# The case of muddled units in temporal discounting

Benjamin T. Vincent*[1] and Neil Stewart†[2]

[1]Division of Psychology, School of Social Sciences, University of Dundee, UK

[2]Warwick Business School, University of Warwick, UK

January 30, 2019

**Abstract**

While parameters are crucial components of cognitive models, relatively little importance has been given to their units. We show that this has lead to some parameters to be contaminated, introducing an artifactual correlation between them. We also show that this has led to the illegal comparison of parameters with different units of measurement – this may invalidate parameter comparisons across participants, conditions, groups, or studies. We demonstrate that this problem affects two related models: Stevens' Power Law and Rachlin's delay discounting model. We show that it may even affect models which superficially avoid the incompatible units problem, such as hyperbolic discounting. We present simulation results to demonstrate the extent of the issues caused by the muddled units problem. We offer solutions in order to avoid the problem in the future or to aid in re-interpreting existing datasets.

*b.t.vincent@dundee.ac.uk

†Neil Stewart was supported by the Economic and Social Research Council (ES/P008976/1, ES/N018192/1) and the Leverhulme Trust (RP2012-V-022)

# 1 The incomparable units problem

## 1.1 A demonstration with Stevens' Power Law

Stevens' 1975 power law is wrong. It is wrong because it is parameterised incorrectly—so that the units of one parameter are muddled with the value of another parameter. The concepts of dimensional invariance (Fourier, 1822; Krantz, 1972; Maxwell, 1891; Stewart, Scheibehenne, & Pachur, 2018) and meaningfulness (Falmagne, 1985; Falmagne & Narens, 1983) can be applied in mathematical models of psychology, by considering the units of psychological parameters. In Stevens' power law, the psychological magnitude $\psi(I)$ of physical magnitude or intensity $I$ is given by

$$\psi(I) = \lambda I^a. \tag{1}$$

Let's consider the perception of visual length, and particularly of lines. If length is measured in the International System of Units (SI) of metres (m), then $I$ has units of m. This means that $I^a$ has units of $\mathrm{m}^a$. If $\psi(I)$ is to be a unitless psychological scale, or at least a scale with its own psychological units, then it must be free of the physical units. This means that $\lambda$ must have the reciprocal units of $I^a$ so that the units cancel out. Thus $\lambda$ must have units of $\frac{1}{\mathrm{m}^a} = \mathrm{m}^{-a}$. The key problem here is that $\lambda$ has units which depend upon $a$. Stewart et al. (2018) have shown, for similar models, how this will lead to estimates of $\lambda$ and $a$ that are highly correlated, and that it is illegal to compare $\lambda$ values across, for example, individuals with different values of $a$. When $a$ differs, the units of $\lambda$ differ, and comparing magnitudes with different units is not permitted.

We have highlighted a problem with the units of $\lambda$ in Stevens' law, though $\lambda$ is often not the parameter of core interest in psychophysical modelling. Instead it is the exponent $a$ that is of primary consideration. The exponent $a$ has been tabulated, in reviews of the experimental literature, for more than 20 physical continua, including loudness, brightness, length and area, tastes and smells, temperature, pressure, texture, vibration, weight, duration, and even electric shocks. The $\lambda$ parameter is of lesser theoretical interest, because it is determined, in part, by the properties of the judgement scale for $\psi$. For example, $\lambda$ will differ depending on whether the scale runs from 0–10 or 0–100. But later in this paper, when we are considering temporal discounting, the analogue of $\lambda$ is of core theoretical interest.

## 1.2 Fechner's Law is dimensionally correct

In contrast to Stevens' law, Fechner's 1860/1966 law

$$\psi(I) = \lambda \ln \left( \frac{I}{I_0} \right) \tag{2}$$

is a meaningful and dimensionally invariant law. The ratio of the physical quantity $I$ and the threshold physical quantity $I_0$ (at which the perception $\psi(I)$ is zero) is unitless, because the units of $I$ and $I_0$ cancel in the ratio. Logarithms are also unitless—they are the power to which the base of the log must be raised, and powers are unitless real numbers. This means that $\lambda$ need have only whatever unit required to match the scale of $\psi(I)$ is measured in.

### 1.3 Fixing the units in Stevens' Law

We propose the modified Stevens' Power Law:

$$\psi(I) = (\gamma I)^a \tag{3}$$

where $\gamma^a = \lambda$ or $\gamma = \lambda^{1/a}$. Note that $a$ is still a dimensionless quantity, but $\gamma$ is in inverse units of $I$, such as $\frac{1}{\text{meters}}$ for length, $\frac{1}{\text{meters}^2}$ for area, $\frac{1}{\text{days}}$ for duration, for example. Also, now $\gamma$ is appealingly independent of $a$. Comparisons between $a$ parameters and between $\gamma$ parameters are allowable in dimensional analysis; but, as we explain above, comparisons of $\lambda$ are not. This fix—moving the constant inside the power—is described in Stewart et al. (2018).

To summarise, studies which have compared values of $\lambda$ across or within individuals are wrong because comparing quantities which are not in the same units is not permitted. It is like comparing 7 metres with 8 seconds and asking which is bigger—nonsensical. Our strong suggestion would be that for such studies, the $\lambda$ parameter is transformed into $\gamma$ and the results be reinterpreted based on these values. The extent to which this may be a problem, and the feasibility of this suggestion, is explored in the remainder of the paper.

## 2 A case study in time perception and temporal discounting

We illustrate the incompatible units problem in the domain of hot affective emotional states. Multiple studies have found that when people undergo a hot affective state manipulation (e.g. by viewing sexually arousing stimuli) then their present bias increases (Ariely & Loewenstein, 2006; Lemley, Asmussen, & Reed, 2015; Van den Bergh & Dewitte, 2008; Wilson & Daly, 2004). That is, they discount future rewards to a greater extent, such that preferences shift toward smaller but sooner rewards compared to larger but later rewards. But what are the cognitive processes that are responsible? Are people's temporal preferences altered in *hot states*[1] because of changes in discount rates, or because of changes in subjective time perception, or some combination of the two?

Regardless of the precise measure used, results of the above studies were interpreted as changes in time discounting caused by the experimental hot state manipulation. Caution must be expressed however as these results could have also been driven partly or wholly by changes in subjective time perception. In an excellent series of studies, Kim and Zauberman (2013) found a similar increase in present bias caused by hot state manipulations, but, because they also measured subjective time perception, were able to conclude that this change in present bias is driven by changes in subjective time perception rather than changes in discount rates. They measured the relationships between subjective time perception and inter-temporal choice for money under control and hot states. In Study 1 they showed that male participant's subjective time perception was altered by viewing pictures of female lingerie models. They used a procedure to estimate perceived durations from objectively stated durations according to Stevens' Power Law (see Equation 1,

[1]Hot state does not have a technical definition, but it refers to a joint physiological and phenomenological state of affairs, often induced by sexual arousal.

where $I$ is measuring the duration). Participants indicated subjective time by adjusting the length of a line on the computer screen, relative to a reference duration of 1 month corresponding to 32.71mm), therefore the units of $I$ was in mm. This resulted in group level fits of $\psi(I) = 0.998\ I^{0.68}$ for the hot condition and $\psi(I) = 0.610\ I^{0.73}$ for the control condition.

One conclusion drawn by the researchers was that participant's subjective time perception was sublinear[2]— based upon the point estimates of the exponents both being below 1, and a non-significant difference between these exponent parameter values. Because the exponent is unitless, this conclusion is not affected by any units problems.

