Behavioural Processes 167 (2019) 103884

Contents lists available at ScienceDirect

Behavioural Processes

journal homepage: www.elsevier.com/locate/behavproc

In search of a definition of reinforcer value: Some successes and failures of the multiplicative hyperbolic model

C.M. Bradshaw

Division of Psychiatry, University of Nottingham, UK

ARTICLE INFO

Keywords:
Value
Choice
Response strength
Multiplicative hyperbolic model
Discounting

ABSTRACT

The concept of 'value' has enjoyed a central position in many theoretical accounts of choice behaviour. Several definitions of 'value' are contrasted in this paper, and one particular approach is defended, whereby value is defined as a dimensionless intervening variable. This definition is a cornerstone of the multiplicative hyperbolic model of choice (MHM), which was proposed twenty years ago as a modification of Mazur's (1987) hyperbolic model of delay discounting. This paper reviews some of the merits and shortcomings of MHM, and suggests some ways in which MHM might be extended and improved. A formal link between 'value' and the related concept of 'response strength' is suggested, and revisions of the model are proposed which may enable it to accommodate several behavioural phenomena not considered in the original formulation. Broadening the scope of MHM comes at the cost of adding to its burden of free parameters, and it is emphasised that addition of any new parameters needs empirical justification. The status of value as a dimensionless intervening variable is upheld; however it is noted that a growing body of empirical evidence for links between neurobiological phenomena and value suggests that interpretation of value as a hypothetical construct may be warranted.

LORD DARLINGTON: What cynics you fellows are!

CECIL GRAHAM: What is a cynic?

LORD DARLINGTON: A man who knows the price of everything and the value of nothing.

Oscar Wilde, *Lady Windermere's Fan*, Act 3

The characters in Wilde's supremely witty social comedy are debating art and romantic love, topics that fall well outside the scope of this paper. However, Lord Darlington's premise that there is a distinction to be drawn between an object's material worth or 'price' and its psychological value is central to the discussion that follows. It will be argued that the concept of 'value' has been defined in various ways by behaviour analysts, behavioural neuroscientists and economists, and one particular definition of reinforcer value will be defended. According to this definition, 'value' takes the form of a dimensionless intervening variable – intervening, that is to say, between the multiplicity of physical attributes of reinforcers and the behaviours that they support. This definition is integral to the multiplicative hyperbolic model of choice (MHM) developed by Ho et al. (1999).

MHM started life as a modified version of Mazur's (1987) hyperbolic model of delay discounting (referred to in this paper as the 'standard hyperbolic model', SHM). It has undergone a certain amount of tweaking in the 20 years since its inception, and has been used to analyse the effects of a number of neurobiological interventions on choice behaviour. The aim of this paper is to present an up-to-date account of MHM, to identify some of its implications for aspects of choice behaviour that were not considered in the original formulation, to identify some of its deficiencies, and to suggest how these might be remedied. The neurobiological applications of MHM have been reviewed elsewhere (Body et al., 2017; Valencia-Torres et al., 2013), and will not be discussed in this paper.

1. The concept of value

The term 'value' has many meanings. The following catalogue is not intended to be exhaustive, but rather to give a flavour of the various uses to which the word has been put and to indicate some implications of these uses for the construction of mathematical models of choice behaviour.

Definitions of value in behavioural sciences fall into two main categories: those that treat value as an objectively measurable quantity, and those that treat it as an essentially unobservable entity that may be either inferred from overt behaviour or defined by the mathematical functions that constitute particular models of behaviour.

Examples of the use of 'value' as an entity that is measurable in physical units can be found in several accounts of operant choice. For

E-mail address: c.m.bradshaw@hotmail.co.uk.

https://doi.org/10.1016/j.beproc.2019.06.007
Received 15 February 2019; Received in revised form 31 May 2019; Accepted 7 June 2019
Available online 06 July 2019
0376-6357/ © 2019 Elsevier B.V. All rights reserved.

example, indifference between two outcomes, one comprising a fixed quantity of food delivered after a delay and the other an adjusting quantity of food delivered immediately, is generally taken to imply that the values of the two outcomes are equal. It is a common practice to express the value of the delayed reinforcer in units of the quantity of the immediate reinforcer (e.g. Myerson and Green, 1995; Richards et al., 1997). According to this useage, value, expressed in physical units, is not a property of the referent reinforcer, but rather a property of a comparison reinforcer that is assumed to have the same value as the referent. This straightforward approach to defining value has much to recommend it, and is a particularly attractive approach when the values of various rewarding outcomes are expressed in terms of quantities of a ubiquitous commodity such as money. However it will be argued below that this has led to the tacit assumption that scales of value and size are interchangeable, which has unfortunate consequences for models whose aim it is to quantify the overall values of complex outcomes. It also suffers from the problem that there is no generally accepted standard with which different outcomes may be compared; thus, for example, the 'value' of a given volume of a sucrose solution of specified concentration is bound to vary, depending on the comparator that the experimenter has chosen to employ.

The term 'objective value' is sometimes used as a generic term for the physical quantity of a reward, as distinct from 'subjective value', which is essentially unmeasurable (e.g. Kable and Glimcher, 2007; Schultz, 2015). This usage has its drawbacks, because 'value' tends to be adopted as a shorthand for both expressions, resulting in the same word being used to denote two quite different concepts. Some authors avoid this difficulty by using the economic term 'utility' in place of 'subjective value', reserving 'value' or 'objective value' as a name for physical dimensions of reinforcers (e.g. Killeen, 2015).[1]

The use of utility as an economic concept may be traced back to the hedonist philosophers Jeremy Bentham and John Stuart Mill, who used this term to refer to the sum total of pleasure derived from an action minus the sum of all suffering caused by the action. The term was adopted by economists as a label for the satisfaction derived by an individual agent from the consumption of a commodity. Initially, utility was perceived as an entirely subjective entity, but Neoclassical economists sought to operationalize the term by defining it in terms of ordered preferences among commodities, expressed by *utility functions*; 'ordinal' utility functions simply arrange commodities in rank order, whereas 'cardinal' utility functions quantify differences among commodities (Hicks and Allen, 1934).[2] Later, Kahneman and Sugden (2005) and others distinguished between subjective and objective aspects of utility, naming them 'experienced' and 'decision' utilities (Read, 2004). The (experienced) utility of a commodity, u(x), is sometimes expressed in terms of imaginary units, *utils*, which are deemed to represent a scale of subjective value that may be applied to qualitatively different rewards (Kable and Glimcher, 2007; Levy and Glimcher, 2011). This practice has been deprecated by many economists as a bogus quantification of an entity that is essentially subjective, unquantifiable and unique to each individual (e.g. Keller, 2015). In the case of ordinal utility functions, *utils* convey little information, since these functions allow no mathematical operators other than $>$, $=$ and $<$. Even in the case of cardinal utility functions, it is arguable that *utils* give a false sense of quantitative precision which cannot be realized by functions based upon numerical scales that have no fixed reference point other than zero (Chipman, 1960; Hausman, 1995). A further blight on the concept of utility is the accusation of circularity. Robinson (1962) expressed this concern most forcefully: "Utility is a metaphysical concept of impregnable circularity; utility is the quality in commodities that makes individuals want to buy them, and the fact that individuals want to buy commodities shows that they have utility" (p. 47).

These obstacles to a coherent account of utility have led many economists to reject the concept altogether (e.g. Robinson, 1962; Keller, 2015). However, two strategies have been adopted by neurobiologists and behaviour analysts in an attempt to place utility on a more scientific footing. These two approaches are outlined in the following discussion which will, however, revert to the term 'value' in place of 'utility', the former term being more widely used than the latter in the behaviour analytic literature.

The first strategy, favoured by modern foraging theorists and neuroscientists, attempts to lend empirical validity to the value concept by linking it to well-defined, quantifiable functions such as reproductive fitness (Kacelnik, 2006; Kacelnik and El-Mouden, 2013) or neuronal activity in specific structures and pathways of the brain (Breton et al., 2014; Rolls, 2014; Schultz, 2015). Models based on this principle generally posit a two-stage 'computation' of value. According to some neurobiological models, for example, the values of particular rewards are first processed by modality-specific pathways; these multiple sources of value then coalesce to yield an overall value that is represented in a different region or pathway of the brain. According to models of this type, overall value constitutes a 'common currency' which animals may use in making comparisons between qualitatively different reinforcers (Conover and Shizgal, 1994; Shizgal and Conover, 1996). This approach will be briefly considered in Section 9 of this paper.

The latter strategy is represented by some behaviour analytic models of choice, including MHM, in which 'value' is perceived not as a substantive entity, but rather as an abstract mathematical entity that aids the description of functional relations between hedonic stimuli or events and overt choice behaviour – in other words, an intervening variable. In contrast to the models described earlier, in which value, an intervening variable, is quantified in physical units, MHM defines value as a dimensionless quantity. This has the advantage that different features of reinforcers may be incorporated into the value function regardless of the dimensions and scales of measurement used to characterise these features. By denying value any empirical status, models of this type evade Skinner's (1950) opprobrium of theoretical models that straddle the boundaries of different 'dimensional systems' (*ibid.* p. 216). However, the dimensional neutrality of value in these models leaves open the possibility of identifying links between value and biological phenomena at some point in the future.

Finally, the word 'value' is, of course, an indispensible descriptor of the numerical quantity of any empirical or theoretical variable, as in phrases such as 'an increasing value of x'. In the present paper, no attempt has been made to eschew this ubiquitous use of the word 'value', and it is hoped that the context will make it clear when 'value' is being used in this general sense and when it is being used in the theoretical senses outlined above.

The nomenclature used in the following discussion is intended to be consistent with that used in previous accounts of MHM. Not

[1] Although 'subjective value' and 'utility' are often used interchangeably by psychologists, ecomonists generally draw a distinction between the two concepts. According to Schultz (2015), "There is a crucial difference between subjective value and utility. Although both subjective value and utility are estimated from measured behavioral choices, utility is a mathematical function of objective value [u(m)] that allows to predict the subjective value even for choices that have not been measured. Such mathematical functions allow to determine whole distributions of subjective values and establish useful terms such as EU [expected utility] for comparison between choice options. Utility is a universal measure of subjective value that does not require immediate behavioral assessment every time and thus constitutes the fundamental variable of economic decision theory" (p. 894). The distinction between value and utility is expressed rather differently by Richardson (1994): "A distinction is often drawn between 'value' which is the result of decision-making in a risk-free environment and 'utility' which is revealed under conditions of risk" (p. 9). This definition of utility corresponds to the proposition of Prospect Theory (Kahneman and Tversky, 1979) that EU is a linear function of subjective value (SV): $EU = \pi.SV$, where π is 'subjective probability' (see Rolls, 2014, for discussion).

[2] The term 'cardinal' in this context has been criticised as being inconsistent with conventional mathematical usage (Chipman, 1960). However, it has become firmly established in the economic literature.

Table 1
Notation used in this paper.

notation used in this paper	definition	units	equivalent in previous work	key reference
V	1 value	dimensionless	V	Rachlin (1971); Ho et al. (1999)
	2 value	same units as q	V	Mazur (1987); Richards et al. (1997)
a	averaging parameter (exponent of generalized mean)	dimensionless	a	Bradshaw (2017)
d	response-reinforcer delay	s	d	Mazur (1987)
$d_{B(50)}$	indifference delay to reinforcer B	s		
d_{Am}	arithmetic mean delay to reinforcer A	s		
H	odds sensitivity parameter	same units as θ	H	Rachlin et al. (1991)
i	immediacy	s^{-1}	i	Rachlin (1971)
K	delay discount parameter	s^{-1}	K	Mazur (1987)
K_H	reinforcement rate corresponding to $R_{max}/2$	reinforcers h^{-1}	r_e	Herrnstein (1970)
q	quantity, size	mg, µl, etc.	A	Mazur (1987)
$q_{B(50)}$	indifference size of reinforcer B	mg, µl, etc.		
q_{Am}	arithmetic mean size of reinforcer A	mg, µl, etc.		
Q	size sensitivity parameter	same units as q	Q	Ho et al. (1999)
r	reinforcement rate	reinforcers h^{-1}		
R	response rate	responses min^{-1}		
R_{max}	maximum response rate	responses min^{-1}	k	Herrnstein (1970)
RS	response strength	behavioural units (e.g. response rate)		
RS_{rel}	relative response strength	dimensionless ($RS_{rel}=RS/\eta$)		
s	denominator exponent in logistic psychometric function	dimensionless	s	Valencia-Torres et al. (2011)
s_{max}	steady-state value of s after extended training	dimensionless	s_{max}	Valencia-Torres et al. (2011)
z	exponent of reinforcer size in delay-discount equation	dimensionless	z	Locey and Dallery (2009)
ε	efficacy (reinforcer-specific maximum value)	dimensionless (0 - 1)		
η	maximum response strength	same units as RS		
θ	'odds against':	$(1/p)$-1	θ	Rachlin et al. (1991)
ϕ	generic metric of a reinforcer	physical units (s^{-1}, mg, etc.)	ψ	Rachlin (2006)
Φ	generic sensitivity parameter	same units as ϕ	k	Rachlin (2006)
Ψ	value of ϕ corresponding to $\eta/2$	same units as ϕ		

Note on subscripts. A and B refer to choice alternatives. In situations where a choice response may lead to one of several outcomes, these outcomes are subscripted by 1, 2, etc., and their arithmetic mean by m; for example, in Mazur's (1984) adjusting-delay schedule the two delays associated with reinforcer A are labelled d_{A1} and d_{A2} and their arithmetic mean is d_{Am}. Different values of q and d used in different phases of an experiment are labelled q_{AI}, q_{AII}, etc.

infrequently, however, the same variable or parameter has been assigned different symbols in different models. For the convenience of the reader, Table 1 lists symbols used in this paper and the corresponding symbols used in previous work.

