevens, J.R., 2014. Evolutionary pressures on primate intertemporal choice. Proc. R. Soc. B: Biol. Sci. 281 (1786), 20140499. https://doi.org/10.1098/rspb.2014.0499.

Takahashi, T., Oono, H., Radford, M.H.B., 2008. Psychophysics of time perception and intertemporal choice models. Physica A 387, 2066–2074. https://doi.org/10.1016/j.physa.2007.11.047.

Vanderveldt, A., Oliveira, L., Green, L., 2016. Delay discounting: pigeon, rat, human-does it matter? J. Exp. Psychol.: Anim. Learn. Cogit. 42 (2). https://doi.org/10.1037/xan0000097. 141 162.

Weatherly, J.N., 2014. On several factors that control rates of discounting. Behav. Proc. 104 (SI), 84–90. https://doi.org/10.1016/j.beproc.2014.01.020.

Yoon, J.H., De La Garza, R., Newton, T.F., Suchting, R., Weaver, M.T., Brown, G.S., Omar, Y., Haliwa, I., 2017. A comparison of Mazur's k and area under the curve for describing steep discounters. Psychol. Rec. 67, 355–363. https://doi.org/10.1007/s40732-017-0220-9.