A second conclusion was that perceived time increased in the hot state, such that a fixed duration was perceived as longer. The scaling parameter ($\lambda$ in our Equation 1) increased from the control to experimental condition. This comparison of $\lambda$ values is not valid, undermining the conclusion that changes in subjective time perception were responsible. As we have seen, the $\lambda$ parameters are in units which are also affected by the exponent. Specifically, the group level constant for the hot condition is $\lambda = 0.998\text{mm}^{-0.68}$ (i.e. units of $\text{mm}^{-0.68}$), and $\lambda = 0.61\text{mm}^{-0.73}$ (i.e. units of $\text{mm}^{-0.73}$) for the control condition. These constants are in different units and therefore cannot be compared. Likewise conducting t-tests or ANOVAs on $\lambda$ values for participant level fits is also illegal, as they are all units of $\text{mm}^a$ where $a$ is different for each participant. Instead, $\lambda$ should be transformed to $\gamma$ (for each participant, which requires the $\lambda$ and $a$ values for each participant). So even though there were non-significant differences in $a$ in the control and hot state groups, the $a$ values will have been different for each participant, and we do not know whether group differences would have been significantly different (at a threshold p–level) when comparing the fitted $\gamma$ values across control and hot conditions. The best we can do without the participant-level data or parameter estimates is to compute $\gamma$ at the group level. This results in $\gamma_{\text{hot}} = 0.998^{(1/0.68)} = 0.997$ and $\gamma_{\text{control}} = 0.610^{(1/0.73)} = 0.508$, but these are just point estimates so we have no way to verify if their are statistically significant differences between $\gamma_{\text{hot}}$ and $\gamma_{\text{control}}$. This is also a highly dubious operation—as we will discuss in more detail later in the paper, transforming $(\lambda, a)$ to $(\gamma, a)$ parameters for group level summary statistics is invalid because $a$ will vary across participants. And so we are unfortunately left with uncertainty about the effect of the hot state manipulations in this experiment on subjective time perception.

Our intention is to point out that we simply cannot make claims based on the comparison of quantities with different units. We do not intend to cast doubt upon the role of subjective time perception in hot state manipulations. Indeed, the basic claim seems reasonable given the findings of a previous study (Zauberman, Kim, Malkoc, & Bettman, 2009) which modelled subjective time with the Webber-Fechner Law (which bypasses these concerns) rather than Stevens' Power Law.

# 3 Implications for delay discounting

This units problem is not just restricted to Stevens' Power Law and magnitude estimation. In the remainder of the paper we outline how this problem filters through into the temporal discounting literature in multiple ways. First we demonstrate that Rachlin's

[2]Technically this should be referred to as a monotonically increasing but decelerating function.

popular discount function suffers from the units problem and we propose a fix. Second, we demonstrate that the popular hyperbolic discount function may also suffer from this problem despite superficially escaping the incompatible units problem.

Here, we illustrate these problems in the domain of inter-temporal choice (also known as delay discounting). The core phenomena of interest here is how agents make trade-offs between the magnitude of a gain (or a loss) and the immediacy of that. For example, the present subjective value of £100 now is greater than £100 in 40 years because future rewards are discounted by some fraction. This could be driven by many reasons including inflation expectations, risk of future rewards not materialising, opportunity costs etc. But how exactly are decisions made about outcomes which occur at different points in time? The general utility-based approach to answering this is to propose that our present subjective value $V$ of a reward $R$ at a given delay $D$ is given by

$$V(R, D) = u(R) \cdot f(D) \tag{4}$$

where $u(R)$ is a utility function relating objective rewards $R$ to subjective values, and $f(D)$ is a discount function which modulates our subjective values as a function of delay. In the discounting literature it is common to assume a linear subjective value function, i.e. the identify function $u(R) = R$ in which case $u(R)$ is in units of pounds, euros, dollars, etc. The focus is instead upon the form of the discount function $f(D)$ which we will explore below.

There are a range of popular discount functions which do not suffer from these unit comparison problems:

- Exponential discounting (Samuelson, 1937) where $f(D) = \exp(-kD)$. $D$ is in time units (e.g., days). Here $k$ is in inverse time units of this (e.g., days$^{-1}$).
- Constant sensitivity function (Ebert, Prelec, & Prelec, 2007) where $f(D) = \exp\left(-(aD)^b\right)$. Here $a$ is in inverse time units, and $b$ is dimensionless.
- The Myerson and Green (1995) hyperboloid where $f(D) = 1/(1 + kD)^s$. Here $k$ is in inverse time units and $s$ is dimensionless.
- Double exponential (McClure, Ericson, Laibson, Loewenstein, & Cohen, 2007) where $f(D) = \omega \exp(-k_1 D) + (1 - \omega) \exp(-k_2 D)$. Although this model might more accurately be called a 'mixture of exponentials'. The mixture component $\omega$ is dimensionless and $k_1$ and $k_2$ are in inverse time units.

Nevertheless even if a discount function's parameters does not suffer from the incompatible units problem, when comparing parameter values (such as $k$) across participants, conditions, or studies, it is important to ensure that they are all in the same units. For example, if the unit of time is days in one paper and years in another, then the $k$s cannot be compared because one will have units of days$^{-1}$ and the other will have units of years$^{-1}$. It is only allowable to compare $k$ values in the same units. Because discount rates vary drastically across species (half lives $(1/k)$ range from seconds to years or decades; Vanderveldt, Oliveira, & Green, 2016) this mistake could easily be made in a meta analysis, for example.

### 3.1 Implications for Rachlin's delay discounting function

Some discount functions suffer from a class of problem where fitted parameters are unknowingly in different units and therefore are not comparable. This problem affects two prominent discount functions. The first is exponential discounting of subjective (i.e. Stevens' power law scaled) time (Takahashi, Oono, & Radford, 2008)

$$f(D) = \exp\left(-kD^s\right) \tag{5}$$

where $\theta = [k, s]$, and the second is the prominent Rachlin (2006) hyperboloid model[3] equating to hyperbolic discounting of (Stevens' power law scaled) subjective time,

$$f(D) = 1/(1 + kD^s). \tag{6}$$

We proceed to illustrate the issues with the Rachlin discount function given its frequent use in the discounting literature, but the issues we highlight also affect Equation 5 (see Appendix A).

$S$ is a power — a unitless real number. This means that comparison of fitted values or posterior distributions of $s$ across participants or studies is allowable under dimensional analysis, as $s$ has the same units (in this case, no units). Of course, there may be other issues around parameter trade offs that make this comparison hard.

However a units problem does arise with the $k$ parameter. Because $D$ is in units of days (for example) then this means that $D^s$ is in units of days$^s$. This follows because $f(D)$ is a unitless fraction, which means the right hand side of Equation 6 must also be unitless. As the numerator 1 is unitless, the denominator $1 + kD^s$ must also be unitless. The $(kD^s)$ term must be unitless, because it is added to 1, which has no units, and one can only add quantities with the same units. This means that $k$ must have units of 1/days$^s$ to cancel with the units of $D^s$. Given that $s$ will vary across participants, then you cannot compare $k$ across participants as they are all in different units. For example, when $s = 1$ then $k$ has units of 1/days but when $s = \frac{1}{2}$ then $k$ has units of $\text{days}^{-\frac{1}{2}} = 1/\sqrt{\text{days}}$.

Based upon our proposed fix to Stevens' Power Law (Equation 3), we propose the modified-Rachlin discounting function:

$$f(D) = 1/(1 + (\kappa D)^s) \tag{7}$$

where $\theta = [\kappa, s]$, and $\kappa^s = k$ or $\kappa = k^{1/s}$. Note that $s$ is still a dimensionless quantity, but $\kappa$ is in units of days$^{-1}$, which is appealingly independent of $s$. Comparisons between $s$ parameters and between $\kappa$ parameters are allowable in dimensional analysis; but, as we explain above, comparisons of $k$ parameters are not. Our fixed-Rachlin function is

[3] We illustrate how Rachlin's discount function is derived from hyperbolic discounting of subjective time perception according to Stevens' Power Law. First we start with hyperbolic discounting of subjective delay, $f(D) = 1/(1 + k\psi(D))$, where $\psi(D)$ is subjective time delay. If we substitute in Equation 1 we obtain $f(D) = 1/(1 + k'\lambda D^s)$. We see that Rachlin's $k$ (Equation 6) is the product of actual discount rates $k'$ and the subjective time scaling parameter, $k = k'\lambda$. To be clear, the $k$ parameter in the Rachlin model conflates the discount rate $(k')$ and a subjective time scaling parameter $(\lambda)$ and it is not possible to identify their values from discounting data alone. Instead, this would require both inter-temporal choice experiments *and* subjective time perception measures for each participant.

still able to be interpreted in terms of hyperbolic discounting of subjective time because $1/(1 + (\kappa D)^s) = 1/(1 + kD^s)$.

