Rosenström, T., Wiesner, K., & Houston, A. I. (2016). Scalar utility theory and proportional processing: What does it actually imply? Journal of Theoretical Biology, 404, 222-235. DOI: 10.1016/j.jtbi.2016.06.003

www.elsevier.com/locate/yjtbi

Scalar utility theory and proportional processing: What does it actually imply?

Tom Rosenström, Karoline Wiesner, Alasdair I. Houston

PII: S0022-5193(16)30131-X
DOI: http://dx.doi.org/10.1016/j.jtbi.2016.06.003
Reference: YJTBI8691

To appear in: *Journal of Theoretical Biology*

Received date: 13 March 2016
Revised date: 31 May 2016
Accepted date: 1 June 2016

Cite this article as: Tom Rosenström, Karoline Wiesner and Alasdair I. Houston Scalar utility theory and proportional processing: What does it actually imply? *Journal of Theoretical Biology,* http://dx.doi.org/10.1016/j.jtbi.2016.06.003

Scalar utility theory and proportional processing: what does it actually imply?

Tom Rosenström[1,2*], Karoline Wiesner[3], Alasdair I Houston[1]

[1]School of Biological Sciences, University of Bristol, UK

[2]Institute of Behavioural Sciences, University of Helsinki, Finland

[3]School of Mathematics, University of Bristol, UK

[*]Corresponding author at: School of Biological Sciences, University of Bristol, UK

Email: tom.rosenstrom@helsinki.fi

Abstract

Scalar Utility Theory (SUT) is a model used to predict animal and human choice behaviour in the context of reward amount, delay to reward, and variability in these quantities (risk preferences). This article reviews and extends SUT, deriving novel predictions. We show that, contrary to what has been implied in the literature, (1) SUT can predict both risk averse and risk prone behaviour for both reward amounts and delays to reward depending on experimental parameters, (2) SUT implies violations of several concepts of rational behaviour (e.g. it violates strong stochastic transitivity and its equivalents, and leads to probability matching) and (3) SUT can predict, but does not always predict, a linear relationship between risk sensitivity in choices and coefficient of variation in the decision-making experiment. SUT derives from Scalar Expectancy Theory which models uncertainty in behavioural timing using a normal distribution. We show that the above conclusions also hold for other distributions, such as the inverse

Gaussian distribution derived from drift-diffusion models. A straightforward way to test the key assumptions of SUT is suggested and possible extensions, future prospects and mechanistic underpinnings are discussed.

Keywords

Scalar expectancy theory; scalar property; Weber's law; decision making; risk preferences

1. Introduction

Ernst Heinrich Weber (1795–1878) was one of the founders of psychophysics. Weber's law states that the resolution of perception diminishes in proportion to the magnitude of the stimulus. That is, if a just noticeable difference between a 10 kg weight and another weight is 10γ kg, where γ is a positive constant, then only differences exceeding 20γ kg can be detected when comparing a weight to a 20 kg weight. Drawing on a body of previous theory (e.g., Gibbon, 1977; Gibbon, Church, Fairhurst, & Kacelnik, 1988), Kacelnik & Brito e Abreu (1998) used Weber's law to propose that the representation of a stimulus of magnitude m in an animal's memory has an error distribution that is Normal with mean m and standard deviation γm, denoted $N(m, \gamma m)$; a theory now known as Scalar Utility Theory, or SUT (Kacelnik & El Mouden, 2013; Marsh & Kacelnik, 2002). The "scalar" parameter γ is a species and stimulus-type (but not stimulus-quantity) specific constant that captures the general resolution of perceptual memory. SUT is a generalization to both delays to reward and reward amounts of the original model for delays only, known as Scalar Expectancy Theory (Gibbon, 1977).