2. Delay discounting: the standard hyperbolic model (SHM)

The concept of delay discounting, that is, the notion that the value of a reward declines as a function of the delay that is interposed between the rewarded response and the primary reward, is no doubt a familiar one to most readers of this journal, and will not be justified in detail here. Ainslie's (1974) classic finding that pigeons' preference for the larger, and more delayed, of two reinforcers, was often reversed when the delays to both reinforcers were extended by the same amount, delivered, in the eyes of many behaviour analysts, the *coup de grâce* to traditional exponential models of delay discounting, and paved the way for a new generation of hyperbolic models, the prototype of which is Mazur's (1987) classic hyperbolic model, SHM.

According to SHM, in its original form, the value of a delayed reinforcer, V, delivered d seconds after a response may be defined thus:

$$V = q \cdot \frac{1}{1 + Kd} \quad , \tag{1}$$

where q is the amount or quantity of the reinforcer expressed in physical units, and K, expressed in units of reciprocal seconds, is a constant of delay discounting (Mazur, 1987). Note that the fraction in the right-hand side of Eq. (1) is dimensionless, and therefore V is expressed in units of reinforcer size (i.e. the same units as q). It will be argued below that this may be seen as a shortcoming of Eq. (1); however, it is important to note that it does not present a problem for an investigator whose main concern is to examine the determinants of indifference between two reinforcers, because indifference between A and B implies that $V_A = V_B$, and so long as q_A and q_B are expressed in the same units (say, mg or µl), a null equation may be constructed, from which V is erased:

$$\frac{q_A}{1 + Kd_A} = \frac{q_B}{1 + Kd_B} \quad . \tag{2}$$

This equation may be transformed into a linear relation between the delays to the two reinforcers:

$$d_{B(50)} = \frac{1}{K} \cdot \left[\frac{q_B - q_A}{q_A} \right] + d_A \frac{q_B}{q_A} \tag{2a}$$

where $d_{B(50)}$ is the indifference delay to B corresponding to any given value of d_A (Mazur, 1987).[3]

[3] The primary purpose of Eq. (2) (and other null equations, e.g. Eq. (5), see below) is to eliminate the unmeasurable intervening variable V, allowing the subordinate intervening variable K to be determined. However the transformation process may distort the estimate of of K and its associated error, which may have implications for any subsequent statistical evaluation of the parameter. It is well known that the error limits surrounding many normally distributed dependent variables deviates markedly from normality following transformation, and conversely, some non-normally distributed variables may be 'normalized' by transformation. In the case of variables that follow negatively accelerated functions, such as pharmacological dose-response curves or psychometric curves, there is often a positive relation between the mean and standard deviation of the data points. Similarly, there is often a positive relation

A potential problem with Eq. (1) is its tacit assumption that the value of a reinforcer is directly proportional to its physical size, q. This issue was recognized by Mazur and Herrnstein (1988), who argued that q (A in their notation) should be perceived not as an exact measure of amount, but rather as a quantity that is monotonically related to amount; however the nature of the monotonic relation was not specified. Locey and Dallery (2009) proposed a modified form of Eq. (1) in which q is raised to a power z, where $z < 1$:

$$V = q^z \cdot \frac{1}{1 + Kd} \ . \tag{1a}$$

This equation defines V as negatively accelerated function of q. The possibility of non-linear scaling of quantity is also recognized in Myerson and Green's (1995) 'hyperboloid' model:

$$V = q \cdot \frac{1}{(1 + Kd)^s} \ , \tag{1b}$$

where the sensitivity parameter s represents the ratio of two exponents that modulate the effects of size and delay.

In the case of electrical medial forebrain bundle (MFB) stimulation, a logistic function relating 'subjective' to 'objective' reinforcer magnitude (the 'reward growth function') has been proposed by Gallistel, Shizgal and their colleagues (Arvanitogiannis and Shizgal, 2008; Breton et al., 2014; Gallistel and Leon, 1991; Hernandez et al., 2010; Leon and Gallistel, 1992, 1998; Simmons and Gallistel, 1994). Like the other non-linear functions discussed above, value or 'subjective magnitude' is expressed in objective units of reinforcer magnitude (in this case, the parameters of the flow of action potentials in the MFB). This point will be briefly addressed in Section 9.

An alternative definition of the function relating size and value (Ho et al., 1999), which dispenses with physical units of measurement, is discussed in the following section.

3. The multiplicative hyperbolic model (MHM)

MHM was formulated in response to the finding that indifference between a small reinforcer delivered after a short delay and a larger reinforcer delivered after a longer delay is influenced by the absolute, as well as the relative, sizes of the reinforcers, suggesting that a term expressing sensitivity to reinforcer size was needed in any adequate model of delay discounting (Bradshaw and Szabadi, 1992; Wogar et al., 1992). Ho et al. (1999) proposed a hyperbolic relation between the size and the value of a (non-delayed) reinforcer:

$$V = \frac{1}{1 + Q/q} \ , \tag{3}$$

where Q is a sensitivity parameter for size. Replacing this expression for q in SHM's delay-discounting equation (Eq. (1)) gives the following definition of value:

$$V = \frac{1}{1 + Q/q} \cdot \frac{1}{1 + Kd} \ . \tag{4}$$

Ho et al. (1999) suggested that other, as yet unspecified, features of reinforcers might be incorporated into the equation, for example the hyperbolic expression proposed by Rachlin et al. (1991) to describe probability discounting:

$$V = \frac{1}{1 + Q/q} \cdot \frac{1}{1 + Kd} \cdot \frac{1}{1 + H\theta} \ , \tag{4a}$$

where θ is the odds against the delivery of a reinforcer (θ = [1/p] -1) and H is an odds-discounting parameter (see also Green and Myerson, 2004).

Expressing delay in reciprocal units ('immediacy', i), Eq. (1) becomes

$$V = \frac{1}{1 + K/i} \ , \tag{1c}$$

allowing Eq. (4a) to be written in a more general form:

$$V = \prod_{x=1}^{n} \left[\frac{1}{1 + \Phi_x/\phi_x} \right] \tag{4b}$$

where ϕ_x is the numerical value of a physical dimension, x (for instance, volume measured in μl, or immediacy measured in reciprocal seconds), Φ_x is the corresponding sensitivity parameter (expressed in the same units), and n is the number of hyberbolic components incorporated into the equation. Eq. (4b) is a generalized form of the definition of value offered by MHM.[4] Value is conceived as a dimensionless quantity between 0 and 1, the effect of each hyperbolic term in the equation being simply to reduce V by a factor. This definition is consonant with Rachlin's (1971, 2006) assertion that the only appropriate measure of value is a relative (dimensionless) quantity and that the use of such a measure affords the opportunity to compare the values of qualitatively different reinforcers, an opportunity that is denied by definitions of value couched in physical units.

It is important to note that MHM's adoption of the rectangular hyperbola to define the size/value relation is not based on the goodness of fit of this function to empirical data. As discussed above, there are evidential reasons for proposing a non-linear, (negatively accelerated) size/value function, consistent with the economic principle of diminishing marginal utility. However, the hyperbola is by no means the only available function that could fulfil this need (see Bradshaw, 2018). From the standpoint of MHM, the principal advantage of the hyperbola, compared to competing functions such as the exponential modulation of reinforcer size (Eq. (1a)), is that the hyperbola preserves the dimensionless quality of the value concept (Rachlin, 2006).

4. Implications of MHM for the construction of indifference functions

Incorporation of a size-sensitivity term into the value equation complicates the determination of indifference points. Substitution of the size-sensitivity term of Eq. (4) into Eq. (2) yields the following linear indifference function:

$$d_{B(50)} = \frac{1}{K} \cdot \left[\frac{Q/q_A - Q/q_B}{1 + Q/q_B} \right] + d_A \frac{[1 + Q/q_A]}{[1 + Q/q_B]} \tag{5}$$

(Ho et al., 1999). K is present in the intercept, and Q in both the slope and the intercept. It follows that no single indifference point provides a basis for discriminating between the influences of delay discounting and size sensitivty on preference. However, a linear plot of $d_{B(50)}$

(*footnote continued*)
between the estimate of the location parameter and its standard error (e.g. K_H in Herrnstein's hyperbola (Eq. (16)); see Bradshaw et al., 1978a, 1981b, 1993), suggesting that logarithmic transformation of the parameter prior to conventional statistical analysis may be appropriate (Bradshaw and Szabadi, 1993). Unfortunately, the error patterns associated with most of the parameters discussed in this paper are unknown. Bayesian approaches to model evaluation that accommodate deviations from the classical normality/homoscedasticity assumption may prove helpful in future investigations of MHM (Martin and Williams, 2017; Villarreal et al., 2019). The availability of software such as Stan (Stan Development Team, 2016) and JAGS (Plummer, 2003) that operate in R (R Core Team, 2018) will no doubt facilitate this endeavour.

[4] The use of immediacy in place of delay draws attention to the symmetrical roles of the size of a primary reinforcer and its temporal displacement from the rewarded response as determinants of the overall value of an outcome, as expressed in Eq. (4b). However, delay is the universally recognised metric of temporal discounting in the psychological and economic literature. Unless specified otherwise, the following discussion will adhere to the conventional useage.

against d_A does enable these two processes to be disentangled. If an intervention alters the slope of this plot it may be assumed that it has affected Q, whereas a change of the intercept with no accompanying change of the slope implicates an effect on K (Ho et al., 1999; Valencia-Torres et al., 2013). In principle, K may be determined from the formula [slope – 1]/intercept, and Q by substitution of the known values of q_A and q_B into the empirical value of the slope. However the reliability of such estimates at the level of the individual subject is often compromised by within-subject variability in the data (see Body et al., 2017).

Eq. (5) offers another means of estimating K and Q. If no delay is imposed on the delivery of reinforcer A ($d_A \approx 0$), Eq. (5) simplifies to

$$d_{B(50)} = \frac{1}{K}\cdot\left[\frac{Q/q_A - Q/q_B}{1 + Q/q_B}\right]. \tag{5a}$$

If two indifference points, $d_{B(50)I}$ and $d_{B(50)II}$, are obtained using values of q_A and q_B chosen in such a way that $(1/q_{AI} - 1/q_{BI}) = (1/q_{AII} - 1/q_{BII})$, the ratio of the indifference delays is

$$\frac{d_{B(50)I}}{d_{B(50)II}} = \frac{1/q_{AI} - 1/q_{BI}}{1/q_{AII} - 1/q_{BII}}\cdot\frac{1 + Q/q_{BII}}{1 + Q/q_{BI}}. \tag{6}$$

The first fraction in the right-hand side cancels to 1, and Q, the only parameter, may be calculated directly from the known values of q and the empirically determined ratio of the indifference delays; this value of Q may then be substituted into Eq. (5a), and the value of K calculated (Valencia-Torres et al., 2011, 2012). Obviously, estimating the values of two parameters from two data points (i.e. the ratio of two indifference delays) is a tall order, but preliminary evidence suggests that the method holds promise, at least at the level of group mean data (Valencia-Torres et al., 2011, 2012).

5. Value as a discriminandum

It is assumed by both SHM and MHM that when faced with a choice between two reinforcers, A and B, an animal will choose the option with the higher value. It follows that proportional choice of B (%B) should be a step function of any relevant dimension of a reinforcer (ϕ: size, delay etc.); that is to say, it should be 100% when $V_B > V_A$, and 0 when $V_A > V_B$. However, this is seldom the case in real life; empirical preference curves usually have a graded form that is characteristically sigmoid when ϕ is plotted on a logarithmic scale. Bezzina et al. (2007) suggested that such preference functions may be regarded as psychometric functions for the discrimination of delay. Fig. 1 shows discrete-trials choice data from 14 rats (Bezzina et al., 2007). Responses on levers A and B delivered, respectively, a smaller reinforcer (q_A) after a shorter delay (d_A) and a larger reinforcer (q_B) after a longer delay (d_B). d_B was varied across successive trial blocks in each session, while d_A was varied across successive phases of the experiment. Fig. 1A shows mean preference functions (i.e., %B vs. d_B) derived from the 14 rats; the six curves correspond to six phases of the experiment in which d_A was systematically varied between 0.5 and 12 s. Fig. 1B shows the same data plotted in semilogarithmic co-ordinates; the fitted functions are generic logistic psychometric functions: $\%B = 100/(1 + [d_B/d_{B(50)}]^s)$, where s expresses the slope of the function. The point of intersection of each curve with the horizontal 50% line is the indifference point, $d_{B(50)}$. Fig. 1C shows $d_{B(50)}$ plotted against d_A; the data are well described by a linear function with a positive intercept, as specified by Eq. (5).

The notion that value, rather than delay, may be viewed as a discriminandum was proposed by Valencia-Torres et al. (2011) in their attempt to simulate performance on an adjusting-delay schedule (Mazur, 1987). In this schedule, the subject makes repeated choices between A and B, A being a smaller reinforcer of size q_A delivered after a short delay d_A, and B being a larger reinforcer of size q_B delivered after a delay d_B which is adjusted according to the subject's choices. If the subject exhibits preference for A in block n, d_B is reduced in block $n+1$, whereas if B is preferred in block n, d_B is increased in block $n+1$. Typical performance on this schedule comprises an oscillating value of d_B which eventually settles at a quasi-equilibrium value which is taken as the indifference delay, $d_{B(50)}$ (Mazur, 1987). The oscillations of d_B can be characterized by a power spectrum, derived by the Fourier transform, the principal indices of which are power within the dominant frequency band and the period of oscillation corresponding to the dominant frequency (Fig. 2; see Bradshaw, 2017, for a brief account of the application of the Fourier transform to biological phenomena). Valencia-Torres et al. (2011) proposed that in each free-choice trial, the selection of A or B is determined by the ratio of V_A to V_B (defined by Eq. (4)), A being generally chosen if $V_A/V_B > 1$, and B being generally chosen if $V_A/V_B < 1$. It was further proposed that selection is affected by psychometric limitations, which may be expressed by the logistic function, $p(B) = 1/(1 + [V_B/V_A]^s)$, where $p(B)$ is the probability of selecting B and s is the slope of the psychometric function. Valencia-Torres et al. (2011) found that simulated performance on the adjusting-delay schedule, derived using this model, resembled the performance of rats exposed to this schedule. This extension of MHM has also been applied to behaviour maintained on an adjusting-magnitude schedule (Bradshaw, 2017, 2018; see below).