The modified-Rachlin function has a number of advantages. First and most obviously, it now becomes legitimate to compare discounting behaviours (using $\kappa$) across participants with different subjective time perception (as specified by $s$). This is a significant advantage because previous comparisons of $k$ across participants or studies will in fact be invalid because they are contaminated by varying values of $s$.

Second, the $\kappa$ parameter is now conveniently always equal to the inverse half life (the delay at which a reward is equal to half its objective value) regardless of the value of $s$. (See this by noting that the delay $D_{\text{half}}$ at which the value of the reward is halved can be substituted into Equation 7 to give $\frac{1}{2} = \frac{1}{1+(\kappa D_{\text{half}})^s}$ which means $2 = 1 + (\kappa D_{half})^s$ and $1 = (\kappa \times D_{half})^s$ and $1^{1/s} = \kappa \times D_{half}$ and thus $1 = \kappa \times D_{half}$.) This was an appealing property of the discount rate in the hyperbolic discount function (Mazur, 1987), but which was lost in the original Rachlin function.

Third, parameter estimation of $(\kappa, s)$ will be improved and more robust. The top panel in Figure 1 shows some simulated data from a delay discounting experiment. The bottom panels show likelihood surfaces for the parameters of the Rachlin and modified-Rachlin discount functions for a single simulated experiment. There is a very clear parameter trade-off which occurs with Rachlin's discount function, as seen by the negatively sloped ridge in the likelihood surface (Figure 1 bottom left). These parameter trade-offs are not noticeable in methods which estimate only point estimate parameters (e.g. Gilroy, Franck, & Hantula, 2017), only those which estimate the full likelihood or posterior surface over parameter space (such as Vincent, 2016). However, parameter correlations *across participants* have been noted in modelling work (such as Peters, Miedl, & Büchel, 2012). This disappears in the likelihood surface of the modified Rachlin function (Figure 1 bottom right). This is especially appealing in the context of Bayesian parameter estimation—the highly anti-correlated structure of the likelihood surface in Figure 1 (bottom left) could pose challenges for some sampling algorithms to accurately estimate the true posterior distribution (see Stewart et al., 2018).

Fourth, a direct consequence of this ridge in the likelihood surface is that errors in estimating the maximum likelihood estimates of true $(k, s)$ parameters will contain undesirable correlational structure. Figure 2 (top) shows the distribution of maximum likelihood estimates from a parameter recovery simulation—200 simulated experiments were run with stochastic choices and maximum likelihood estimation of an observer with fixed parameters. Because the 200 observers were identical, with fixed $(k, s)$, scatter of the estimates away from the true $(k, s)$ crosshairs represents error in parameter recovery which is caused by the stochasticity of the binary responses to the choices. The result is that errors in the maximum likelihood parameters are undesirably correlated. Figure 2b shows that the modified-Rachlin function fixes this problem, we no longer have this parameter trade-off in the maximum likelihood estimates.

We propose that existing research with $(k, s)$ estimated from the Rachlin function can, and should, be transformed to our modified parameters $(\kappa, s)$ so that comparison between participants and studies become valid and relevant. This transformation is a valid approach—we found that a maximum likelihood procedure to estimate $(k, s)$ are accurate, and map on precisely (after the $\kappa = k^{1/s}$ transformation) to parameter estimates of $(\kappa, s)$ directly (see Figure 2c, d). The correlation coefficient between $s$ estimated from

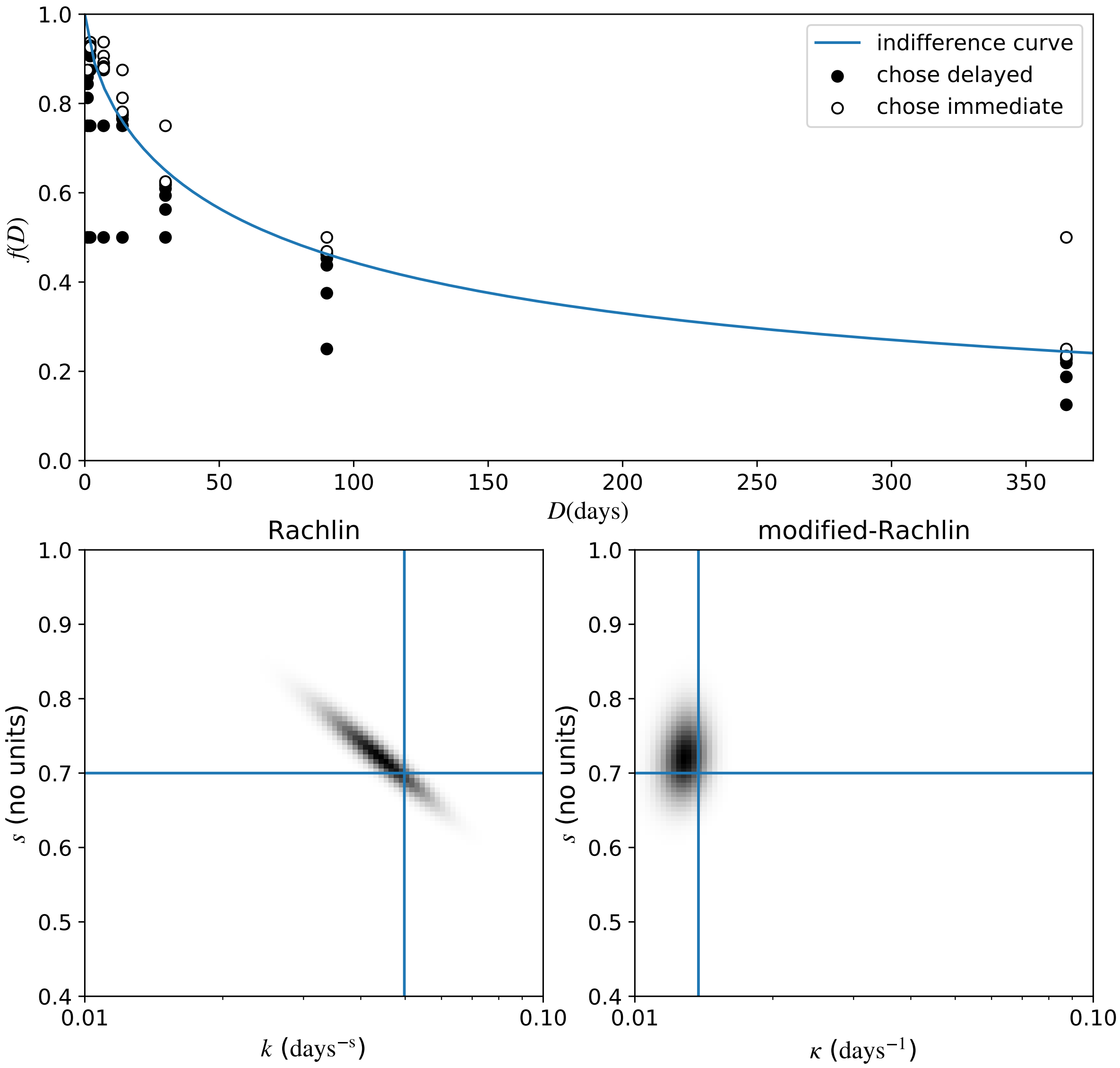


Figure 1: Likelihood surfaces of the simulated data (top) where ($k = \exp(-3), s = 0.7$) for the Rachlin model (bottom left; Equation 6) and modified Rachlin model (bottom right; Equation 7). Experimental designs (rewards and delays) were generated using the adaptive procedure described by Frye et al. (2016). See Appendix B for simulation details. The code to generate this figure is available at https://osf.io/uscmd/.

the Rachlin and modified Rachlin function was virtually equal to 1, within 5-6 decimal places. This was also the case for the correlation coefficient between $\kappa$ (transformed from the $k$ recovered from the Rachlin function) and the $\kappa$ recovered from the modified Rachlin function. This is good news—assuming rigorous maximum likelihood estimation procedures were followed, we do not believe that estimation with the Rachlin function would introduce systematic errors in the actual parameters estimated, just that the $k$ parameter is contaminated as described above. If there is doubt however about the accuracy of past maximum likelihood procedures, the most prudent approach would be to estimate $(\kappa, s)$ directly from the archived raw intertemporal choice data.