SUT has had some notable success in explaining effects of risk in decision making (Bateson & Kacelnik, 1995a, 1995b; Brito e Abreu & Kacelnik, 1999; Buhusi & Meck, 2005;

Kacelnik & Bateson, 1996; Reboreda & Kacelnik, 1991; Shafir, Reich, Tsur, Erev, & Lotem, 2008), but has received surprisingly little theoretical attention since the seminal work of Gibbon (1977) and Gibbon et al. (1988), despite the rapidly accumulating empirical data. Therefore, we provide an updated theoretical review and summary of the predictions of SUT and dispel some common misbeliefs regarding them. We specifically concentrate on the challenge posed for future research in a recent review of proportional processing: "… to model what behaviour influenced by proportional processing would look like …" (Akre & Johnsen, 2014).

A central goal in the study of animal and human behaviour has been to understand risk sensitivity in choice preferences (see Kacelnik & El Mouden, 2013; Rieskamp et al., 2006, for reviews). Some researchers approached the problem through functional explanations, deriving models that explain what kind of state-dependent risk sensitivity should evolve by natural selection (Caraco, Martindale & Whittam, 1980; Barnard et al., 1985; McNamara & Houston, 1992; Houston & McNamara, 1999). While theoretically justified, the models failed to explain the ubiquitous partial preferences of animals (variable choices in the same task despite the same conditions; McNamara &Houston 1987; Shapiro et al 2008). By reference to findings of psychophysics, SUT and its precursors were able to propose a proximate mechanism that intrinsically explains both the partial preferences and the findings on risk sensitivity reviewed below (Gibbon et al., 1988; Reboreda & Kacelnik, 1991; Kacelnik & El Mouden, 2013). However, SUT does not explain switches in risk sensitivity as a function of the external environment and the animal's internal state (Caraco, Martindale & Whittam, 1980; Houston & McNamara, 1999).

Researchers since Tinbergen (1963) have recognised the need to integrate the levels of explanation (Kacelnik & Bateson, 1996; Kacelnik & El Mouden, 2013; McNamara & Houston,

2009). Provided that our curiosity is not satisfied by a simple statement that all the levels are likely to play a role, such integration requires an understanding of the fundamental components to be integrated. This paper aims to contribute to this wider discussion by improving the level of theoretical understanding about the implications of SUT.

A paradigmatic application of SUT involves understanding an animal's behaviour when it is forced to choose between two options with the same arithmetic mean, one having no variance ("fixed" or "safe" option) and the other ("variable" or "risky" option) involving either a variable amount of reward or variable delay to reward (e.g. Bateson & Kacelnik, 1995b; Kacelnik & Brito e Abreu, 1998; Kacelnik & El Mouden, 2013). Simple arguments that identify the value of an option with its arithmetic mean suggest that animals should be indifferent in such experiments, but instead, they are typically found to favour the variable option for delays in reward and often prefer the fixed option for reward amounts (Kacelnik & Bateson, 1996; Kacelnik & El Mouden, 2013). It has been thought that this constitutes broad support for SUT, but despite many empirical studies, researchers have not carefully outlined the theoretical predictions of SUT, apparently thinking that the general trends of risk proneness for delays and risk aversion for amounts are what SUT predicts (Kacelnik & Bateson, 1996, 1997; Kacelnik & Brito e Abreu, 1998; Kacelnik & El Mouden, 2013). Here we outline the fuller flexibility of the model's predictions, providing a possible basis for resolution of conflicts, as well as new ways to support or disprove SUT experimentally. We also make connections with related literature.

This paper is organised in the following manner. First we explain and define the SUT model. Then we outline its, often surprising, predictions in a number of different contexts. Specifically, we consider previously misrepresented model predictions of SUT, different accounts of rational decision making in the context of SUT, and efficient ways to test SUT. After

this section on model predictions, we discuss extensions, future prospects, and mechanistic underpinnings of SUT, and then conclude the paper.