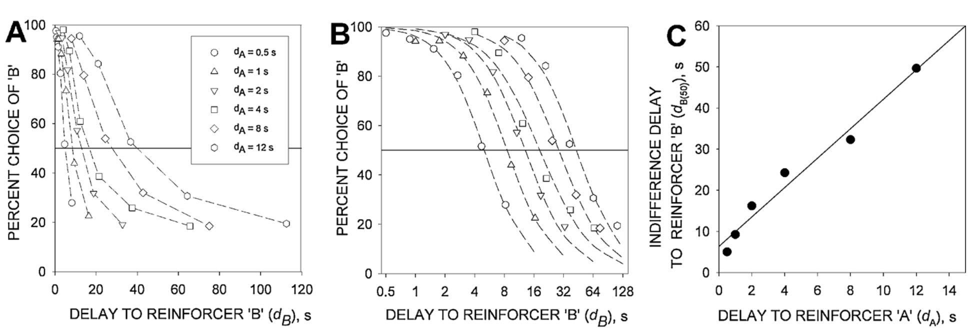

Fig. 1. Mean performance of 14 rats on a series of progressive delay schedules (data from Bezzina et al., 2007). **A**: Preference functions showing percent choice of a larger reinforcer, B, for a range of delays to that reinforcer (d_B). The six curves were obtained with different delays to a smaller reinforcer, A (d_A: the delays are shown in the inset). **B**: The same data are shown in semi-logarithmic co-ordinates; the curves are logistic functions fitted by the least squares method. The intersection of each curve with the horizontal 50% line is the indifference point, $d_{B(50)}$. **C**: Indifference function ($d_{B(50)}$ plotted against d_A: Eq. (5)), fitted by the least squares method.

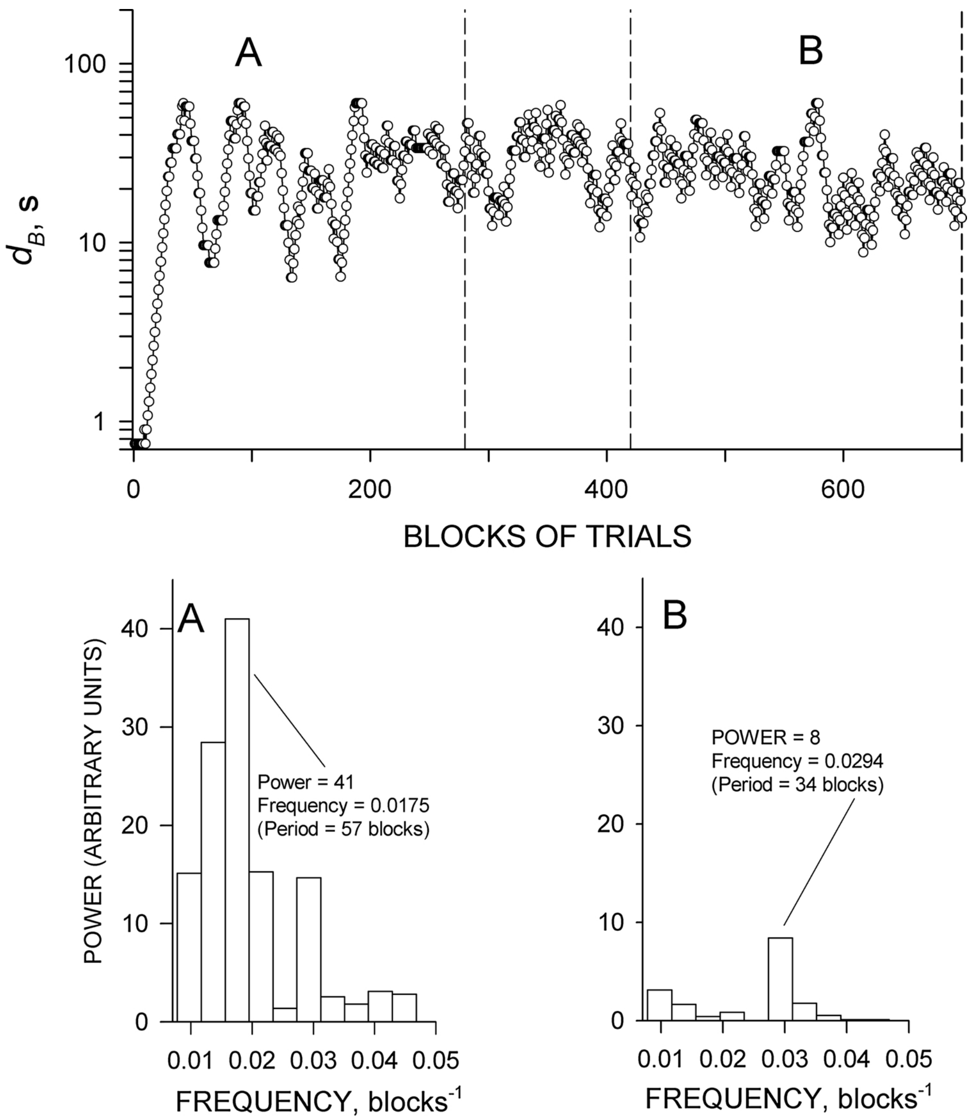

Fig. 2. Example of one rat's performance on the adjusting-delay schedule. *Upper graph:* Adjusting delay to the larger reinforcer (d_B, s) plotted against blocks of trials (trial blocks 1–700), the sizes of the reinforcers (q_A, q_B) were 25 and 100 µl of a 0.6 M sucrose solution. *Lower panels:* Power spectra derived from the Fourier analysis of the d_B data from segments A and B, demarcated by the vertical broken lines in the upper graph. Power is plotted against frequency (blocks^{-1}). The period of oscillation corresponding to the dominant frequency band, and the power within that band are shown for each segment. Note the reduction of the power of oscillation in B, following extended exposure to the schedule (data from Valencia-Torres et al., 2011).

A noteworthy feature of this extension of MHM is its minimal reliance on short-term memory as a mechanism of preference. As presently formulated, the model proposes that preference in the free-choice trials of each block is determined by the values of the reinforcers experienced in the preceding forced-choice trials of that block; no allowance is made for an influence of extended sequences of blocks of trials. This may seem unrealistic, and may therefore be viewed as a shortcoming of the model. Further experimental work is needed to establish whether a more complex model may be needed, which takes into account the cumulative effect of experience of reinforcers obtained in a sequence of trial blocks. Such a model might incorporate gradients of influence of the kind propounded by Killeen (1994, 2011). In the meantime, however, the simple, albeit naïve, version of the model seems to offer a qualitatively adequate account of the oscillations of delays and magnitudes in conventional adjusting schedules (Valencia-Torres et al., 2011; Bradshaw, 2017, 2018).

6. Reinforcer-specific value maxima: the concept of 'efficacy'

It is apparent from Eq. (4b) that when the numerical values of all the dimensions of a reinforcer greatly exceed their respective sensitivity parameters (i.e. when $\phi_1 > > \Phi_1$, ... $\phi_n > > \Phi_n$), V will approach its maximal value (1.0). Thus, for example, the value of a sucrose solution will be maximal if it is delivered in large quantities with minimal delay. According to the original formulation of MHM, the parameters represented by the generic Φ provide a basis for distinguishing between the effectiveness of different reinforcers. Take, for example, the case of a reinforcer delivered without delay after a response; *ceteris paribus*, Eq. (3) (the simplest version of Eq. (4b)) will apply, in which V is determined solely by the quantity of the reinforcer, q, and the size-sensitivity parameter, Q. Since Q represents the physical amount of the reinforcer that generates the half-maximal value ($q = Q$ when $V = 0.5$), it follows that this parameter will be smaller in the case of a richer reinforcer (say a 1 M sucrose solution) than in the case of a weaker reinforcer (say a 0.1 M sucrose solution). However, it also follows from Eq. (3) that both reinforcers should be capable of generating the maximum attainable value ($V = 1.0$) if given in sufficient quantites. This seems highly unlikely; sensory physiology and common sense rebel against the notion that the maximum value of the concentrated solution might be replicated simply by a providing a larger volume of the more dilute one.

This difficulty may be overcome by the addition of a new parameter. In Eq. (7), the new parameter, ε, specifies the 'efficacy' of a reinforcer, defined as the maximum value of a particular type or quality of reinforcer.

$$V = \frac{\varepsilon}{1 + Q/q} \quad . \tag{7}$$

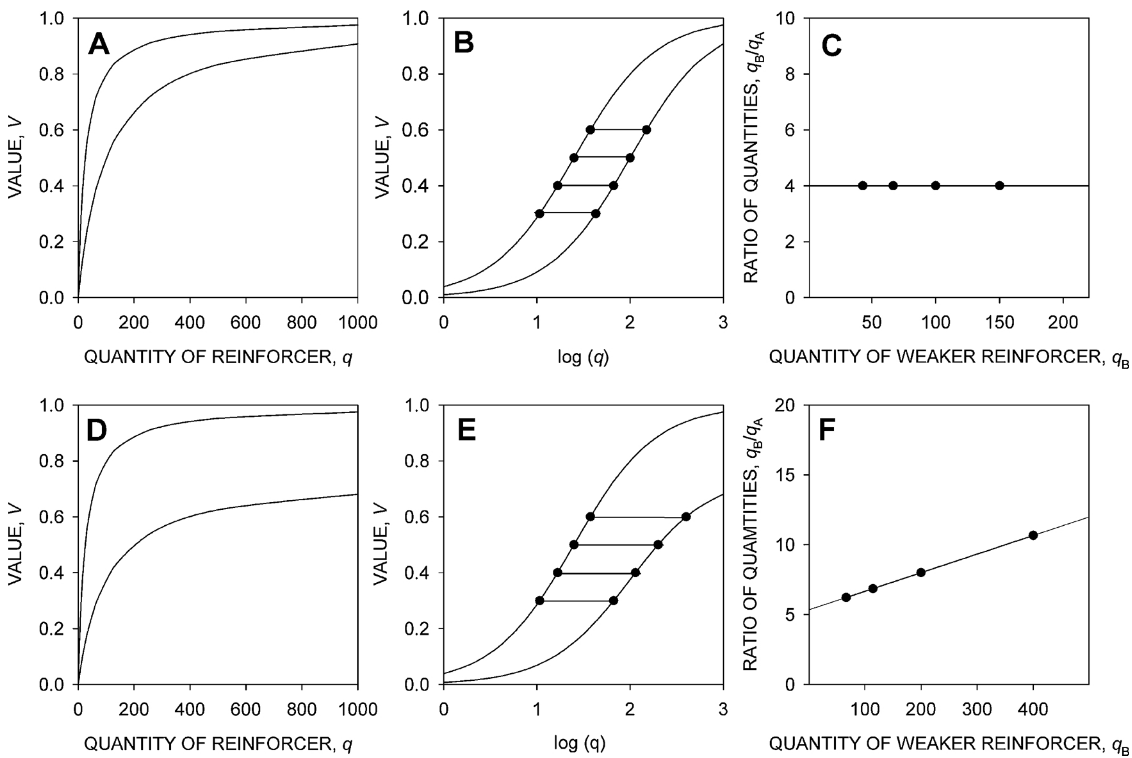

Fig. 3. Proposed effect of 'efficacy', ε, on reinforcer value. **A**: *ordinate*, value; *abscissa*, reinforcer size, *q*. The two curves represent the value/size function for a strong and a weak reinforcer differing only with respect to size sensitivity, *Q* (Eq. (3)). **B**: The same functions in semi-logarithmic co-ordinates. The points joined by horizontal lines indicate equi-valued sizes of the two reinforcers. **C**: Relation between the ratio of the equi-valued sizes of the two reinforcers and the size of the weaker reinforcer. Note that the ratio of the equi-valued sizes is unaffected by the absolute size of the weaker reinforcer (Eq. (8b)). **D**, **E**, and **F**: Corresponding functions for two reinforcers that differ in efficacy, ε, as well as size sensitivity, *Q*. Note the positive slope of the linear function in **F**, reflecting the different efficacies of the two reinforcers (Eq. (9b)).

If A is a stronger reinforcer than B, $\varepsilon_A > \varepsilon_B$. In the case of a maximally efficacious reinforcer, $\varepsilon = 1$ and Eq. (7) becomes identical to Eq. (3). In choice schedules in which the same type of reinforcer is provided in both alternatives (e.g., the same brand of food pellet, or the same concentration of sucrose), efficacy may be ignored, because ε cancels out of the indifference equation. However, the impact of efficacy on choice may be revealed in situations when choices are made between qualitatively different reinforcers. Consider, for example, the case of choice between a concentrated sucrose solution, A, and a more dilute solution, B. $Q_A < Q_B$, and if no allowance is made for different efficacies of the two solutions (i.e., $\varepsilon_A = \varepsilon_B = 1$), Eq. (3) applies to both reinforcers. At indifference, $V_A = V_B$, and expansion using Eq. (3) gives

$$\frac{1}{1 + Q_A/q_A} = \frac{1}{1 + Q_B/q_B} , \tag{8}$$

which simplifies to

$$q_B = q_A \cdot \frac{Q_B}{Q_A} , \tag{8a}$$

or

$$\frac{q_B}{q_A} = \frac{Q_B}{Q_A} . \tag{8b}$$

Eq. (8a) specifies a simple linear relation between the equally valued volumes of the two reinforcers, with zero intercept. The ratio of the equi-valued quantities of A and B (q_B/q_A) is defined by the ratio of the sensitivity parameters and is constant across all quantities of B (Eq. (8b)).