To probe this mapping between $(k, s)$ and $(\kappa, s)$ further, we repeated the parameter

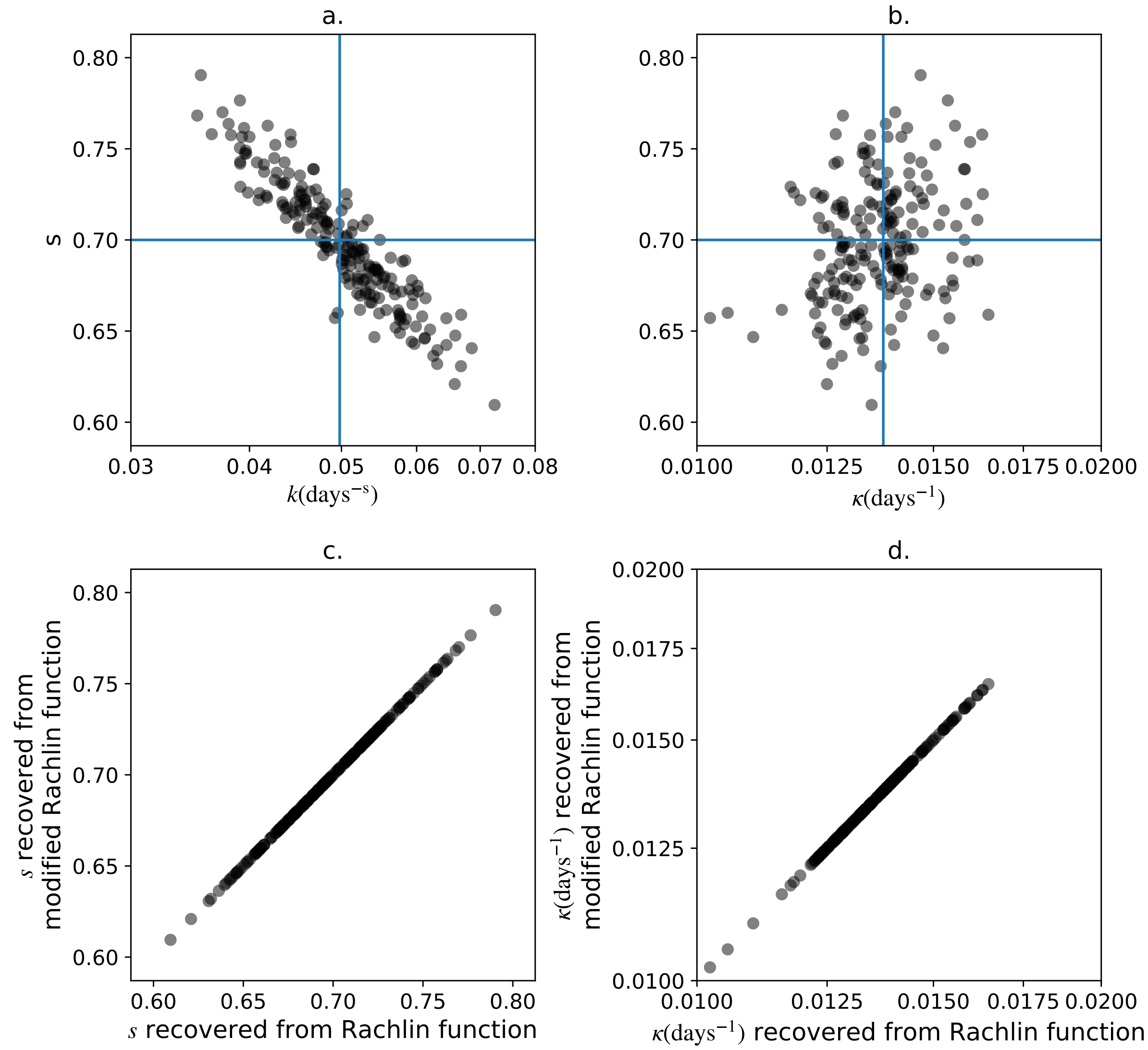


Figure 2: Robustness of ML estimation to stochastic response data. A set of 200 experiments (akin to that shown in Figure 1) were conducted on a simulated participant with fixed true parameter values (shown by crosshairs; $k = \exp(-3)$, $s = 0.7$, thus $\kappa = \exp(-3)^{1/0.7}$) and stochastic responses. Maximum likelihood estimation was used to estimate parameters for the Rachlin (panel a) and modified Rachlin (panel b) functions. Points represent the maximum likelihood parameters for each simulated dataset. The parameter estimation procedure was found to be robust—conducting MLE on data using the Rachlin or the modified Rachlin functions will result in identical maximum likelihood estimates, see main text for details. This was demonstrated by near perfect correlations between $s$ from both equations (panel c) being almost exactly 1, and likewise for $k$ transformed to $\kappa$, and $\kappa$ (panel d). See Appendix B for simulation details. The code to generate this figure is available at `https://osf.io/uscmd/`.

recovery approach (from Figure 2) but extended this for multiple true parameter values in Figure 3. Figure 3 shows true parameter values chosen from a grid over $(\kappa, s)$ space, along with recovered parameter values using maximum likelihood estimation. The results are in line with the intuition from Figure 2, that there is a straight 1-to-1 mapping between $(k, s)$ and $(\kappa, s)$ parameter spaces. That is, it should be possible to accurately map to $(\kappa, s)$ directly from existing estimates of $(k, s)$ obtained from Rachlin's function. There are two concerns which remain however.