1.1. The model

1.1.1. Basic definitions

SUT is a model of perceptual memory's accuracy that is to be applied only after the animal's learning process can be considered as having stabilised in the sense that further experiences of the same experimental setting no longer change the expected behaviour (i.e., observed behaviour is stationary, but some learning processes may well be in operation). Pertaining to the paradigmatic application of SUT, Figure 1A shows how SUT assumes that the animal represents a choice option with fixed reward or delay in its memory, whereas Figure 1B summarises how representation of the variable option arises from a mixture of such simple representations. It is further assumed that the only way to access these representations is to draw a random sample from the associated distribution (Kacelnik & Brito e Abreu, 1998). Because Weber's law implies a higher memory variance (inaccuracy) for the random outcomes with the greater magnitude, the mixture distribution of the two equally likely outcomes is heavily skewed (thick line in Fig. 1B). Thus, when the animal draws independent samples from its memory representations for the fixed option (denoted S_F) and for the variable option (S_V), the random variable S_F is probably slightly higher than S_V for a given realisation, although more rarely, S_V is much higher (the joint sample is more likely to derive from the lower-right quadrant in Fig. 1C). Thus, one finds that $P(S_F > S_V) > ½$ despite the experimenter using equal mean values for both the choice alternatives. If S_F and S_V represent delays to reward, the animal prefers the lower option (most frequently the variable

one), and is thus risk prone; in contrast, if S_F and S_V are amounts of reward, the animal prefers the higher value (most frequently the fixed option), and so is risk averse (Kacelnik & Bateson, 1996, 1997; Kacelnik & El Mouden, 2013).

Because SUT assumes independent draws from the memory distributions, the probability density function of the joint memory distribution relevant for the experiment factorises as $f_{F,V} = f_F \times f_V$, and the integral over the lower diagonal in Figure 1C has a particularly convenient form:

$$P(S_F > S_V) = \int_{-\infty}^{\infty} f_F(x) F_V(x) dx, \qquad (1)$$

where $F_V(x) = P(S_V < x)$ is the cumulative distribution function of the variable option. SUT can be extended for more general applications as follows.

We consider a 'general experiment' where an animal can choose one of the finite number of options indexed by i, each having a set of associated rewards of values $m_i = \{m_{i,1}, \ldots, m_{i,k(i)}\}$ that occur with respective probabilities $\pi_i = \{\pi_{i,1}, \ldots, \pi_{i,k(i)}\}$, where $k(i)$ is the number of possible reward values ("outcomes") under alternative i. Here, SUT assumes that the animal has accurately learned the probabilities of stimuli (π's) but has perceptual memory, estimation and/or retrieval inaccuracy in their magnitudes (in m's), so that its representation for the outcome of option i is captured by the mixture probability density

$$f_i(x) = \sum_{j=1}^{k(i)} \pi_{i,j} N(x; m_{i,j}, \gamma m_{i,j}), \qquad (2)$$

where $N(\cdot;$ mean, s.d.) refers to a Normal density and γ is a representation noise, or error parameter (the scalar of "SUT"). This noise has a standard deviation $\gamma m_{i,j}$ that is a scalar multiple of the stimulus magnitude $m_{i,j}$.

Gibbon (1977) was aware of the possible problematic confounding of the distribution of π's and that of the stimulus-magnitude estimates as "both are produced by the subject" and he presented an analysis that "ignores this problem", but subsequent work has not considered it. Since the model has empirical merits, we also ignore this conceptual problem initially, but discuss it later on (in sections 3.1 and 3.2). Our aim is to derive the predictions of SUT as an empirical model; we do not claim that conscious estimates of stimulus probabilities are easier to make than estimates of stimulus quantity, although estimation of longer time intervals can be very challenging if one is not explicitly counting seconds.

For simplicity, we also use reward amounts (preference for high stimulus values) whenever the arguments for amounts *versus* delays are just complements, unless specifically mentioned otherwise; that is, among memories of several outcomes, the animal prefers the largest one. In other words, an outcome is preferred if it exceeds the maximum of the alternatives. Because the distribution of a maximum is just a product of individual distributions,[1] it follows from equations (1) and (2) and these general properties of probabilities that the specific probability of choosing option i from the finite set of alternatives is

$$P(S_i > \max_{j \neq i}\{S_j\}) = \int_{-\infty}^{\infty} f_i(x) \prod_{j \neq i} F_j(x) dx, \qquad (3)$$

where F_j is the cumulative distribution function of the density f_j that was defined in equation (2). This is SUT for the "general experiment" we introduced.