In contrast, if $\varepsilon_A \neq \varepsilon_B$, a more complex indifference relation ensues:

$$\frac{\varepsilon_A}{1 + Q_A/q_A} = \frac{\varepsilon_B}{1 + Q_B/q_B} , \tag{9}$$

which specifies a non-linear relation between the equi-valued quantities of A and B:

$$q_A = \frac{Q_A}{\frac{\varepsilon_A}{\varepsilon_B} . (1 + Q_B/q_B) - 1} . \tag{9a}$$

This relation may linearized by taking the ratio of the equi-valued quantities of the two reinforcers:

$$\frac{q_B}{q_A} = q_B \cdot \frac{1}{Q_A} \cdot \left(\frac{\varepsilon_A}{\varepsilon_B} - 1\right) + \frac{\varepsilon_A}{\varepsilon_B} \cdot \frac{Q_B}{Q_A} . \tag{9b}$$

Eq. (9b) specifies a linear relation between q_B/q_A and q_B, with a positive slope and a positive intercept.[5]

Fig. 3 shows the effect of the efficacy term on the indifference relation. The left-hand graphs show hyperbolic value functions defined by Eqs. (3) and (7). In the upper graph, both curves were derived using Eq. (3) (i.e., $\varepsilon_A = \varepsilon_B = 1$), the left curve (A) representing a lower value of *Q* than the right curve (B) (A being the more potent of the two reinforcers). In the lower graph, the values of *Q* are the same as in the upper graph, but in this case, the values of ε differ between the two reinforcers (Eq. (7)): $\varepsilon_A = 1$ and $\varepsilon_B = 0.75$. Semi-logarithmic transformation of these plots (middle graphs) emphasises the effect of the differing values of ε: when

[5] Efficacy, ε, in this model is formally identical to 'intrinsic activity' in classical pharmacological receptor theory, which constrains the maximum effect of an agonist and which may take a value of 1.0 (full agonist), between 0 and 1.0 (partial agonist or 'dualist'), or 0 (competitive antagonist) (Ariens et al., 1964). Intrinsic activity is a property of a drug's putative action at pharmacological receptors. It differs from pharmacological efficacy, which incorporates both drug-specific and tissue-specific components, and which may take a value greater than unity (Furchgott and Bursztin, 1967). Eq. (9b) is formally identical to an equation that describes the relation between the ratio of equi-effective concentrations of a full and a partial agonist and the concentration of the partial agonist (Kenakin, 1997, p. 280).

the two reinforcers differ only with respect to Q (upper middle graph), the two curves are essentially parallel, and the displacement of the B curve relative to the A curve is uniform. In contrast, when ε also differs between the two reinforcers (lower middle graph) the proportional discrepancy between equi-valued quantities of the two reinforcers increases as a function of q. This is reflected in the different plots defined by Eqs. (8b) and (9b) (upper and lower right-hand graphs).

Of course, caution is always warranted when advocating the intrusion of additional parameters into a model. However, it will be argued at a later point in this paper that there is already some evidence, albeit indirect evidence, that an efficacy parameter is needed in the quantification of value. Future experiments may allow direct assessment of the necessity, or otherwise, of the efficacy parameter by model comparison using the Akaike or Bayesian Information Criterion (Sakamoto et al., 1986; Schwartz, 1978). In any case, the patent difference between the relations specified by Eqs. (8b) and (9b) would seem to be an eminently testable (but as yet untested) implication of the efficacy concept.

7. Complex reinforcers and the problem of averaging

In the foregoing discussion, the problem of teasing apart the hypothetical processes of delay discounting and size sensitivity has been addressed mainly by algebraic manipulation. Because, *ex hypothesi*, value cannot be measured directly, null equations are constructed from which V is eliminated. Applications of Eq. (5) typify this approach (e.g. Kheramin et al., 2002; Bezzina et al., 2007, 2009). Since this equation contains two parameters, K and Q, it is necessary to measure the principal dependent variable, $d_{B(50)}$, across a range of values of the independent variable, d_A. Conventional function-fitting methods are then used to extract the values of the two parameters. In the case of Eq. (5), as discussed above, they may be calculated from the slope and intercept of the linear indifference function.

A different approach is exemplified by an adjusting-delay schedule devised by Mazur (1984), in which choices are made between reinforcers that differ with respect to delay but not size. Selection of A is followed by the delivery of a reinforcer of size q_A after a delay d_{A1} or d_{A2}, with equal probability, whereas selection of B is followed by delivery of a reinforcer of the same size ($q_B = q_A$) after a delay that is adjusted according to the subject's choices. According to SHM, indifference is defined as follows:

$$\frac{q_B}{1 + K.\ d_{B(50)}} = 0.5\cdot\left[\frac{q_A}{1 + K.\ d_{A1}} + \frac{q_A}{1 + K.\ d_{A2}}\right] \tag{10}$$

(Mazur, 1984). Since $q_B = q_A$, the qs cancel out, and solving for $d_{B(50)}$ yields

$$d_{B(50)} = \frac{d_{Am} + K.\ d_{A1}.\ d_{A2}}{1 + K.\ d_{Am}}\ , \tag{10a}$$

or

$$K = \frac{d_{Am} - d_{B(50)}}{d_{B(50)}\cdot\ d_{Am} - d_{A1}.\ d_{A2}}\ , \tag{10b}$$

where d_{Am} is the arithmetic mean of d_{A1} and d_{A2}. The deletion of the qs from Eqs. (10a) and (10b) means that these equations are derivable from both SHM and MHM.

da Costa Araújo et al. (2010) used a similar adjusting schedule to assess sensitivity to reinforcer magnitude, free of contamination by the influence of delay. Choices were made between reinforcer A, of size q_{A1} or q_{A2} occurring with equal probability, and an adjusting reinforcer, B, whose size increased and decreased in successive trial blocks according to the subject's preference. In this case, SHM offers a very simple prediction for the indifference magnitude of B:

$$q_{B(50)} = 0.5.\ (q_{A1} + q_{A2}),\ \text{or}\ q_{B(50)} = q_{Am}\ , \tag{11}$$

where q_{Am} is the arithmetic mean of q_{A1} and q_{A2}.

As a corollary of SHM, Eq. (11) does not incorporate a size-sensitivity parameter. Taking size sensitivity into account, MHM offers a more complex definition of indifference:

$$\frac{1}{1 + Q/q_{B(50)}} = 0.5\cdot\left[\frac{1}{1 + Q/q_{A1}} + \frac{1}{1 + Q/q_{A2}}\right], \tag{12}$$

from which the indifference magnitude and the size-sensitivity parameter may be calculated:

$$q_{B(50)} = \frac{q_{A1}\cdot\ q_{A2} + Q.\ q_{Am}}{q_{Am} + Q}\ ; \tag{12a}$$

$$Q = \frac{q_{A1}\cdot\ q_{A2} - q_{B(50)}\cdot\ q_{Am}}{q_{B(50)} - q_{Am}}\ . \tag{12b}$$

(Bradshaw, 2017). In the special case where $q_{A1} = q_{A2}$, the first term of the numerator becomes $q_{Am}{}^2$, and Eq. (12a) simplifies to Eq. (11).

It may be noted that this reasoning depends upon the assumption, made by both SHM and MHM, that the overall value of a complex reinforce is defined thus:

$$V_A = 0.5\cdot[V_{A1} + V_{A2}] \tag{13}$$

or, in a more general form,

$$V = \sum_{i=1}^{n} (p_i\cdot V_1)\ /n \tag{13a}$$

where p_i is the probability associated with each component of the complex reinforcer and n is the number of components (Mazur, 1984). It should also be noted that Eq. (13) is incompatible with the proposal by Ho et al. (1999) that probability of reinforcement may be incorporated into the definition of value as a hyperbolic odds-discounting factor (Rachlin et al., 1991). In fact it is questionable whether the principle of hyperbolic odds discounting, as expressed in Eq. (4), is applicable to complex reinforcers in the adjusting-magnitude schedule. Take, for instance, the typical situation in which $p_{A1} = p_{A2} = 0.5$. By definition, $\theta_{A1} = \theta_{A2} = 1$, and (since $p_B = 1$) $\theta_B = 0$. Under these circumstances, Eq. (4) leads to the following definition of indifference between A and B:

$$\frac{1}{1 + Q/q_{B(50)}} = \left[\frac{1}{1 + Q/q_{A1}} + \frac{1}{1 + Q/q_{A2}}\right]\cdot\frac{1}{1 + H} \tag{14}$$

In the special case where $q_{A1} = q_{A2}$, Eq. (14) reverts to the expected solution expressed by Eq. (11) if and only if $H = 1$. For any other value of H, Eq. (14) leads to the implausible conclusion that when there is no variance in q_A, $q_{B(50)} \neq q_A$.

The above theoretical argument suggests that hyperbolic odds discounting is an unsatisfactory basis for combining the values of the components of a complex reinforcer. However, recent empirical evidence suggests that the use of untransformed probability, as in Eqs. (13) and (13a), is also problematic. The left-hand panel of Fig. 4 shows the theoretical indifference magnitudes $q_{B(50)}$ corresponding to a range of values of q_{Am}, with q_{A1} held constant and q_{A2} systematically varied; the probabilities of reinforcement following a response are $p_{A1}, p_{A2} = 0.5$, $p_B = 1$. The diagonal straight line is defined by Eq. (11) (SHM), and the downward curving line by Eq. (12a) (MHM). The points in the right-hand panel show results obtained from eight rats trained under the adjusting-magnitude schedule using a 0.6 M sucrose solution as the reinforcer (Bradshaw, 2017). The failure of both equations is obvious; clearly, a new approach is needed. One way of accommodating the data within MHM is to abandon the arithmetic mean as a method for combining the values of the components (Eq. (13)), and replace it by a more liberal averaging method – the generalized, or Hölder, mean:

$$V = \left[\left(\sum_{i=1}^{n} V_i^a\right)/n\right]^{\frac{1}{a}}, \tag{15}$$

where a is an 'averaging parameter'. In the adjusting magnitude

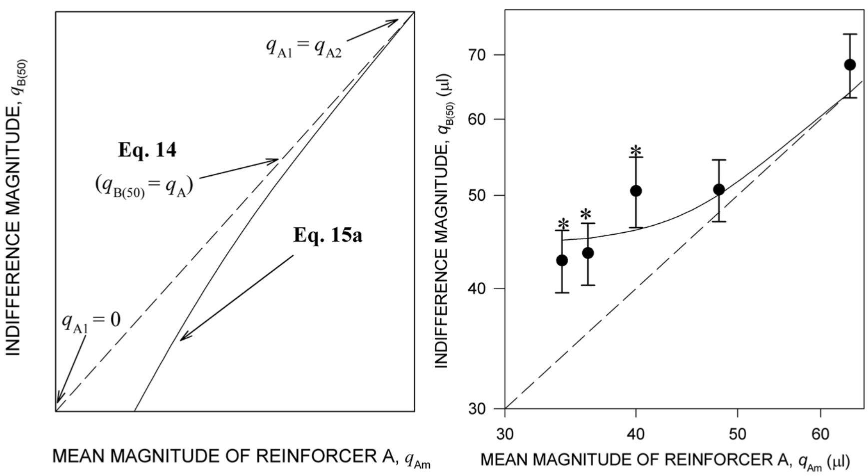

Fig. 4. ***Left-hand graph***. Predicted relation between the indifference magnitude of reinforcer B, $q_{B(50)}$, and the arithmetic mean magnitude of reinforcer A, q_{Am} (axes are $\log_{10}$-transformed). Broken diagonal line shows the prediction of SHM (Eq. (11)), where $q_{B(50)} = q_{Am}$. The continuous line shows the prediction of the original version of MHM (Eq. (12a)), Note that the presence of the size-sensitivity parameter, Q, in the equation causes the function to lie below the diagonal. ***Right-hand graph***. Relation between indifference magnitude of reinforcer B ($q_{B(50)}$, µl) and the mean magnitude of reinforcer A (q_{Am}, µl): mean data (± s.e.m.) from eight rats trained under an adjusting-magnitude schedule. Note that at low values of q_{Am}, $q_{B(50)}$ lies above the diagonal (* $p <$ 0.05). The curve was drawn from Eq. (15a) with Q = 186 and a = 2.51. (Modified from Bradshaw, 2017.)

schedule, indifference occurs when $V_B = V_A$. If V_B, V_{A1} and V_{A2} are all defined by Eq. (3), and Eq. (15) is used to derive the value of the complex reinforcer A from V_{A1} and V_{A2}, an indifference curve may be computed for $q_{B(50)}$. If $a > 1$, this indifference curve lies above the linear indifference function specified by SHM (Eq. (11)).[6] The curve in the right-hand panel of Fig. 4 is a plot of this function with Q set at 186 µl, the average value of this parameter found in previous experiments in which the same sucrose solution (0.6 M) was used as the reinforcer. The best fit to the data was obtained with a value of a of 2.51 (Bradshaw, 2017). Further evidence for a value of $a > 1$ was obtained using a range of concentrations of a sucrose reinforcer (Bradshaw, 2018); in this case, the deviation of the adjusting magnitude from the linear indifference function was inversely related to sucrose concentration, an effect that may reflect the opposing influences of a and Q on the overall value of the complex reinforcer (see following section).

The *ad hoc* intrusion of the new parameter, a, might be seen as yet an another unwelcome violation of the parsimony of MHM. However, it carries two potential advantages. Firstly, it expresses a principle that may be testable in various situations other than the adjusting-magnitude schedule; that is, the finding that $a > 1$ implies that whenever the same response delivers more than one outcome, higher valued outcomes exert a greater influence than lower valued outcomes on the combined value of the outcomes. Secondly, use of the generalized mean obviates the need for probability as an explanatory principle. Both SHM and the original version of MHM tacitly or explicitly assume that animals' choices are influenced by information about the probabilities of future outcomes (in the parlance favoured by many neuro- and behavioural economists, 'expectation' or 'expectancy'); this information must be gleaned by the cumulation of many experiences of these outcomes. The use of the generalized mean in the present version of MHM removes the need assume that animals bring any direct knowledge of probability to bear on their choices. This will be discussed in greater detail in the following section.

On the debit side, adoption of the generalized mean has an unfortunate consequence for one approach to the assessment of delay discounting and size sensitivity. There have been several attempts to use single indifference points obtained using adjusting-delay and adjusting-magnitude schedules in order to differentiate between the effects of interventions on these two processes (da Costa Araújo et al., 2010; McClure et al., 2014; Moschak and Mitchell, 2014). Since a affects the computation of overall value of complex reinforcers in both schedules, a selective effect of an intervention on one or other schedule may reasonably be ascribed to the operation of either K or Q. However if the intervention affects indifference points in both schedules, there is no way of distentangling effects on these parameters from effects on a. Moreover, the contribution of a to the determination of the overall value of complex reinforcers contaminates numerical estimates of K and Q obtained using Eqs. (10b) and (12b).