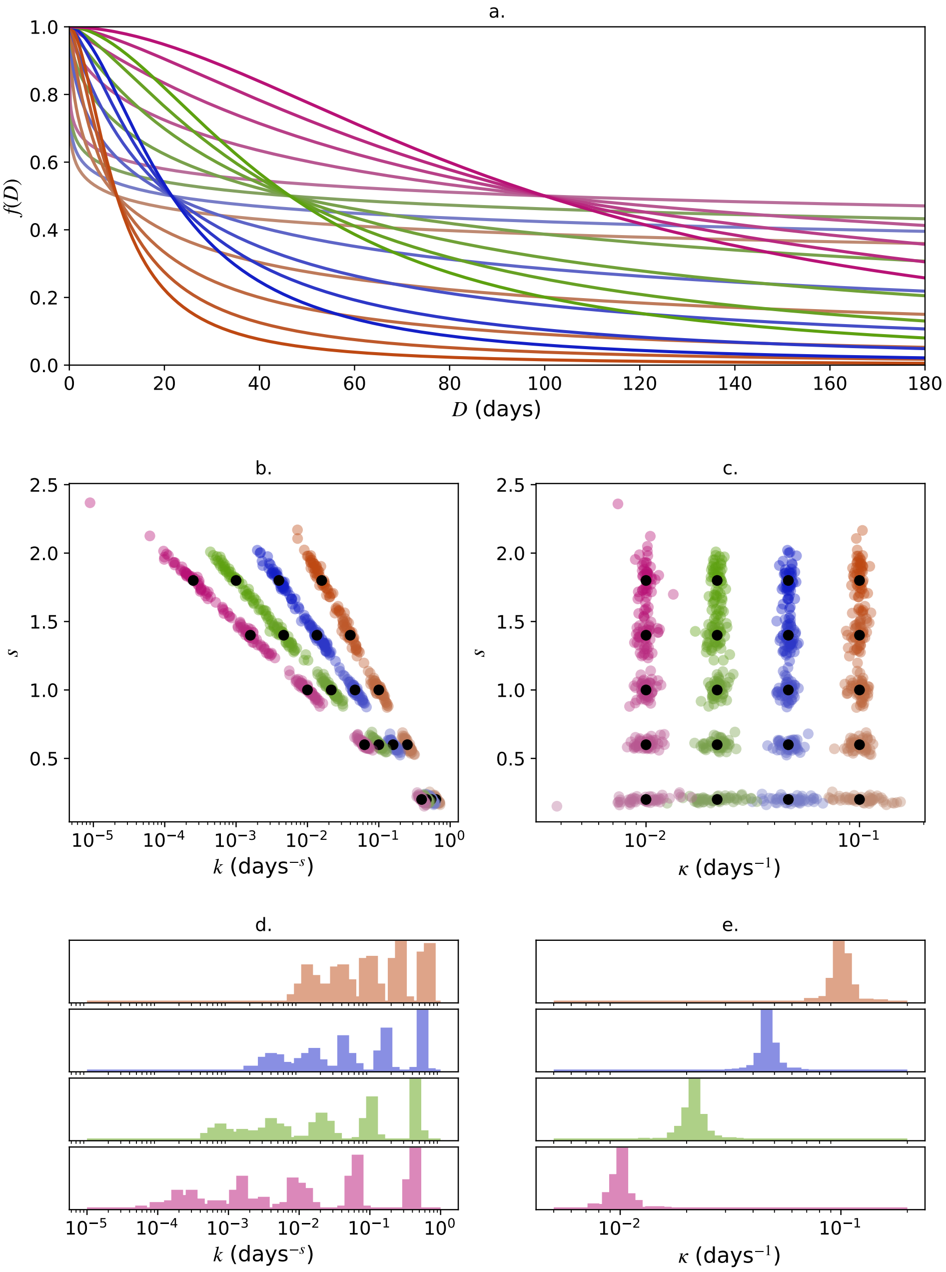


Figure 3: Comparing the Rachlin $(k, s)$ and our modified Rachlin $(\kappa, s)$ parameter spaces. Panel (a) shows a series of discount functions with parameter values uniformly spaced in modified-Rachlin parameter space $(\kappa, s)$ (black points in (c)). Different colours represent different true $\kappa$ values, and saturation represents different $s$ values. Corresponding true parameter values in the Rachlin $(k, s)$ space are shown in panel (b) – comparing panels (b) and (c) shows the nature of the mapping between $(k, s)$ and $(\kappa, s)$. Coloured points in (b) & (c) correspond to inferred parameter values based on simulated experiments. We can see that the inferred parameters are centred on the true parameter values (black points). The change in estimation precision over the parameter space is caused by the ability of the simulated adaptive delay discounting procedure to constrain the plausible parameter values. Panels (d) and (e) show histograms of inferred $k$ and $\kappa$ values respectively for simulated participants with fixed true $\kappa$ values. We can see that the inferred $\kappa$ values are independent of $s$, but the inferred $k$ values are contaminated by $s$, and so group level inferences about $k$ will be skewed by participants with varying $s$ values. See Appendix B for simulation details. The code to generate this figure is available at https://osf.io/uscmd/.

The first concern is that when past results are re-examined and $k, s$ is transformed into $\kappa, s$, this may well merit reinterpretation of existing findings in the literature. For example, differences between $k$ between groups or conditions could have been interpreted (wrongly) as relevant differences in discount rates between participants, groups, or conditions. But because $k$ is contaminated by $s$, these differences could have been caused by changes in subjective time perception. This is clear to see in Figure 3(a–c). As stated, $\kappa$ is unrelated to $s$ and corresponds to the inverse half life[4]. It is clear from Figure 3 (b), that increases in $k$ could either be caused by an increase in $k$ while $s$ remains constant, or by $k$ remaining constant, and a decrease in $s$. This should hopefully underscore the importance of revisiting published studies which make theoretical claims about discounting behaviour on the basis of changes in $k$ obtained from Rachlin's hyperboloid function.

Our second concern is that conversion of existing parameter estimates of $(k, s)$ from the Rachlin discount function to our proposed $(\kappa, s)$ parameterisation should be done with care. As we eluded to in the case study above, this conversion is only valid when conducted on participant level parameters, not on group mean or median parameter values. To get a sense of why this is the case, we can see from Figure 3 b that group mean or median values of $k$ will be disproportionately influenced by participants with high $s$ values. This is shown further in the histograms Figure 3 d-e. For example, consider a number of participants with the same discounting behaviour (same values of $\kappa$) but with different subjective time perception (different values of $s$). If we fit with the modified-Rachlin function, then our group level estimate of average $\kappa$ will be accurate and independent of the varying $s$ values. However, if we fit the same set of participants (ie a column of points) with the Rachlin function then our group level estimate of $k$ will be undesirably influenced by the variation of $s$. To summarise, researchers wishing to convert $(k, s)$ parameters into our proposed superior $(\kappa, s)$ parameter space must do so on a participant level, not on a group mean or median level.

## 3.2 Concerns extend to hyperbolic discounting

We also propose that a related problem may befall the accuracy of the discount rate obtained from the classic hyperbolic[5] discount function (Mazur, 1987)

$$f(D) = 1/(1 + kD) \tag{8}$$

where $\theta = k$. Superficially, this function does not suffer from the incompatible units problem—$k$ is simply in units of days$^{-1}$ (or $1/k$ is measured in days) and we can compare $k$ values across participants. Or can we?

Our core concern is that we *may not* be able to draw relevant conclusions about time discounting from changes in discount rates $k$ from the hyperbolic discount function, either within participant changes from condition to condition, or between participants. If this concern is valid, this may have broad consequences which may require revisiting previous results.

[4]Although $\kappa$ is still the product of both a discounting process and the constant term in Steven's Power Law.

[5]Equation 8 is not actually a hyperbolic function, but we will stick with this convention as it has been adopted wholesale in the discounting literature. We refer the reader to (Rasmusen, 2008) for further insights on this point.

The problem revolves around the fact that the Mazur (1987) hyperbolic discount function is a special case of the Rachlin discount function[6] when the exponent $s$ equals 1. So *if* it actually is the case that people hyperbolically discount subjective time ($s \neq 1$) rather than objective time (when $s = 1$), then analyses based upon the hyperbolic discount function will suffer from the omitted-variable problem. Expecting that papers all contain accurate and comparable estimates of $k$ just because the unit of $k$ does not contain $s$ is not a good solution. If $s$ differs across participants but is left out of our model, all of our $k$ values will be *differently* systematically biased, and thus not comparable.

We propose that this is a real problem. Claims about discounting *behaviour* (i.e. choices made in inter-temporal choice tasks) previously attributed to changes in discount *rates*, but may have been partly down to changes in subjective time perception. For example, one of the most highly cited empirical works on delay discounting shows that discounting is higher in current smokers than ex-smokers, than never smokers (Bickel, Odum, & Madden, 1999). However we also have empirical support for atypical time perception in addictive disorders (stimulant-dependent participants *over*-estimate time), with the explicit suggestion that this may influence broad lack of impulse control (Wittmann, Leland, Churan, & Paulus, 2007). On the other end of the spectrum, patients with anorexia nervosa display some of the lowest observed discounting behaviour (Bartholdy et al., 2017; Decker, Figner, & Steinglass, 2015; Steinglass et al., 2012) also *under*-estimate time (Vicario & Felmingham, 2018). So to what extent is this discounting behaviour caused by changes in discount rates versus subjective time perception? We therefore mirror the call of Kim and Zauberman (2018) that research needs to disentangle the relative contributions of subjective time perception and discount rates. Until we have a clearer understanding here, it *may* be premature to claim that changes in discounting behaviour is straightforwardly attributable to changes in discount rates alone.

In order to estimate the extent of the problem, we conducted further simulations. Figure 4 shows the degree of bias in the hyperbolic discount rate $k$ parameter as a function of true $(\kappa, s)$ parameters from the modified Rachlin function. For simulated observers who have linear time perception ($s = 1$; equal to hyperbolic discounting) we can recover discount rates with no systematic bias. Worryingly, we find systematic biases in the estimates of $k$ for observers who do discount subjective time ($s \neq 1$). We see systematic overestimates of $k$ for decelerating time perception ($s < 1$) and systematic underestimates of $k$ for accelerating time perception ($s > 1$). These biases are not subtle—for example, a normalised estimate of +2 means $k_{\text{estimate}}$ is twice the true $\kappa$ value, and a normalised estimate of -0.5 means $k_{\text{estimate}}$ is half of the true $\kappa$ value. These simulations suggest that *if* we accept that subjective time perception influences preferences in inter-temporal choice tasks, and that participant's subjective time perception is uncontrolled for, then claims of changes in discount rates could be conflated with subjective time perception.