[1] If $F_Y(y) = \mathrm{P}(Y \leq y)$ and $F_Z(z) = \mathrm{P}(Z \leq z)$ are probability distributions of independent random variables Y and Z, then the distribution of their maximum is $\mathrm{P}(\max\{Y, Z\} \leq x) = \mathrm{P}(\{Y \leq x\} \cap \{Z \leq x\}) = \mathrm{P}(Y \leq x)\mathrm{P}(Z \leq x) = F_Y(x)F_Z(x)$. Here, "$\cap$" refers to both the conditions holding, and the same argument generalizes to an arbitrary (finite) number of variables.

Although SUT is based on the assumption of drawing a single sample from a "reference memory", Kacelnik & Brito e Abreu (1998) expressed reservations regarding this assumption, and Todd & Kacelnik (1993) modelled the joint effects of "reference" (long-term) and "working" (recent-experience) memory on choice behaviour. Here we nevertheless focus on the standard single-sample formulation of SUT, because we think it is important to properly understand one model before comparing it with other models. Furthermore, some recent psychological data (Rauhut & Lorenz, 2011; Vul & Pashler, 2008) and influential ideas about neural computation (Dayan & Abbot, 2001; Deco & Rolls, 2006; Hopfield, 1982) seem in line with the single sampling assumption, as further explained in the Discussion section.

1.1.2. Zero rewards or delays

It is not immediately obvious that this version of SUT really applies to the entire general experiment. For example, Shafir et al. (2008) considered a specific example of a choice between a certain reward ($\pi_{1,1} = 1$, $m_{1,1} = 3$) *versus* a variable one with an approximately equal mean ($\pi_{2,1} = 0.8$, $\pi_{2,2} = 0.2$, $m_{2,1} = 4$, $m_{2,2} = 0$), but previous work has not explicitly addressed what happens when an option delivers a reward of zero ($m_{i,j} = 0$) with a positive probability ($\pi_{i,j} > 0$).

Shafir and colleagues (2008) did not explicitly make reference to SUT, but noticed that if an animal draws samples from its memory and has perfectly accurate estimates for both m's and π's, then for a given draw the risky option will yield precisely the value "4" with probability 0.8 and value "0" with probability 0.2. The "safe" option, in turn, always yields precisely the value "3". In repeated testing, the "safe" option "wins" (the relation $S_{\text{safe}} > S_{\text{risky}}$ is satisfied) only every fifth trial, and thus the animal will pick the "safe" option with a probability 0.2; its behaviour is

characterized as "risk prone". In contrast, if the animal cannot discriminate at all between the values "3" and "4" but readily discriminates between nothing and something, then in half of the "risky" mental draws yielding a reward (proportion 0.8 of all draws), the draw from the "safe" option will nevertheless "win", and thus the animal now picks the safe option with probability 0.6 (i.e., probability 0.2 for no reward from "risky" option plus 0.5×0.8 from the now indiscernible rewards). Thus, the introduction of perceptual/memory error has changed the animal's behaviour in this experiment from risk prone [i.e., P(choose "safe") = 0.2 < 0.5] to risk-averse [i.e., P(choose "safe") = 0.6 > 0.5]. Shafir et al. (2008) showed that humans and honeybees indeed changed from risk-prone to risk-averse when the perceptual accuracy was experimentally reduced by using a very easy task and then a much harder task. How can SUT handle this case where zero-valued rewards are introduced?

A formal mathematical explanation is given in the Appendix A, but here it suffices to observe Figure 2A and take it for granted that the distribution $N(x; 0, \gamma 0)$ is Dirac's delta functional $\delta(x)$ that assigns a probability mass 1 to the point $\{x = 0\}$ and no mass elsewhere. Dirac's δ is not a function in the ordinary sense, but it is the outcome of completing the limiting process in Figure 2A and it can be treated as a probability measure whose cumulative distribution function gets a value zero if $x < 0$, and 1 if $x > 0$ (Klenke, 2008; Rudin, 1991). Whereas this complication of zero rewards is thus solved by some knowledge of measure-theoretic calculus and functional analysis, Figure 2A reveals a further complication.