8. Implications for risk sensitivity

Choice between an outcome that is 'certain' and one that is variable or 'uncertain' is often referred to as 'risky choice'. Adjusting-delay and adjusting-magnitude schedules constitute examples of risky choice (Mazur, 2004). By convention, choice is regarded as 'risk insensitive' if the indifference point coincides with the arithmetic mean value of the manipulated dimension of the uncertain outcome. Thus, in the case of Mazur's (1984) adjusting-delay schedule, risk insensitivity is said to occur when $d_{B(50)} = d_{Am}$. When $d_{B(50)} < d_{Am}$, choice is regarded as 'risk prone', because the subject is assumed to have equated the value of the variable outcome A with a 'better' value of outcome B. Conversely, when $d_{B(50)} > d_{Am}$, choice is regarded as 'risk averse', because the subject is assumed to have equated the value of the variable outcome A with a 'worse' value of outcome B. The same logic may be applied to the adjusting-magnitude schedule (da Costa Araújo et al., 2010): in this case, choice is regarded as risk insensitive when $q_{B(50)} = q_{Am}$, risk prone when $q_{B(50)} > q_{Am}$, and risk averse when $q_{B(50)} < q_{Am}$ (Mazur, 2004).

It is generally agreed that animals tend to be risk prone when making choices between reinforcers that differ with respect to delay (Bateson and Kacelnik, 1995; Cicerone, 1976; Davison, 1969; Mazur, 1984; Rider, 1983; for reviews, see Kacelnik and Bateson, 1996; Kacelnik and El-Mouden, 2013). This is the trend predicted by both SHM and MHM. Indifference is defined by Eq. (10a), from which it can be shown that $d_{B(50)}$ is necessarily less than d_{Am} (Bradshaw, 2017; Mazur, 1984).

The literature on choice between reinforcers of different sizes is more controversial. In such cases, risk proneness (Bradshaw, 2017,

[6] The indifference magnitude, $q_{B(50)}$, is given by by the following equation:

$$q_{B(50)} = Q \cdot \frac{[(V_{A1}{}^{a} + V_{A2}{}^{a})/2]^{1/a}}{1 - [(V_{A1}{}^{a} + V_{A2}{}^{a})/2]^{1/a}} \tag{15a}$$

Notwithstanding its somewhat daunting appearance, this equation contains only two parameters, Q and a. It simplifies to Eq. (12a) in the special case of a = 1. The finding that the adjusting magnitude may exceed the arithmetic mean magnitude of a complex reinforcer is unexpected both from the standpoint of SHM (which specifies that the adjusting magnitude should be equal to the arithmetic mean) and from the standpoint of the original version of MHM (which specifies that the adjusting magnitude should be less than the arithmetic mean). Introduction of the generalized mean enables MHM to accommodate the unexpected finding; however it is by no means the only way in which this could be accomplished. For a comparison of the present approach with an alternative solution to the problem involving exponential modulation of reinforcer magnitude, see Bradshaw (2017).

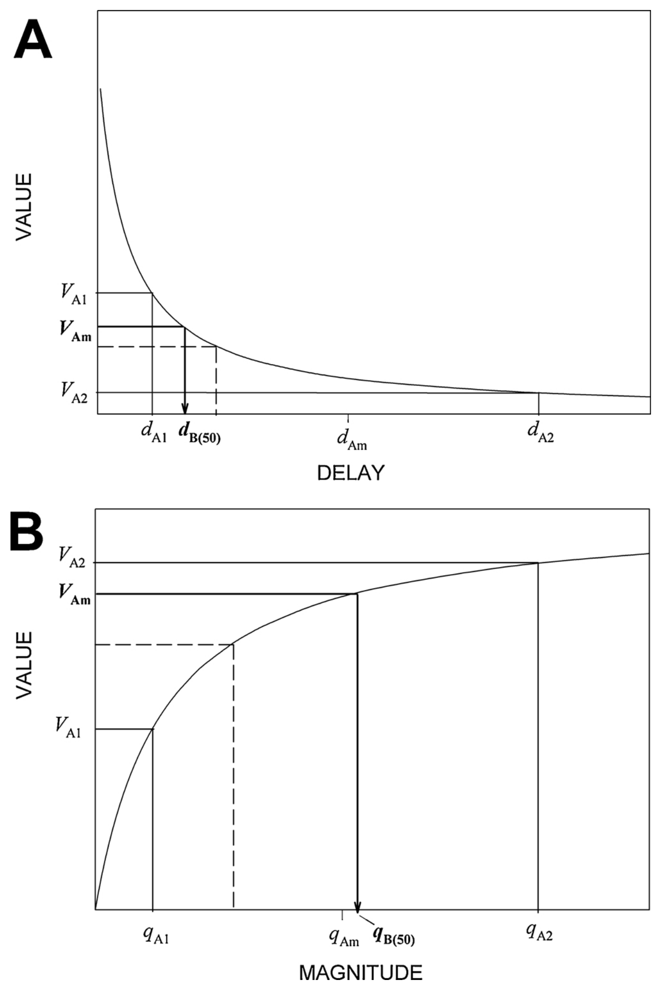

Fig. 5. Hypothetical relations between reinforcer value and delay and magnitude of reinforcement, illustrating the implications of MHM for risky choice in the adjusting-delay and adjusting-magnitude schedules. Curves are defined by hyperbolic functions (Eq. (3)). **A.** Adjusting-delay schedule (after Mazur, 1984). The continuous straight lines indicate the values corresponding to the two delays of reinforcer A (d_{A1} and d_{A2}). The horizontal broken line indicates the arithmetic mean value of reinforcer A, and the broken drop line indicates the corresponding delay (i.e. the value of $d_{B(50)}$ predicted by SHM). The bold horizontal line indicates the upward displacement of the average value of A induced by the 'averaging parameter', a, and the bold drop line indicates the corresponding predicted value of $d_{B(50)}$. Note that MHM predicts a further leftward displacement of $d_{B(50)}$ from the mean delay to reinforcer A (d_{Am}) than that predicted by SHM (i.e. an enhancement of risk proneness). **B.** Adjusting-magnitude schedule (da Costa Araújo et al., 2010). The continuous straight lines indicate the values corresponding to the two magnitudes of reinforcer A (q_{A1} and q_{A2}). The horizontal broken line indicates the arithmetic mean value of reinforcer A, and the broken drop line indicates the corresponding reinforcer magnitude, which lies to the left of the mean magnitude of reinforcer A (q_{Am}) (i.e., risk aversion). The bold horizontal line indicates the upward displacement of the average value of A induced by the 'averaging parameter', a, and the bold drop line indicates the corresponding predicted value of $q_{B(50)}$. Note that MHM predicts that the effect of a is to induce a rightward displacement of $q_{B(50)}$ (i.e. an attenuation of risk aversion). (Reproduced from Bradshaw, 2018.).

2018; Essock and Rees, 1974; Mazur, 1985; Mobini et al., 2000, 2002; Young, 1981), risk-insensitivity (Reboreda and Kacelnik, 1991; Staddon and Innis, 1966; Wunderle and O'Brien, 1985) and risk-aversion (Bateson and Kacelnik, 1995; Clements, 1990; Hastjarjo et al., 1990; Menlove et al., 1979) have all been reported. This is unexpected from the point of view of SHM, which predicts risk insensitivity (Eq. (11)). It is also unexpected from the point of view of MHM in its original form, which predicts risk aversion (Eq. (12a); see left-hand panel of Fig. 4). This problem may be resolved by recourse to the generalized mean as an averaging principle in the computation of the overall value of complex reinforcers (see above, section 7). Fig. 5 (Bradshaw, 2018) shows how the averaging parameter interacts with K and Q in determining the overall value of a complex reinforcer. In the case of a complex reinforcer comprising two delays (graph A), the effect of K is to shift $d_{B(50)}$ leftwards; thus $d_{B(50)} < d_{Am}$, (risk proneness). The effect of the averaging parameter, a, is to bias the average value of A, V_{Am}, upwards, further reducing $d_{B(50)}$, i.e. enhancing risk proneness. In the case of a complex reinforcer comprising two reinforcer magnitudes (graph B), the effect of Q is to shift $q_{B(50)}$ leftwards, in the direction of risk-proneness, this effect being greater in the case of lower values of Q. However, in this case the upward displacement of V_{Am} induced by a has the effect of displacing $q_{B(50)}$ rightwards, attenuating the effect of Q and shifting preference in the direction of risk aversion. Thus, while risk proneness is always to be expected when delays are combined in a complex reinforcer, the averaging of reinforcer magnitudes may produce risk proneness, risk insensitivity or risk aversion, depending on strengths of influence of Q and a.[7]

Risk sensitivity has an important status in models of choice based on evolutionary principles (Kacelnik and Bateson, 1996; McNamara and Houston, 1992; Real and Caraco, 1986), but bears no fixed relation to the parameters of MHM. The concept of 'risk' is generally taken to imply an awareness of, or sensitivity to, probability. As we have seen in the previous section, the current version of MHM has dispensed with the notion of probability as a discriminandum, replacing it with the notion of biased averaging of value. Nevertheless, MHM may help to explain why some investigations of risk sensitivity using operant techniques have yielded seemingly inconsistent or paradoxical results.

9. From value to response strength

Ultimately, the viability of any behavioural model composed of mathematically defined intervening variables depends upon the sturdiness of its links to external events and overt behaviour. Links to the external world are generally defined by the conditions of the experiments used to build and test the model, and therefore present no major theoretical difficulties. However, links to behaviour often present significant practical and theoretical obstacles. Indeed, it is arguable that its inability to forge firm links between its intervening variables and overt behaviour was the undoing of the most ambitious of all quantitative models of voluntary behaviour, Hull's (1943) *Principles of Behavior*.

Many contemporary models of choice behaviour, including MHM, circumvent the problem of linking value to behaviour by the use of null methods. These methods make it possible to construct indifference equations from which value is expunged, allowing direct access to the subordinate intervening variables (K, Q, etc.). Null methods are extremely powerful and versatile, and it is fair to say that most of the insights into choice behaviour that form the basis of models such as SHM and MHM derive from the application of these methods. Nevertheless, insofar as intervening variables like value are seen to be essential to the construction of such models, there is a very present need for robust links between these variables and overt behaviour. Without such links, concepts like value and utility will always be vulnerable to accusations of circularity.

One obvious difficulty in identifying reliable links between value and behaviour is the wide range of behavioural measures that have been used to assess the effects of rewarding outcomes. The concept of

[7] The logic of biased averaging based on the generalized mean can, of course, be extended to 'risky choice' situations in which reinforcers are omitted entirely. For example, if response A is reinforced in only one out of two free-choice trials in an adjusting-magnitude schedule, the overall value of the complex reinforcer is $[(V_A{}^a + 0)/2]^{1/a}$.

response strength offers a means of reducing the number of dependent variables with which a theory of value may need to contend (Herrnstein, 1970; Killeen and Hall, 2001; Skinner, 1938). While Skinner and innumerable behaviour analysts after him favoured response rate as a convenient measure of the 'strength' of an operant, many other measures, including latency, probability, force, and persistence in extinction, have also been advocated. The usefulness of response strength as an intervening variable depends upon the functional equivalence of these measures. Killeen and Hall (2001) examined the intercorrelations of several of these measures and concluded that a single intervening variable, response strength, accounted for a large proportion of the variance of these measures in a number of free-operant schedules. Skinner's advocation of response rate as a reliable measure of response strength was vindicated by the exceptionally high correlation between this dependent variable and the principal component.

Might response strength help MHM to bridge the gap between value and behaviour? Although it is too early to answer this question with confidence, there are some interesting pointers as to how the response strength concept might be put to good use in the further development of MHM. As it stands, MHM aspires to provide a finite set of equations that relate the physical attributes of reinforcing outcomes to the hypothetical entity, value. These equations, which have the hyperbolic form of Eq. (4b), are represented by the arrow labelled I in Fig. 6, connecting the physical attributes of reinforcers, $\phi_{i,\ j\ldots}$, with value, V. Value, in turn, may be regarded as a determinant of response strength, RS, via an as yet unspecified intervening function (II). RS is expressed by a variety of behavioural measures, represented as $B_{1,2\ \ldots}$ in the figure.[8] An additional type of relation is indicated by the function labelled III in the figure: the relation between specific attributes of the reinforcing outcome and indices of response strength, such as response rate. Although not incorporated into the current version of MHM, relations of this type are well established in the quantitative analysis of behaviour, perhaps the best known exemplar being Herrnstein's (1970) hyperbolic response strength equation. A typical application of this equation to free-operant performance takes response rate, R, maintained on variable-interval schedules, as the behavioural manifestation of response strength, and rate of reinforcement, r, as the independent variable. There is abundant evidence that operant performance in this paradigm can be well described by the following hyperbolic equation (Herrnstein, 1970):

$$R = \frac{R_{max} \cdot r}{K_H + r} \quad \text{or} \quad R = R_{max} \cdot \frac{1}{1 + K_H/r}\ , \tag{16}$$

where $R_{\max}$ is a parameter expressing the maximum response rate and K_{H} is the reinforcement rate corresponding to $R_{\max}/2$ (Bradshaw and Szabadi, 1988, 1989, 1993; de Villiers, 1977; de Villiers and Herrnstein, 1976; Herrnstein, 1970). In a comprehensive review of the literature up to 1977, de Villiers (1977) found that Eq. (16) provided an excellent account of data derived from experiments employing a variety of reinforcer variables, including magnitude, intensity and immediacy, and response measures, including rate of free-operant responding, response latency and running speed. In the light of such evidence it may be appropriate to express Eq. (16) in a more general form:

$$RS = \eta \cdot \frac{1}{1 + \Psi/\phi}\ , \tag{16a}$$

where ϕ is an objective attribute of the reinforcer expressed in physical units, η is the maximum response strength, Ψ is the value of ϕ corresponding to $\eta/2$.