One line of evidence suggests that this may be a real problem for conclusions based on hyperbolic discount rates alone. When discount functions are pitted against each other to explain inter-temporal choice behaviour, 2–parameter hyperboloid models (including the Rachlin hyperboloid) fit behavioural data better than the hyperbolic model (McKerchar et al., 2009). That study only assessed goodness of model fit however and did not assess either fit to out-of-sample data (e.g. as in cross validation) or compare model metrics

[6] And of the Myerson and Green (1995) hyperboloid: $f(D) = 1/(1 + kD)^s$

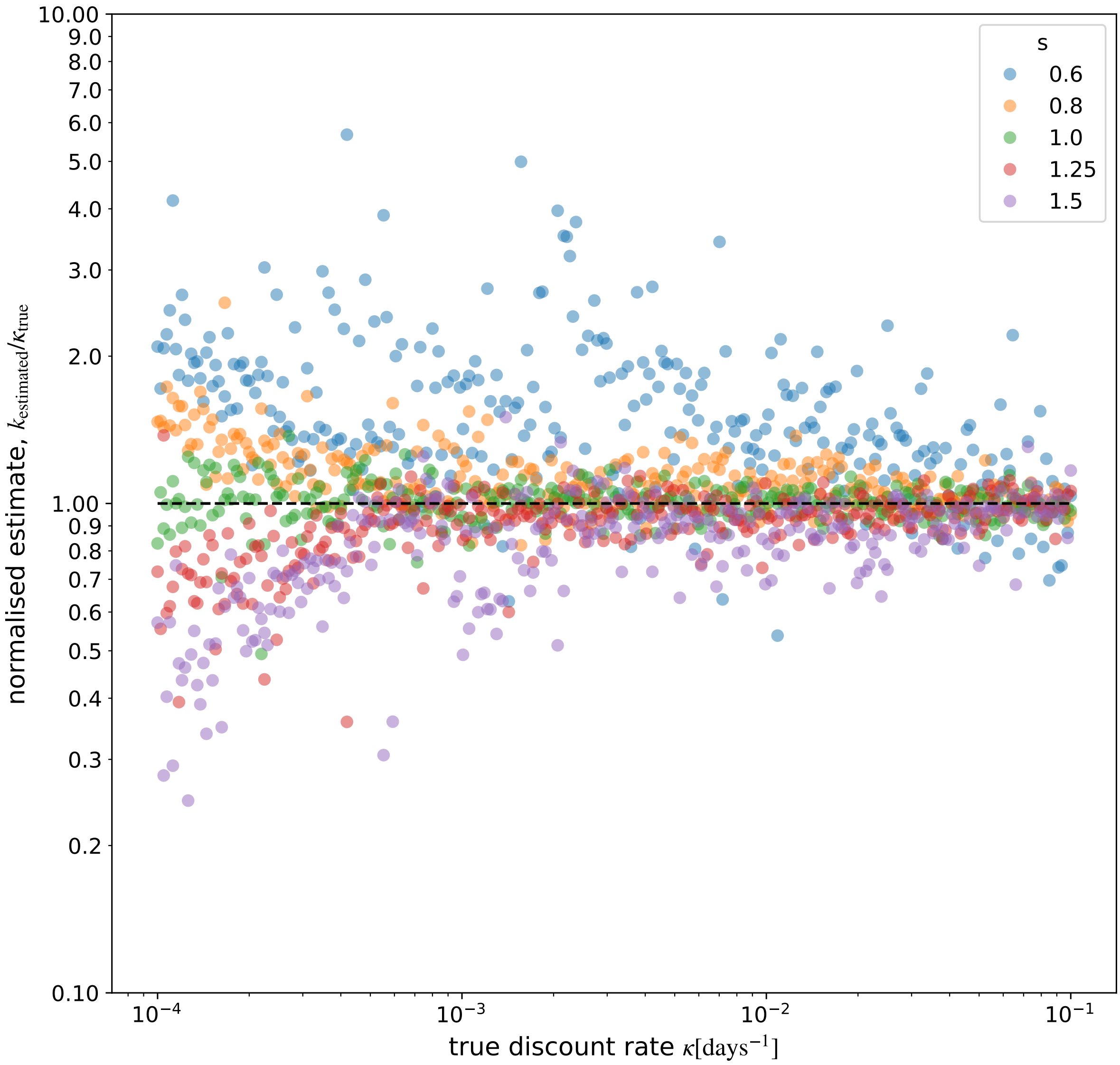


Figure 4: Estimation biases of $k$ from the hyperbolic discount function, based upon inter-temporal choice data for simulated observers who discount according to the modified-Rachlin function. Each point represents a simulated observer with a true $\kappa$ corresponding to the x-axis position and true $s$ value as shown by the colour (see legend). The y-axis shows normalised error $k_{\text{estimated}}/\kappa_{\text{true}}$ such that a value of 1 means no error. We see no systematic bias when observers discount linear time, $s = 1$. But we see systematic underestimates for accelerating time perception ($s > 1$) and systematic overestimates for decelerating time perception ($s < 1$). The code to generate this figure is available at https://osf.io/uscmd/.

which add a penalty for the additional parameter. Franck, Koffarnus, House, and Bickel (2015) did however report BIC (Bayesian Information Criterion; which penalises models with more parameters) scores for fits to individuals. They report the proportion of participants for which a range of models were the most probable to have generated the data as: Rachlin (34.3%), Myerson and Green (27.0%), hyperbolic (18.0%), Laibson (10.8%), exponential (8.1%), and a control model (1.8%). Given the hyperbolic discount function was the most probable model for only 18.0% of participants, this is not strong support for linear subjective time perception. This suggests that in many cases $s \neq 1$

and so estimates of $k$ from the hyperbolic discount function will be contaminated by subjective time perception and not solely reflect discounting processes. This potentially poses a problem for some established findings in the discounting literature based upon participant, group, or condition differences in $k$ values from the hyperbolic discount function.

A second line of evidence can be drawn from the Kim and Zauberman (2013) and Zauberman et al. (2009) studies we have already seen. Taken together these studies provide compelling evidence that subjective time perception is decelerating , and varies across participants and/or experimental conditions. We propose that this may be a serious issue—previous results claiming that inter-temporal choice is affected by discounting may need to be revisited in order to assess the the confound of subjective time perception.

## 3.3 Plotting log parameter values

Before we conclude, we add a note on plotting or reporting transformed parameter values. Researchers often report, or plot, the logarithm of the $k$ parameter from the hyperbolic discounting model, $\log(k)$. But what is $\log(k)$? Following dimensional analysis leads us to the answer. Recall that $k$ is the inverse half life in the hyperbolic discounting model. That is, $k$ is the inverse of the number of days it takes for the present value to be half that of the delayed outcome: If it takes 10 days for the value to drop by half, then $k = 1/10$ per day. Dimensional analysis requires that the number to which a logarithm is applied is unitless. For this reason, a standard reference level is required. The level can be set at any value (e.g., $k_{\text{reference}} = 1$ per day or $k_{\text{reference}} = 3.141592654$ per day, or any value). However if researchers just take the logarithm of the numerical value of $k$ without regard to the units, they have effectively selected a reference value of 1 unit. In this example for $k$, that would be a reference level of $k_{\text{reference}} = 1$ per day.