1.1.3. Alternative fixed-outcome densities

Although we see that SUT limits the unwieldy negative draws from memories of amounts (the probability mass below 0 in Fig. 2A) by the fact that the standard deviation tends to zero in proportion to the mean, these nevertheless exist in the model. Kacelnik & Brito e Abreu (1998) argue that negative memories are "unrealistic and can be ignored" by starting the integration in equation (1) from 0 rather than minus infinity. If this were the case, then P(choose "safe") would not be predicted to equal 1 – P(choose "risky") in an experiment where an animal must choose either the "safe" or the "risky" option (i.e. P would not be a probability measure but a sub-probability, or finite measure; Klenke, 2008). A better way to do the truncation of the integral would be to replace the underlying normal distributions with truncated normal distributions (Appendix B). Numerical experiments show that the use of truncated normal distributions provides mostly the same predictions (Figure 2B) and negative values may turn out to have an interpretation (see discussion section on interpreting γ). Thus, when it is inconsequential, we use the simplest SUT model with Normal distributions and unrestricted integrals, and only discuss what the results would be for alternative distributions.

In addition to the Normal distribution, the Inverse Gaussian distribution can produce the scalar property, is consistent with certain kinds of theoretical timing networks and experimentally found properties of neurons involved in time processing, and provides a slightly better fit to behaviour in interval timing tasks than the Normal distribution (Simen, Balci, de Souza, Cohen, & Holmes, 2011; see Figure 2C and Appendix B). Furthermore, it has been suggested that related mechanisms could be involved in the processing of quantity and numerosity in addition to time (Buhusi & Meck, 2005; Matell & Meck, 2004; Merchant,

Harrington, & Meck, 2013). The distribution arises naturally in the context of drift-diffusion processes used to model the timing of behaviour in humans and animals, but it is not strictly defined at zero. Inverse Gaussian distributions can be used in the same computations as Normal and Truncated Normal distributions, however, if the value at zero rewards is taken to be a delta distribution as for the Normal distribution (see Appendix B for details). Here, all the three approaches yield the heuristic predictions of Shafir et al. (2008) qualitatively, but not quantitatively, as we now show.

Figure 2D plots the SUT-predicted probability of choosing the "safe" option in the example experiment as a function of the perceptual-memory noise γ (note that $1/\gamma$ then parameterizes discriminability). We see that at the limit of perfect discriminability ($\gamma = 0$) the animal is predicted to be risk prone and when the discriminability deteriorates (γ increases) the animal is predicted to be risk averse, but only barely; certainly not nearly as much as heuristic computations and empirical results in Shafir et al. (2008) would suggest. This is because one-dimensional SUT is not flexible enough to allow a situation where the animal would easily (i.e., nearly perfectly) discriminate no reward from reward, while at the same time, have no ability to discriminate between two distinct rewards (eliminating 'negative' memories by the use of truncated normal helps a bit here, and use of Inverse Gaussian even more; Fig. 2D). To capture such a situation, SUT should either relax its defining scalar-hypothesis or make use of stimulus variations in several dimensions [the experiment of Shafir et al. (2008) used several 'perceptual dimensions', since their easy task was presented as numbers and the difficult one as dots]. On the other hand, it is the falsifiable predictions (restricted flexibility) that make SUT an attractive model, and motivates our efforts to fully outline these predictions.

In summary, the central assumptions of SUT are: (i) the process of retrieving perceptual memories can be approximated by a process of random sampling from a probability distribution, (ii) the distribution of retrieved "memories" for a fixed outcome is subject to the scalar property, (iii) memories of variable outcomes correspond to mixtures of fixed outcomes with mixing weights equal to probabilities of the outcomes, and (iv) the animal (possibly human) is in a trained state. "Trained state" means here that there is no uncertainty regarding π in the general experimental paradigm; the animal knows the outcome probabilities, but still has memory/perceptual imprecision in the quantity of delay to or amount of reward. We adopt these previously implicitly introduced assumptions here, but feel that we have provided a more rigorous and explicit definition of SUT than has been available in the previous literature. We will next proceed to outline the qualitative predictions of SUT.