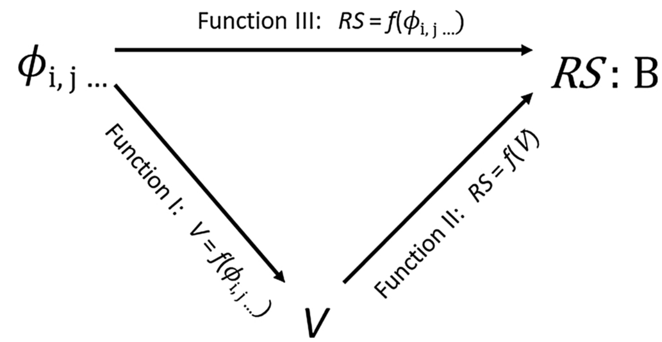

Fig. 6. Theoretical links between the physical attributes of reinforcers ($\phi_{i,j,\ldots}$), value (V) and response strength (RS). Function I is a hypothetical hyperbolic equation typified by Eq. (4b); function III is Herrnstein's (1970) hyperbolic response strength equation (Eq. (16)); function II is assumed also to be hyperbolic (Eq. (17)) for reasons given in the text. The relation between RS and overt behaviour is unknown, but is assumed in this paper, for the sake of simplicity, to be simply proportional.

It is now possible to deduce the form of the intervening function II: $RS = f(V)$. The arrangement of the three functions shown in Fig. 6 has been the subject of detailed analysis by Black and Leff (1983) in the context of classical pharmacological receptor theory. These authors showed that if functions I and III are hyperbolic, function II must be either linear or hyperbolic (see also Paton and Rothschild, 1965, for a related proof). We may apply these authors' logic to the definition of V and RS. We take Eq. (16a) as the generic response strength equation (function III), and a general form of Eq. (3) as the theoretical definition of V (function I):

$$V = \frac{1}{1 + \Phi/\phi}. \tag{3a}$$

Rearranging the terms of Eq. (3a), we get

$$\phi = \frac{V.\Phi}{1 - V}\ , \tag{3b}$$

and substitution of Eq. (3b) into Eq. (16a) then yields

$$RS = \frac{\eta.\ \Phi.\mathrm{V}}{\Psi + V\,(\Phi - \Psi)}\ . \tag{17}$$

This is a hyperbolic expression relating RS and V. Fig. 7 shows three patterns of effect of value on response strength expressed by Eq. (17). These three patterns are defined by the numerical relation between the two sensitivity parameters Φ and Ψ. When $\Phi > \Psi$ (graph A), RS is an increasing negatively accelerated function of V. Equality of Φ and Ψ (graph B) defines the only circumstance in which the relation between RS and V is not hyperbolic: in this case, RS is linearly related to V, the slope of the function being determined by η. When $\Phi < \Psi$ (graph C), RS is an increasing positively accelerated function of V. Black and Leff opined that the B and C scenarios were implausible in any pharmacological system and therefore concluded that Φ was necessarily greater than Ψ. Although it seems intuitively reasonable to make the same assumption in the develpment of MHM, it may be prudent to keep all three options on the table pending further empirical evidence. In particular, it will be important to establish whether alternative B can be excluded, because this option renders the concept of response strength more or less superfluous in the development of MHM.

The two parameters of the response strength equation (Eq. (16)) are differentially sensitive to 'response-related' and 'reinforcer-related' manipulations. Response-related manipulations such as changes in response-force requirement generally alter $R_{\max}$ without affecting K_{H} (Bradshaw et al., 1981a, 1983a; see below, section 10), whereas reinforcer-related manipulations such as changes in the volume of a sucrose reinforcer or the level of food deprivation affect K_{H} without

[8] We assume, for the sake of simplicity, that the response measures (rate, force, etc) are directly proportional to RS. This allows the chain connecting ϕ to B to be composed of two links: $V = f(\phi)$ and $RS = f(V)$, with $B \propto RS$. The addition of a third link, $B = f(RS)$, may prove necessary in due course. However for the present purposes, an additional intervening step would complicate the picture without affecting the argument that follows.

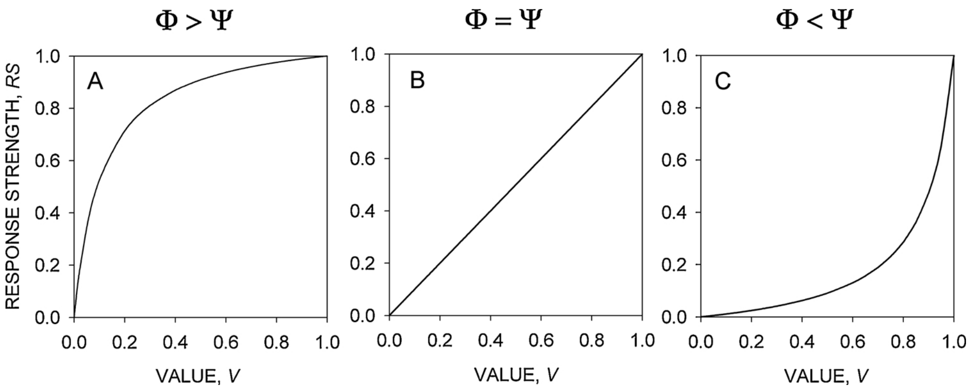

Fig. 7. Possible forms of the function relating value (V) and response strength (RS) (Eq. (17)). The three graphs illustrate the theoretically possible relations, based on the difference between the two sensitivity parameters Φ and Ψ. **A.** If $\Phi > \Psi$, RS is an increasing negatively accelerated hyperbolic function of V. **B.** If $\Phi = \Psi$, RS is directly proportional to V. **C.** if $\Phi < \Psi$, RS is a positively accelerated hyperbolic function of V. It is argued that the function shown in **A** is the most plausible alternative.

affecting R_{max} (Bradshaw et al., 1981b, 1983b; Heyman and Monaghan, 1987). There are, however, some notable exceptions; in particular, there are several reports of changes in some reinforcer-related variables affecting R_{max} (Bradshaw et al., 1978a, 1978b; Dallery et al., 2000; McDowell and Dallery, 1999; Snyderman, 1983; Shah et al., 1991). These findings are difficult to reconcile with the original interpretation of this parameter as the maximum response rate or the totality of emitted behaviour (Herrnstein, 1970, 1974; for reviews, see McDowell, 2013, and Olarte-Sánchez et al., 2015). The inclusion of the efficacy parameter ε into the value equation (Eq. (7)) may allow these problematic results to be explained away; however, empirical support for the inclusion of this parameter will be needed before this suggestion can be taken seriously (see above, section 6).

The formal link between value and response strength proposed in this section presupposes a hyperbolic response strength equation (Herrnstein, 1970) that is applicable to all behavioural measures of 'response strength.' In the case of response rate, the case for a hyperbolic function is supported by an abundance of empirical evidence (see above for references). Other response measures, including time allocation (Beardsley and McDowell, 1992; Billington and DiTomasso, 2003), initiation of response 'bouts' (Shull, 2011), running speed and reciprocal response latency (for references and re-analysis, see de Villiers, 1977; de Villiers and Herrnstein, 1976), may also yield to the descriptive power of Herrnstein's (1970) equation; however a systematic validation of the equation on a broad range of indices of response strength has not yet been undertaken.

Before leaving this topic, it is worth noting that, unlike value (as defined by MHM), response strength, as defined above, is not dimensionless. In Eq. (16), R and R_{max} are expressed in the same units, as are RS and η in the generalized form of Herrnstein's equation (Eq. (16a)).[9] This suggests that response strength may be an appropriate point of contact between MHM and quantitative models of electrical brain stimulation reinforcement in which action potential frequency in the medial forebrain bundle (MFB) is postulated to constitute a 'common currency' that may mediate the effectiveness of qualitatively different 'natural' reinforcers such as food and water, as well as direct stimulation of the MFB (Breton et al., 2014; Gallistel and Leon, 1991; Shizgal, 1997). In the current terminology of MHM, activity in the MFB cannot be an expression of value, which is dimensionless; it may, however, be thought of as a manifestation of response strength, measurable in physiological units (spikes s^{-1}).

10. Negative value, punishment and response cost

Ho et al. (1999) suggested that aversive consequences of responding could be represented in MHM as 'negative value', V^-, which derives from the mutiplicative combination of a set of hyperbolic functions that quantify the impact of the delay, magnitude, and other unspecified aspects of the aversive event (Estle et al., 2007). According to Ho et al. (1999) the overall value of a compound outcome consisting of a positive and a negative consequence is the algebraic sum of V^+ and V^-. Although this is a plausible account of the combination of positive and negative values in some situations, it is not the only way in which positive and negative influences may interact under the rubric of MHM. In this section, two types of interaction will be considered: those in which an aversive event reduces the overall value of an outcome (*value reduction*), and those that reduce response strength via an interaction with the parameters of Eq. (17) (*response strength reduction*).

Taking value reduction first, the interaction envisaged by Ho et al. (1999) may be illustrated by a compound outcome consisting of a reward of size q^+ delivered after a delay d^+, and a punisher of size q^- delivered after a delay d^-. This resembles the classic approach-avoidance conflict situation which was studied in great depth by Miller (1959). Miller's favoured experimental paradigm was the 'conflict alley' in which rats ran towards a goal box in which they received a food reward, an electric shock, or both. When approach and avoidance were assessed separately, strength of pull towards the goal increased as a function of the subjects' nearness to the goal, while strength of pull away from the goal also increased as a function nearness to the goal (Brown, 1948; Bugelski and Miller, 1938). When the reward and punishment were presented together, the subjects tended to advance some way towards the goal and then stop – the 'conflict point' (Miller et al., 1959). Miller's interpretation of these results and many similar findings was based on an extension of Hull's (1938) goal gradient hypothesis, in which an approach gradient (strength of approach plotted against nearness to the goal) intersected with a steeper avoidance gradient (strength of retreat plotted against nearness to the goal). For the sake of explication, Miller assumed linear gradients; however he took pains to point out that this was an over-simplification, and that the true form of the gradients awaited further empirical and theoretical analysis (Miller, 1959).

Fig. 8 shows how Ho et al.'s (1999) postulate of opposing positive and negative values may be applied to approach-avoidance conflict. Functions defining V^+ and V^- (based on Eq. (4b)) replace Miller's linear gradients:

$$V^+ = \frac{1}{1 + Q^+/q^+}\cdot\frac{1}{1 + K^+/i^+};\quad V^- = \frac{1}{1 + Q^-/q^-}\cdot\frac{1}{1 + K^-/i^-};$$
$$V = V^+ - V^-. \tag{18}$$

In order for the model to generate a 'conflict point' at which $V = 0$, $K^- \neq K^+$. The data amassed by Miller and his colleagues point towards steeper gradients for avoidance than for approach (Miller, 1959). This suggests that in the case of the appetitive and aversive stimili used in their experiments, $K^- > K^+$.

The above analysis brings Miller's (1959) insights within the ambit of MHM. It is also relevant to a wider range of phenomena, including

[9] Relative response strength, RS_{rel}, defined as RS/η in Eq. (16a), or R/R_{max} in Eq. (16), is dimensionless.

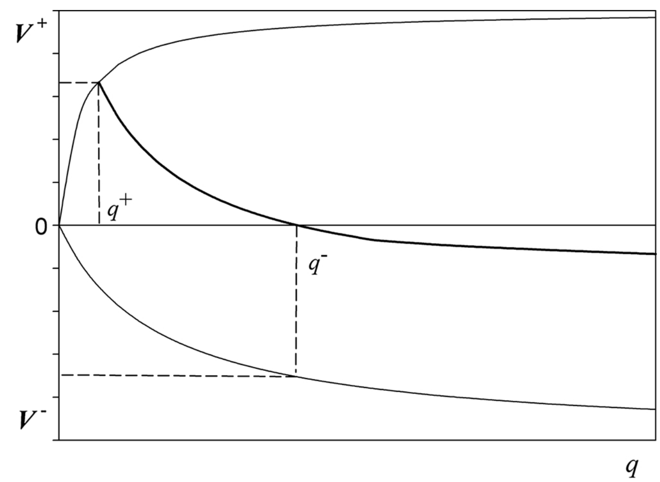

Fig. 8. Arithmetic combination of positive and negative values (V^+ and V^-) defined by hyperbolic value functions (the graph shows the simplest version, in which the size/value function is defined by Eq. (3)). The thick continuous line shows the algebraic sum of a fixed positive value and increasing negative values, which intersects with the zero line when $V^+ = V^-$ (analagous to the 'conflict point' defined by Miller, 1959).

those situations, commonly encountered in the world outside the laboratory, where outcomes may be composed of positive and negative consequences presented at different times after a choice response (as in the familiar 'buy now, pay later' contingency).

The competing influences of positive and negative consequences are also felt in any situation in which a commodity is acquired by effort or expense. This is often referred to as 'response cost'; however this label may obscure an important distinction between situations in which the 'cost' is incurred as a consequence of a prior choice response and situations in which the cost is paid per unit response. In the former case, the cost may be perceived as negative value which detracts from the overall value of the outcome as defined by Eq. (18) (i.e., 'value reduction'). In the latter case, however, the cost may be seen as a penalty for responding, the impact of which is likely to be felt in the response strength equation (Eq. (16)) (i.e., 'response strength reduction'). Examples of these two types of 'response cost' are shown in Fig. 9 which summarises the results of two experiments in which human subjects responded for monetary reinforcement on variable-interval schedules specifying a broad range of reinforcement rates (Bradshaw et al., 1977, 1978b). Each experiment included one condition in which no punishment was imposed and another condition in which punishment consisting of monetary penalties was superimposed on the positive reinforcement schedules. In one experiment, (Bradshaw et al., 1977) punishment was imposed according to a variable-ratio schedule, whereas in the other experiment (Bradshaw et al., 1978b) it was imposed according to a variable-interval schedue. In both cases, the punishment contingency resulted in an increase in the reinforcement rate corresponding to the half-maximal response rate (K_H), suggesting that the punishment conditon had imposed 'negative value' which detracted from the 'positive value' of the reinforcer. However, in the case of the variable-ratio punishment condition, the maximum response rate (R_{max}) was also greatly reduced, suggesting that the cost per unit response specified by the variable-ratio punishement schedule may have had an additional depressant effect on response strength.