For example, say $k = 2$ per day and suppose the experimenter is using logarithms to the base 10. $\log_{10}(2) = 0.30103$. But this number 0.30103, which is "$\log_{10}(k)$", really should be written as $\log_{10}\left(\frac{2 \text{ per day}}{1 \text{ per day}}\right) = 0.30103$. This means that $k = 2$ per day is $10^{0.30103} = 2$ times larger than the reference level of $k = 1$ per day. And with natural logarithms $\log_e\left(\frac{2 \text{ per day}}{1 \text{ per day}}\right) = 0.6931472$. This means that $k = 2$ per day is $e^{0.6931472} = 2$ times larger than the reference level of $k = 1$ per day.

Although $\log(k)$ has no units, it should be understood is a logarithm of the number of times larger $k$ is than some reference level $k$. So when reporting or plotting $\log(k)$ or $\log(\kappa)$, one should report the base of the logarithm.[7]

[7]Perhaps the most prominent example of this is the reporting of sound levels on the decibel scale. For example the ear-drum-rupturing 150 dB level of a jet at takeoff at a distance of 25m, often reported as "150 dB" is really "150 dB SPL" or "150 dB sound pressure level". Sound pressure level $L_p = 20\log_{10}\left(\frac{p}{p_0}\right)$, where $p$ is the sound pressure level of the jet measured in any unit of pressure and $p_0$ is the reference level, measured in the same unit. $p_0$ is typically set at 20 $\mu$Pa or 20 micro Pascals (which Wikipedia says the loudness of a mosquito flying 3 m away). This means that the "150" means the jet at 25 m is $10^{150/20} = 31,622,777$ times louder than a mosquito at 3 m.

# 4 Conclusions

The importance of the units of psychological parameters in perceptual and cognitive models is underappreciated. This has led to the muddled units problem where researchers illegally compare parameter values with different units. Further, this causes parameters to become polluted by other parameters, which changes their meaning and interpretation, potentially invalidating research conclusions.

We have illustrated the problem with Stevens' Power Law and proposed a simple-to-implement re-paraterisation which can (a) allow past research to be reevaluated, and (b) avoid the parameter incompatibility problem in the future. We have also show how the problem affects the study of subjective time perception and temporal discounting. Using simulations, we demonstrated this using the Rachlin temporal discounting function, and show how the units of the temporal discounting parameter are contaminated by the units of the subjective time perception parameter, and show how our re-paraterisation avoids this problem.

More subtle, but still deeply problematic, is that the units issue may well afflict models which superficially avoid the parameter identification problem. This was illustrated with the hyperbolic discount function, which is used very often in the delay discounting literature.

We end with a series of important, but potentially alarming, recommendations in relation to cognitive modelling. First, researchers should routinely report the units of their psychological parameters. This will help reduce the possibility of erroneous comparison of parameter estimates in different units. Ideally this reporting will apply to both axis labels of plots in parameter space as well as reporting of parameter values in tables or the main text. For example, reporting that $k = 0.5$ is not sufficient; reporting $k = 0.5\ \text{days}^{-1}$ or $k = 0.5$ per day is preferred. It is also typical to report log transformed $k$ values ($\ln(k)$). While log transformed values have no units, they do have a base which does have units and the base and its units should be reported. The same goes for $\kappa$ or $\ln(\kappa)$ in our modified Rachlin discount function.

Our second recommendation is that cognitive modellers might routinely consider the units of their models during model formulation, in order to avoid the incompatible units problem. For example, in the Introduction, we show that Stevens' Law is not dimensionally correct, but that Fechner's Law is.

Our third recommendation is for the readers of the existing literature. When interpreting existing models, and especially their parameterisation and parameter estimation, readers should have in mind the units of the parameters. If there is a problem, then a solution involving re-parameterisation needs to be found and this may necessitate revisiting the theoretical claims made. For example, in conducting a meta-analysis of loss aversion, Walasek, Mullett, and Stewart (2018) had to obtain the raw choice-level data and re-estimate prospect theory's loss aversion parameter for each participant. For any non-linear re-parameterisation, transforming group level average parameter values will not be sufficient.

We also have recommendations relating to delay discounting and subjective time perception. Our fourth suggestion is that exponential discounting of power-scaled subjective time (Takahashi et al., 2008) should be disfavoured and treated with caution. Instead, focus should be placed on the constant sensitivity function (Ebert et al., 2007, see Appendix A), or on exponential discounting of Weber-Fechner time perception (which is equivalent to the Myerson (2004) hyperboloid (Takahashi et al., 2008)). And finally, fifth, we suggest that researchers should switch to using our modified Rachlin discount function from this point onwards. Published research findings based upon participant, group, or condition differences in the contaminated discount rate parameter ($k$) from Rachlin's discount function may need to be re-examined (Jones & Rachlin, 2009; Kralik & Sampson, 2012; Mazur, 2007; Myerson, Green, & Morris, 2011; Peters et al., 2012; Schneider, Peters, Peth, & Büchel, 2014).

# A Fix for exponential discounting of power scaled time

For the sake of completeness, our proposed fix to the exponential discounting of (Stevens' power law scaled) subjective time (Takahashi et al., 2008) would be

$$f(D) = \exp\left(-(\kappa D)^s\right) \tag{9}$$

where $\kappa$ and $s$ are parameters. Our proposed fix is in fact equivalent to the constant sensitivity function of Ebert et al. (2007).

# B Simulation methods

The code to generate the figures is available at `https://osf.io/uscmd/`. We used the Python programming language, version 3.6 (Python Software Foundation, `https://www.python.org/`).

## B.1 Experimental design

Our simulated inter-temporal choice tasks used the heuristic adaptive experimental design procedure described by Frye et al. (2016) to select pairs of immediate and delayed prospects. Each experimental trial consisted of an immediate prospect $P^A = (R^A, D^A)$ and a delayed prospect $P^B = (R^B, D^B)$, each of which consists of a reward $R$ and a delay $D$. Prospect A was always immediate, so $P^A$ was always equal to 0. We defined 8 valid delay levels, $D^B$ could take on values of 1, 2, 7, 14, 30, $30 \times 3$, 365, $365 \times 5$ days, and we used 8 trials per delay level. This resulted in each simulated experiment consisting of 64 trials. In summary, the choice presented to the simulated participant was between $P^A$ and $P^B$). Responses were generated as a biased coin flip (Bernoulli trial), see below.

## B.2 Likelihood methods

We used maximum likelihood estimation methods for parameter estimation. We either evaluate the likelihood over a grid of possible parameter values (grid approximation), or we use a maximisation procedure to maximise the probability of the data given the parameters. Practically, we minimised the negative log likelihood of the data being generated by a given set of parameters $\theta$

$$NLL = -\Sigma_{t=1}^{T} \log(P(\text{data}_t|\theta)). \tag{10}$$

The likelihood of the data on a trial for given parameters was modelled as a biased coin flip, i.e. a Bernoulli trial,

$$P(\text{data}_t|\theta) = \text{Bernoulli}(P(R_t = 1|P_t^A, P_t^B, \theta)). \tag{11}$$

where $P(R_t = 1|P_t^A, P^B, \theta)$ is the probability of choosing $P_t^B$ (coded as $R_t = 1$).

We defined the data as

$$\text{data}_t = (P_t^A, P_t^B, R_t). \tag{12}$$

where $P^A$ and $P^B$ are prospects (see above) and $R_t$ is the response on trial $t$ of $T$ trials in total.

We defined the response probability as

$$P(R_t = 1|P_t^A, P^B, \theta) = \epsilon + (1 - 2\epsilon) \cdot \Phi\left(\frac{V(P_t^B) - V(P_t^A)}{\alpha}\right). \tag{13}$$

In the second term, $\Phi$ is the standard cumulative normal distribution which forms a psychometric function mapping the difference between present subjective values of the rewards to a response probability. We set a fixed value of $\alpha = 4$, which is the slope of this psychometric function and can be thought of as a 'comparison acuity' parameter—lower values mean greater response accuracy for prospects with similar present subjective values (see Vincent, 2016, for details). The first term deals with response errors, where $\epsilon$ was fixed at 0.01. The function $V(P)$ converts a prospect (consisting of a reward and its delay) into a present subjective value (see Equation 4). We assume a linear value function, $u(R) = R$ as is common in the discounting literature.