11. Conclusions

MHM started life as an extension of SHM (Mazur, 1987), a highly influential and successful model of inter-temporal choice. By incorporating an explicit statement of the relation between the size of a reinforcer and its psychological value, MHM brought together the twin principles of hyperbolic delay discounting and the law of diminishing marginal utility. This paper has considered various ways in which MHM might be expanded, and has pointed out some of its strengths and weaknesses. Perhaps its greatest strength is its ability to accommodate a range of behavioural phenomena, a range that this paper has attempted to extend. However the cost of this added strength is the intrusion of a number of new parameters that were not envisaged in the original version of MHM. It has been argued that the existence of some of these intruders might be justified by future experiments, but there can be

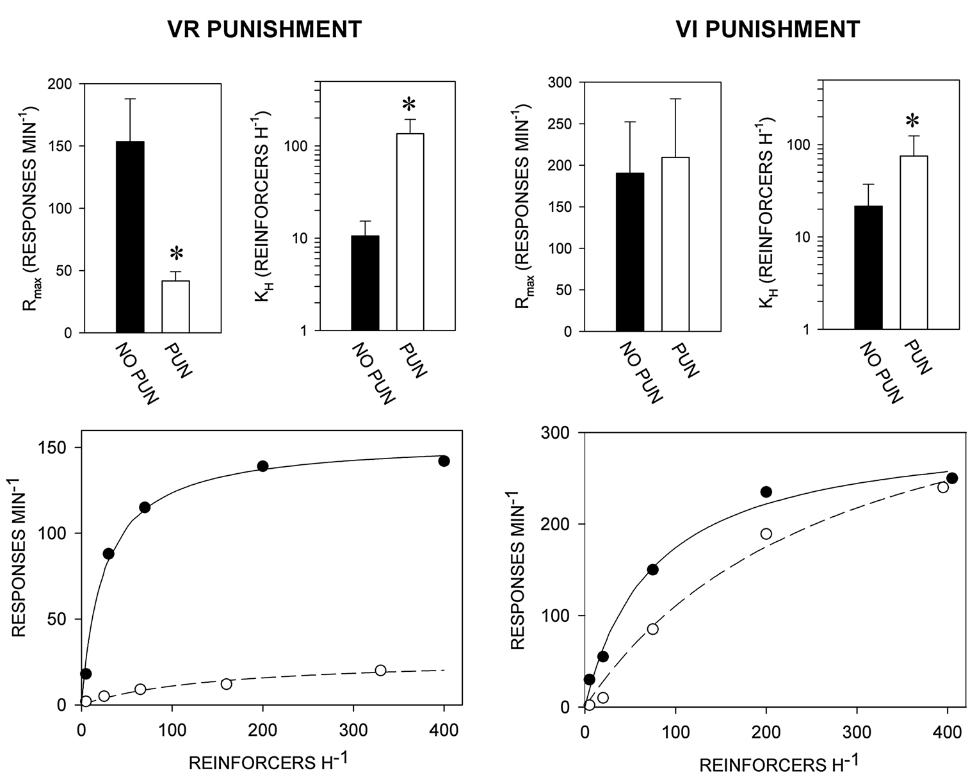

Fig. 9. Effect of two punishment schedules on the parameters of Herrnstein's (1970) response strength equation (mean data from Bradshaw et al., 1977, 1978b). Human subjects earned monetary reinforcement by pressing buttons under variable-interval schedules specifying a broad range of reinforcement rates. Hyperbolic functions were fitted to the response-rate data, and estimates of R_{max} (the maximum response rate) and K_H (the reinforcement rate corresponding to $R_{max}/2$) were obtained (upper panels of the figure). In each experiment, a punishment contingency was superimposed upon the positive reinforcement schedules in one condition of the experiment; Punishment (subtraction of monetary earnings) was delivered on a variable-ratio schedule in one experiment, and on a variable-interval schedule in the other experiment. Both punishment schedules induced a substantial increase in K_H. Variable-ratio punishment, but not variable-interval punishment, also induced a substantial reduction of R_{max}. The graphs in the lower part of the figure show response-rate data from a representative subject from each experiment.

little doubt that their presence will be a persistent source of difficulty in attempts to subject the model to decisive experimental test.

As presently formulated, MHM is a descriptive model; it is not prescriptive, teleological or 'normative'. Unlike many economic models, it makes no assumptions about the 'rationality' or 'adaptiveness' of decisions, and unlike many models based on evolutionary theory it makes no assumptions about animals' inherited predisposition to maximize their chances of short- or long-term survival or their reproductive fitness. By constructing a web of intervening variables and functional relations, it aspires to produce a quantitative descriptive account of choice behaviour and to generate testable predictions about choices in controlled situations. Thus, in proposing that an animal faced with a choice between two options will select the option with the higher value, MHM and other models like it are not offering an insight into the psyche of animals, but simply stating a defining feature of the intervening variable 'value'. The usefulness of this intervening variable, and the host of subordinate intervening variables that help to define it (K, Q, etc.), resides in their ability to classify and quantify functional relations. In this respect, 'value' is no different from a concept like 'thirst', which enables multiple environmental and biological influences to be tied to multiple behavioural phenomena (Hinde, 1970; Killeen and Hall, 2001; Miller, 1959).

However, as noted by MacCorquodale and Meehl (1948), science has a habit of gravitating towards reductionist explanations, to seek explanations for complex phenomena in terms of more fundamental principles. A concept that starts life as a puristic intervening variable may in the course of time find itself promoted to the status of a hypothetical construct and, eventually, to that of a physical entity. For MHM, value remains an intervening variable, but a rapidly growing body of evidence for reliable links between value and specific neural structures and neurophysiological events (Body et al., 2017; Breton et al., 2014; Rolls, 2014; Schultz, 2015; Shizgal, 1997; Smith et al., 2018) suggests that its promotion to the status of a hypothetical construct may be in order.

Acknowledgement

The author is immensely grateful to Peter R. Killeen for stimulating discussions about the topics discussed in this paper and for his wise advice about the improvement of the first draft.

References

Ainslie, G.W., 1974. Impulse control in pigeons. J. Exp. Anal. Behav. 21, 485–489.
Ariens, E.J., Simonis, A.M., van Rossum, J.M., 1964. Drug-receptor interaction: interaction of one or more drugs with one receptor system. In: Ariens, E.J. (Ed.), Molecular Pharmacology: The Mode of Action of Biologically Active Compounds. Academic, New York, pp. 119–286.
Arvanitogiannis, A., Shizgal, P., 2008. The reinforcement mountain:allocation of behavior as a function of the rate and intensity of rewarding brain stimulation. Behav.Neurosci 122, 1126–1138.
Bateson, M., Kacelnik, A., 1995. Preferences for fixed and variable food sources: variability in amount and delay. J. Exp. Anal. Behav. 63, 313–329.
Beardsley, S.D., McDowell, J.J., 1992. Appliction of Herrnstein's hyperbola to time allocation of naturalistic human behavior maintained by naturalistic social reinforcement. J. Exp. Anal. Behav. 57, 177–185.
Bezzina, G., Cheung, T.H.C., Asgari, K., Hampson, C.L., Body, S., Bradshaw, C.M., Szabadi, E., Anderson, I.M., Deakin, J.F.W., 2007. Effects of quinolinic acid-induced lesions of the nucleus accumbens core on inter-temporal choice: a quantitative analysis. Psychopharmacology 195, 71–84.
Bezzina, G., Cheung, T.H.C., Body, S., Deakin, J.F.W., Anderson, I.M., Bradshaw, C.M., Szabadi, E., 2009. Quantitative analysis of the effects of lesions of the subthalamic nucleus on inter-temporal choice: further evidence for enhancement of the incentive value of food reinforcers. Behav. Pharmac. 20, 437–446.
Billington, E., DiTomasso, N.M., 2003. Demonstrations and applications of the Matching Law in education. J. Behav. Educ. 12 981–104.
Black, J.W., Leff, P., 1983. Operational models of pharmacological agonism. Proc. Roy. Soc. Lond. B. 220, 141–160.
Body, S., Bradshaw, C.M., Szabadi, E., 2017. Delayed reinforcement: neuroscience. In: Stein, J. (Ed.), Reference Module in Neuroscience and Biobehavioral Psychology. Elsevier, Amsterdam. https://doi.org/10.1016/B978-0-12-809324-5.02713-9.
Breton, Y.-A., Conover, K., Shizgal, P., 2014. The effect of probability discounting on reward seeking: a three-dimensional perspective. Front. Behav. Neurosci. 8, 1–13 art. 284.
Bradshaw, C.M., 2017. Some properties of an adjusting-magnitude schedule of reinforcement: implications for models of choice. Behav. Proc. 140, 19–32.
Bradshaw, C.M., 2018. Effects of sucrose concentration on choice in an adjusting-magnitude schedule. Behav. Proc. 157, 562–569.
Bradshaw, C.M., Szabadi, E., 1988. Quantitative analysis of human operant behavior. In: Davey, G.C.L., Cullen, C. (Eds.), Human Operant Conditioning and Behavior Modification. Wiley, Chichester pp 225 259.
Bradshaw, C.M., Szabadi, E., 1989. Central neurotransmitter systems and the control of operant behaviour by "natural" positive reinforcers. In: Liebman, J., Cooper, S.J. (Eds.), The Neuropharmacological Basis of Reward. Oxford University Press, New York, pp. 320–376.
Bradshaw, C.M., Szabadi, E., 1992. Choice between delayed reinforcers in a discrete-trials schedule: the effect of deprivation level. Q. J. Exp. Psychol. 44B, 1–16.
Bradshaw, C.M., Szabadi, E., 1993. Herrnstein's equation: data from 110 rats. Psychol. Rep. 73, 1355–1361.
Bradshaw, C.M., Ruddle, H.V., Szabadi, E., 1981a. Studies of concurrent performances in humans. In: Bradshaw, C.M., Szabadi, E., Lowe, C.F. (Eds.), Quantification of Steady-State Operant Behaviour. Elsevier/North Holland, Amsterdam, pp. 79–90.
Bradshaw, C.M., Ruddle, H.V., Szabadi, E., 1981b. Relationship between response rate and reinforcement frequency in variable-interval schedules: II effect of the volume of sucrose reinforcement. J. Exp. Anal. Behav. 35, 262–269.
Bradshaw, C.M., Szabadi, E., Bevan, P., 1977. Effect of punishment on human variable-interval performance. J. Exp. Anal. Behav. 27, 275–279.
Bradshaw, C.M., Szabadi, E., Bevan, P., 1978a. Relationship between response rate and reinforcement frequency in variable-interval schedules: the effect of the concentration of sucrose reinforcement. J. Exp. Anal. Behav. 29, 447–452.
Bradshaw, C.M., Szabadi, E., Bevan, P., 1978b. Effect of variable-interval punishment on the behavior of humans in variable-interval schedules of monetary reinforcement. J. Exp. Anal. Behav. 29, 161–166.
Bradshaw, C.M., Szabadi, E., Ruddle, H.V., 1983a. Herrnstein's equation: effect of response-force requirement on performance in variable-interval schedules. Behav. Anal. Lett. 3, 93–100.
Bradshaw, C.M., Szabadi, E., Ruddle, H.V., Pears, E., 1983b. Herrnstein's equation: effect of deprivation level on performance in variable-interval schedules. Behav. Anal. Lett. 3, 267–273.
Brown, J.S., 1948. Gradients of approach and avoidance responses and their relation to level of motivation. J. Comp. Physiol. Psychol. 41, 450–465.
Bugelski, R., Miller, N.E., 1938. A spatial gradient in the strength of avoidance responses. J. Exp. Psychol. 23, 494–505.
Chipman, J.S., 1960. The foundations of utility. Econometrica 28, 193–224.
Cicerone, R.A., 1976. Preference for mixed versus constant delay of reinforcement. J. Exp. Anal. Behav. 25, 257–261.
Clements, K.L., 1990. Risk-aversion in the foraging blue jay, Cyanocitta cristala. Anim. Behav. 40, 182–195.
Conover, K.L., Shizgal, P., 1994. Competition and summation between rewarding effects of sucrose and lateral hypothalamic stimulation in the rat. Behav. Neurosci. 108, 537–546.
da Costa Araújo, S., Body, S., Valencia-Torres, L., Olarte-Sánchez, C.M., Bak, V.K., Deakin, J.F.W., Anderson, I.M., Bradshaw, C.M., Szabadi, E., 2010. Choice between reinforcer delays versus choice between reinforcer magnitudes: differential Fos expression in the orbital prefrontal cortex and nucleus accumbens core. Behav. Brain Res. 213, 269–277.
Dallery, J., McDowell, J.J., Lancaster, J.S., 2000. Falsification of matching theory's account of single-alternative responding: herrnstein's k varies with sucrose concentration. J. Exp. Anal. Behav. 73, 23–43.
Davison, M.C., 1969. Preference for mixed-interval versus fixed-interval schedules. J. Exp. Anal. Behav. 12, 247–252.
de Villiers, P.A., 1977. Choice in concurrent schedules and a quantitative formulation of the law of effect. In: Honig, K., Staddon, J.E.R. (Eds.), Handbook of Operant Behavior. Prentice Hall, Englewood Cliffs, NJ.
de Villiers, P.A., Herrnstein, R.J., 1976. Toward a law of response strength. Psychol. Bull. 83, 1131–1153.
Essock, S.M., Rees, E.P., 1974. Preference for and effects of variable as opposed to fixed reinforcer duration. J. Exp. Anal. Behav. 21, 89–97.
Estle, S.J., Green, L., Myerson, J., Holt, D.D., 2007. Discounting of monetary and directly consumable rewards. Psychol. Sci. 18, 58–63.
Furchgott, R.F., Bursztin, P., 1967. Comparison of dissociation constants and of relative efficacies of selected agonists acting of parasympathetic receptors. Ann. N.Y. Acad. Sci. 144, 882–892.
Gallistel, C.R., Leon, M., 1991. Measuring the subjective magnitude of brain stimulation reward by titration with rate of reward. Behav.Neurosci. 105, 913–925.
Green, L., Myerson, J., 2004. A discounting framework for choice with delayed and probabilistic rewards. Psychol. Bull. 130, 769–792.
Hastjarjo, X., Silberberg, A., Hursh, S.R., 1990. Risky choice as a function of amount and variance in food supply. J. Exp. Anal. Behav. 53, 155–161.
Hausman, D.M., 1995. The impossibility of interpersonal utility comparisons. Mind 104, 473–490.
Hernandez, G., Breton, Y.-A., Conover, K., Shizgal, P., 2010. At what stage of neural processing does cocaine act to boost pursuit of rewards? PLoS One 5, e15081.
Herrnstein, R.J., 1970. On the law of effect. J. Exp. Anal. Behav. 13, 243–266.
Herrnstein, R.J., 1974. Formal properties of the matching law. J. Exp. Anal. Behav. 21, 159–164.
Heyman, G.M., Monaghan, M.M., 1987. Effects of changes in response requirement and deprivation on the parameters of the matching law equation: New data and review. J. Exp. Psychol.: Anim. Behav. Proc. 13, 384–394.
Hicks, J.R., Allen, R.G.D., 1934. A reconsideration of the theory of value. Part I. Economica, New Series 1, 52–76.

Hinde, R.A., 1970. Animal Behaviour, 2nd edition. McGraw-Hill, New York.
Ho, M.-Y., Mobini, S., Chiang, T.-J., Bradshaw, C.M., Szabadi, E., 1999. Theory and method in the quantitative analysis of "impulsive choice" behaviour: implications for psychopharmacology. Psychopharmacology 146, 362–372.
Hull, C.L., 1938. The goal gradient hypothesis applied to some "fieldforce" problems in the behavior of young children. Psychol. Rev. 45, 271–299.
Hull, C.L., 1943. Principles of Behavior. Appleton-Century-Crofts, New York.
Kable, J.W., Glimcher, P.W., 2007. The neural correlates of subjective value during intertemporal choice. Nature Neurosci. 10, 1625–1633.
Kacelnik, A., 2006. Meanings of rationality. In: Hurley, S., Nudds, M. (Eds.), Rational Animals? Oxford Universsity Press, Oxford.
Kacelnik, A., Bateson, M., 1996. Risky theories – the effects of variance on foraging decisions. Amer. Zool. 36, 402–434.
Kacelnik, A., El-Mouden, C., 2013. Triumphs and trials of the risk paradigm. Anim. Behav. 86, 1117–1129.
Kahneman, D., Sugden, R., 2005. Experienced utility as a standard of policy evaluation. Environ. Resour. Econ. Rev. 32, 161–181.
Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica 47, 263–291.
Keller, L., 2015. The problem with the concept of utility and its measurement. Pol. J. Polit. Sci. 1 (4), 6–44.
Kenakin, T., 1997. Pharmacologic Analysis of Drug-Receptor Interaction, 3rd edn. Lippincott-Raven, Philadelphia.
Kheramin, S., Body, S., Mobini, S., Ho, M.-Y., Velazquez-Martinez, D.N., Bradshaw, C.M., Szabadi, E., Deakin, J.F.W., Anderson, I.M., 2002. Effects of quinolinic acid-induced lesions of the orbital prefrontal cortex on intertemporal choice: a quantitative analysis. Psychopharmacology 165, 9–17.
Killeen, P.R., 1994. Mathematical principles of reinforcement. Behav. Brain Sci. 17, 105–172.
Killeen, P.R., 2011. Models of trace decay, eligibility for reinforcement, and delay of reinforcement gradients, from exponential to hyperboloid. Behav. Proc. 87, 57–63.
Killeen, P.R., 2015. The arithmetic of discounting. J. Exp. Anal. Behav. 103, 249–259.
Killeen, P.R., Hall, S.S., 2001. The principle components of response strength. J. Exp. Anal. Behav. 75, 1113–1134.
Leon, M., Gallistel, C.R., 1992. The function relating the subjective magnitude of brain stimulation reward to stimulation strength varies with the site of stimulation. Behav. Brain Res. 52, 183–192.
Leon, M., Gallistel, C.R., 1998. Self-stimulating rats combine subjective reward magnitude and subjective reward rate multiplicatively. J. Exp. Psychol.: Anim. Behav. Proc. 24, 265–277.
Levy, D.J., Glimcher, P.W., 2011. Comparing apples and oranges: using reward-specific and reward-general subjective value representations in the brain. J. Neurosci. 31, 14693–14707.
Locey, M.L., Dallery, J., 2009. Isolating behavioral mechanisms of inter-temporal choice: nicotine effects on delay discounting and delay sensitivity. J. Exp. Anal. Behav. 91, 213–224.
MacCorquodale, K., Meehl, P.E., 1948. On a distinction between hypothetical constructs and intervening variables. Psychol. Rev. 55, 95–107.
Martin, S.R., Williams, D.R., 2017. Outgrowing the Procrustean bed of normality: the utility of Bayesian modeling for asymmetrical data analysis. PsyArXiv, Aug 30, 1–22. https://doi.org/10.31234/osf.io/26m49.
Mazur, J.E., 1984. Tests of an equivalence rule for fixed and variable reinforcer delays. J. Exp. Psychol.: Anim. Behav. Proc. 10, 426–436.
Mazur, J.E., 1985. Probability and delay of reinforcement as factors in discrete-trial choice. J. Exp. Anal. Behav. 43, 341–351.
Mazur, J.E., 1987. An adjusting procedure for studying delayed reinforcement. In: Commons, M.L., Mazur, J.E., Nevin, J.A., Rachlin, H.C. (Eds.), Quantitative Analyses of Behavior, Vol. 5: The Effects of Delay and Intervening Events on Reinforcement Value. Erlbaum, Mahwah N.J, pp. 55–73.
Mazur, J.E., 2004. Risky choice: selecting between certain and uncertain outcomes. Behav. Analyst Today 5, 190–202.
Mazur, J.E., Herrnstein, R.J., 1988. On the functions relating delay, reinforcer value, and behavior. Behav. Brain Sci. 11, 690–691.
McClure, J., Podos, J., Richardson, H.N., 2014. Isolating the delay component of impulsive choice in adolescent rats. Front. Integ. Neurosci 8, 1–9 article 3.
McDowell, J.J., 2013. On the theoretical and empirical status of the matching law and matching theory. Psychol. Bull. 139, 1000–1028.
McDowell, J.J., Dallery, 1999. Falsification of matching theory: changes in the asymptote of Herrnstein's hyperbola as a function of water deprivation. J. Exp. Anal. Behav. 72, 251–268.
McNamara, J.M., Houston, A.I., 1992. Risk-sensitive foraging: a review of the theory. Bull. Math. Biol. 54, 355–378.
Menlove, R.L., Inden, H.M., Madden, E.G., 1979. Preference for fixed over variable access to food. Anim. Learn. Behav. 7, 499–503.
Miller, 1959. Liberalization of basic S-R concepts: extensions to conflict behavior, motivation, and social learning. In: Koch, S. (Ed.), Psychology: A Study of a Science; Volume 2. General Systematic Formulations, Learning, and Special Processes. McGraw Hill, New York.
Miller, N.E., Brown, J.S., Lewis, H., 1959. A Theoretical and Experimental Analysis of Conflict Behavior: Approach-avoidance Conflict As a Function of Strength of Drive and Strength of Shock (Unpublished; Described in Miller, 1959).
Mobini, S., Chiang, T.-J., Ho, M.-Y., Bradshaw, C.M., Szabadi, E., 2000. Effects of central 5-hydroxytryptamine depletion on sensitivity to delayed and probabilistic reinforcement. Psychopharmacology 152, 390–397.
Mobini, S., Body, S., Ho, M.-Y., Bradshaw, C.M., Szabadi, E., Deakin, J.F.W., Anderson, I.M., 2002. Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and probabilistic reinforcerment. Psychopharmacology 160, 290–298.
Moschak, T.M., Mitchell, S.H., 2014. Partial inactivation of nucleus accumbens core decreases delay discounting in rats without affecting sensitivity to delay or magnitude. Behav. Brain Res. 268, 159–168.
Myerson, J., Green, L., 1995. Discounting of delayed rewards: models of individual choice. J. Exp. Anal. Behav. 64, 263–272.
Olarte-Sánchez, C.M., Valencia-Torres, L., Cassaday, H.J., Bradshaw, C.M., Szabadi, E., 2015. Quantitative analysis of performance on a progressive-ratio schedule: effects of reinforcer type, food deprivation and acute treatment with Δ^9-tetrahydrocannabinol (THC). Behav. Proc. 113, 122–131.
Paton, W.D.M., Rothschild, A.M., 1965. The changes in response and in ionic content of smooth muscle produced by acetylcholine action and by calcium deficiency. Br. J. Pharmac. 24, 437–448.
Plummer, M., 2003. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. Hornik, K., Leisch, F., Zeileis, A. (Eds.), Proceedings of the 3rd International Workshop on Distributed Statistical Computing.
R Core Team, 2018. R: A Language and Environment for Stastical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Rachlin, H.C., 1971. On the tautology of the matching law. J. Exp. Anal. Behav. 15, 249–251.
Rachlin, H.C., 2006. Notes on discounting. J. Exp. Analysis Behav. 85, 425–435.
Rachlin, H.C., Raineri, A., Cross, D., 1991. Subjective probability and delay. J. Exp. Anal. Behav. 55, 223–244.
Read, D., 2004. Utility Theory From Jeremy Bentham to Daniel Kahneman. LSE Working Paper LSEOR 04-64. London School of Economics, London.
Real, L., Caraco, T., 1986. Risk and foraging in stochastic environments. Ann. Rev. Ecol. Syst. 17, 371–390.
Reboreda, J.C., Kacelnik, A., 1991. Risk sensitivity in starlings: variability in food amount and food delay. Behav. Ecol. 2, 301–308.
Richards, J.B., Mitchell, S.H., de Wit, H., Seiden, L.S., 1997. Determination of discount functions in rats with an adjusting amount procedure. J. Exp. Anal. Behav. 67, 353–366.
Richardson, J., 1994. Cost utility analysis: what should be measured? Soc. Sci. Med. 39, 7–21.
Rider, D.P., 1983. Preference for mixed versus constant delays of reinforcement: effect of probability of the short, mixed delay. J. Exp. Anal. Behav. 39, 257–266.
Robinson, J., 1962. Economic Philosophy. Penguin Books, Harmondsworth, U.K.
Rolls, E.T., 2014. Emotion and Decision-making Explained. Oxford University Press, Oxford.
Sakamoto, Y., Ishiguro, M., Kitagawa, G., 1986. *Akaike Information Criterion Statistics*. Reidel, Dordrecht, Holland.
Schultz, W., 2015. Neuronal reward and decision signals: from theories to data. Physiol. Rev. 95, 853–951.
Schwartz, G., 1978. Estimating the dimensions of a model. Ann. Statistics 6, 461–464.
Shah, K., Bradshaw, C.M., Szabadi, E., 1991. Relative and absolute reinforcement frequency as determinants of choice in concurrent variable-interval schedules. Q. J. Exp. Psychol. 43B, 25–38.
Shizgal, P., 1997. Neural basis of utility estimation. Curr. Opin. Neurobiol. 7, 198–208.
Shull, R.L., 2011. Bouts, changeovers, and units of operant behavior. Eur. J. Bejav. Anal. 12, 49–72.
Simmons, J.M., Gallistel, C.R., 1994. Saturation of subjective reward magnitude as a function of current and pulse frequency. Behav. Neurosci. 108, 151–160.
Shizgal, P., Conover, K., 1996. On the computation of utility. Curr. Direct. Psychol. Sci. 5, 37–43.
Skinner, B.F., 1938. The Behavior of Organisms: an Experimental Analysis. Appleton-Century-Crofts, New York.
Skinner, B.F., 1950. Are theories of learning necessary? Psychol. Rev. 57, 193–216.
Smith, B.J., Montarosso, J.R., Wakslak, C.J., Bechara, A., Read, S.J., 2018. A meta-analytic review of brain activity associated with intertemporal decisions: evidence for an anterior-posterior tangibility axis. Neurosci. Biobehav. Rev. 86, 85–98.
Snyderman, M., 1983. Bodyweight and response strength. Behav. Anal. Lett. 3, 255–265.
Staddon, J.E.R., Innis, N.K., 1966. Preference for fixed vs. Variable amounts of reward. Psychon. Sci. 4, 193–194.
Stan Development Team, 2016. RStan: The R Interface to Stan. R Package Version 2.14.1. Retrieved from. http://mc-stan.org/.
Valencia-Torres, L., da Costa Araújo, S., Olarte-Sánchez, C.M., Body, S., Bradshaw, C.M., Szabadi, E., 2011. Transitional and steady-state choice behavior under an adjusting-delay schedule. J. Exp. Anal. Behav. 95, 57–74.
Valencia-Torres, L., Olarte-Sánchez, C.M., Body, S., Bradshaw, C.M., Szabadi, E., 2013. Investigations of the neurobiological bases of inter-temporal choice. Eur. J. Behav. Anal. 14, 215–239.
Valencia-Torres, L., Olarte-Sánchez, C.M., da Costa Araújo, S., Body, S., Bradshaw, C.M., Szabadi, E., 2012. Nucleus accumbens and delay discounting in rats: evidence from a new quantitative protocol for analysing inter-temporal choice. Psychopharmacology 219, 271–283.
Villarreal, M., Velázquez, C., Baroja, J.L., Segura, A., Bouzas, A., Lee, M.D., 2019. Bayesian methods app;oed to the generalized matching law. J. Exp. Anal. Behav. 111, 252–273.
Wogar, M.A., Bradshaw, C.M., Szabadi, E., 1992. Choice between delayed reinforcers in an adjusting-delay schedule: the effects of absolute reinforcer size and deprivation level. Q. J. Exp. Psychol. 45B, 1–13.
Wunderle, J.M., O'Brien, T.G., 1985. Risk aversion in hand-reared bananaquits. Behav. Ecol. Sociobiol. (Print) 17, 371–380.
Young, J.S., 1981. Discrete-trial choice in pigeons: effects of reinforcer magnitude. J. Exp. Anal. Behav. 35, 23–29.