# References


Ariely, D., & Loewenstein, G. (2006). The heat of the moment: The effect of sexual arousal on sexual decision making. *Journal of Behavioral Decision Making*, *19*, 87–98.

Bartholdy, S., Rennalls, S., Danby, H., Jacques, C., Campbell, I. C., Schmidt, U., & O'Daly, O. G. (2017, March). Temporal Discounting and the Tendency to Delay Gratification across the Eating Disorder Spectrum. *European Eating Disorders Review*, *25*(5), 344–350.

Bickel, W. K., Odum, A. L., & Madden, G. J. (1999, October). Impulsivity and cigarette smoking: delay discounting in current, never, and ex-smokers. *Psychopharmacology*, *146*(4), 447–454.

Decker, J. H., Figner, B., & Steinglass, J. E. (2015, November). On Weight and Waiting: Delay Discounting in Anorexia Nervosa Pretreatment and Posttreatment. *Biological Psychiatry*, *78*(9), 606–614.

Ebert, J. E. J., Prelec, D., & Prelec, D. (2007, September). The Fragility of Time: Time-Insensitivity and Valuation of the Near and Far Future. *Management Science*, *53*(9), 1423–1438.

Falmagne, J. C. (1985). *Elements of psychophysical theory*. New York: Oxford University Press.

Falmagne, J. C., & Narens, L. (1983). Scales and meaningfulness of quantitative laws. *Synthese*, *55*, 287–325. doi: doi:10.1007/BF00485824

Fechner, G. T. (1966). *Elements of psychophysics* (H. E. Adler, Trans.). New York: Holt, Rinehart, & Winston. (Original work published 1860)

Fourier, J. (1822). *Théorie analytique de la chaleur*. Paris: Firmin Didot.

Franck, C. T., Koffarnus, M. N., House, L. L., & Bickel, W. K. (2015, January). Accurate characterization of delay discounting: a multiple model approach using approximate Bayesian model selection and a unified discounting measure. *Journal of the Experimental Analysis of Behavior*, *103*(1), 218–233.

Frye, C. C. J., Galizio, A., Friedel, J. E., DeHart, W. B., & Odum, A. L. (2016, January). Measuring Delay Discounting in Humans Using an Adjusting Amount Task. *Journal of Visualized Experiments*(107), 1–8.

Gilroy, S. P., Franck, C. T., & Hantula, D. A. (2017, May). The discounting model selector: Statistical software for delay discounting applications. *Journal of the Experimental Analysis of Behavior*, *107*(3), 388–401.

Jones, B. A., & Rachlin, H. (2009, January). Delay, Probability, and Social Discounting in A Public Goods Game. *Journal of the Experimental Analysis of Behavior*, *91*(1), 61–73.

Kim, B. K., & Zauberman, G. (2013, May). Can Victoria's Secret change the future? A subjective time perception account of sexual-cue effects on impatience. *Journal of Experimental Psychology: General*, *142*(2), 328–335.

Kim, B. K., & Zauberman, G. (2018, June). Psychological time and intertemporal preference. *Current Opinion in Psychology*, *26*, 90–93.

Kralik, J. D., & Sampson, W. W. L. (2012, March). A fruit in hand is worth many more in the bush: Steep spatial discounting by free-ranging rhesus macaques (Macaca mulatta). *Behavioural Processes*, *89*(3), 197–202.

Krantz, D. H. (1972). A theory of magnitude estimation and cross-modality matching. *Journal of Mathematical Psychology*, *9*, 168–199. doi: doi:10.1016/0022-2496(72)90025-9

Lemley, S. M., Asmussen, L., & Reed, D. D. (2015). Mr. right versus Mr. right now: A discounting-based approach to promiscuity. *Behavioural Processes*, *115*, 117–122.

Maxwell, J. C. (1891). *A Treatise on Electricity and Magnetism: Magnetism. Electromagnetism*. Dover Publications.

Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In M. L. Commons, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analyses of behavior* (pp. 55–73). Hillsdale, NJ: Erlbaum.

Mazur, J. E. (2007, August). Rats' choices between one and two delayed reinforcers. *Learning & Behavior*, *35*(3), 169–176.

McClure, S. M., Ericson, K. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2007, May). Time discounting for primary rewards. *The Journal of Neuroscience*, *27*(21), 5796–5804.

McKerchar, T. L., Green, L., Myerson, J., Pickford, T. S., Hill, J. C., & Stout, S. C. (2009, June). A comparison of four models of delay discounting in humans. *Behavioural Processes*, *81*(2), 256–259.

Myerson, J. (2004, September). A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin*, *130*(5), 769–792.

Myerson, J., & Green, L. (1995, November). Discounting of delayed rewards: Models of individual choice. *Journal of the Experimental Analysis of Behavior*, *64*(3), 263–276.

Myerson, J., Green, L., & Morris, J. (2011, March). Modeling the Effect of Reward Amount on Probability Discounting. *Journal of the Experimental Analysis of Behavior*, *95*(2), 175–187.

Peters, J., Miedl, S. F., & Büchel, C. (2012, November). Formal Comparison of Dual-Parameter Temporal Discounting Models in Controls and Pathological Gamblers. *PLoS ONE*, *7*(11), 1–12.

Rachlin, H. (2006, May). Notes on Discounting. *Journal of the Experimental Analysis of Behavior*, *85*(3), 425–435.

Rasmusen, E. B. (2008). Some Common Confusions About Hyperbolic Discounting. *SSRN Electronic Journal*.

Samuelson, P. A. (1937). A note on measurement of utility. *The Review of Economic Studies*, *4*(2), 155.

Schneider, S., Peters, J., Peth, J. M., & Büchel, C. (2014, April). Parental inconsistency, impulsive choice and neural value representations in healthy adolescents. *Translational Psychiatry*, *4*(4), 1–8.

Steinglass, J. E., Figner, B., Berkowitz, S., Simpson, H. B., Weber, E. U., & Walsh, B. T. (2012, May). Increased Capacity to Delay Reward in Anorexia Nervosa. *Journal of the International Neuropsychological Society*, *18*(4), 773–780.

Stevens, S. S. (1975). *Psychophysics*. New York: Wiley.

Stewart, N., Scheibehenne, B., & Pachur, T. (2018). *Psychological parameters have units: A bug fix for stochastic prospect theory and other decision models.* Manuscript submitted for publication.

Takahashi, T., Oono, H., & Radford, M. H. B. (2008). Psychophysics of time perception and intertemporal choice models. *Physica A: Statistical Mechanics and its Applications*, *387*(8-9), 2066–2074.

Van den Bergh, B., & Dewitte, S. (2008). Bikinis instigate generalized impatience in intertemporal choice. *Journal of Consumer Research*, *35*, 85–97.

Vanderveldt, A., Oliveira, L., & Green, L. (2016, April). Delay discounting: Pigeon, rat, human–does it matter? *Journal of Experimental Psychology: Animal learning and cognition*, *42*(2), 141–162.

Vicario, C. M., & Felmingham, K. (2018, April). The Perception of Time Is Underestimated in Adolescents With Anorexia Nervosa. *Frontiers in Psychiatry*, *9*, 110–8.

Vincent, B. T. (2016, December). Hierarchical Bayesian estimation and hypothesis testing for delay discounting tasks. *Behavior Research Methods*, *48*(4), 1608–1620.

Walasek, L., Mullett, T. L., & Stewart, N. (2018). *A meta-analysis of loss aversion in risky contexts.* SSRN Working Paper. doi: doi:10.2139/ssrn.3189088

Wilson, M., & Daly, M. (2004, May). Do pretty women inspire men to discount the future? *Proceedings of the Royal Society B: Biological Sciences*, *271*, S177–S179.

Wittmann, M., Leland, D. S., Churan, J., & Paulus, M. P. (2007, October). Impaired time perception and motor timing in stimulant-dependent subjects. *Drug and Alcohol Dependence*, *90*(2-3), 183–192.

Zauberman, G., Kim, B. K., Malkoc, S. A., & Bettman, J. R. (2009). Discounting Time and Time Discounting: Subjective Time Perception and Intertemporal Preferences. *Journal of Marketing Research*, *46*(4), 543–556.