2. Results

2.1. Model predictions

2.1.1. SUT predicts both risk aversion and proneness for both amounts and delays

Although it has been implied that SUT always predicts animals to be risk averse in the face of variability in reward amount and risk prone for delays to reward (Bateson & Kacelnik, 1995b; Kacelnik & Bateson, 1996; Kacelnik & Brito e Abreu, 1998; Kacelnik & El Mouden, 2013), the formal treatment of the experiment in Shafir et al. (2008) in the "Model" section of this paper suggested that this is not true when the experiment involves a zero-valued outcome. The associated empirical experiment also found risk proneness for amounts (Shafir et al., 2008). As we now show, SUT can predict both risk proneness and risk aversion for reward amounts

depending on the values of γ and π, even when zero-valued rewards are not involved. This is of interest because there are several empirical studies where animals demonstrate risk proneness for amount (De Petrillo, Ventricelli, Ponsi, & Addessi, 2015; Haun, Nawroth, & Call, 2011; Ratikainen, Wright, & Kazem, 2010; Xu & Kralik, 2014).

We consider an experiment that defines the "safe" option as ($\pi_{1,1} = 1$, $m_{1,1} = m$) and the "risky" option as ($\pi_{2,1} = 1 - \theta$, $\pi_{2,2} = \theta$, $m_{2,1} = s$, $m_{2,2} = s + (m - s)/\theta$), where $0 < \theta < 1$ and $0 < s < m$. That is, the safe option always yields a reward m, whereas the risky option yields a smaller reward with probability $(1 - \theta)$ or a bigger reward with probability θ, the average being m, nevertheless. Notice that this is the 'paradigmatic' experiment in Figure 1, with the exception that θ can take values other than 0.5. Figure 3 shows the contour plots for the probability of choosing the "safe" option [i.e., P(choose safe); left column] and the animal's indifference regions [i.e., P(choose safe) ≈ 0.5; right column] as a function of both θ and γ. It can be seen that SUT yields all the conceivable cases: depending on the experimental parameter θ and the animal's perceptual noise γ, the animal is predicted to be risk prone, risk averse, or indifferent to risk. Thus, there is no general qualitative prediction of SUT; specific predictions need to be determined on a case by case basis (almost exactly the same results are obtained for truncated normal and inverse Gaussian distributions; not shown, but available from the authors upon request). However, in the special case where the means of the alternatives are equal and the risky option has two equiprobable outcomes (i.e., $\theta = 0.5$), SUT consistently predicts risk aversion for amount or indifference to risk, but not risk proneness (see Fig. 3). We now consider this special case.

When a risky option has two equiprobable outcomes, SUT predicts indifference between the risky option and the safe option precisely when the magnitude of the safe option equals the

geometric mean of the two equiprobable outcomes in the risky option (Bateson & Kacelnik, 1995b). Because SUT assumes that choice is not based on any explicit concept of risk and that the animal only tries to maximize its immediate reward, one could think that this geometric mean represents the "point of subjective equality" for the animal, wherein the value of the safe option "looks" or "is remembered" equal to the average value of the risky option, thereby resulting in the animal's indifference. This provides a link between SUT and empirical literature on the bisection of an interval, where animals and humans display the point of subjective equality at the geometric mean of the interval's end points (Allan & Gibbon, 1991; Church & Deluty, 1977; Jordan & Brannon, 2006; Pearson, Roitman, Brannon, Platt, & Raghavachari, 2010; Platt & Davis, 1983; Stubbs, 1976). In this context, the geometric mean has also been referred to as the "certainty equivalent" of the variable-option value (Kacelnik & Brito e Abreu, 1998). Because the geometric mean is less than or equal to the arithmetic mean, certainty equivalence implies risk aversion or indifference for amounts, but not risk proneness. In Appendix C, we show that the geometric-mean property (Bateson & Kacelnik, 1995b) does not hold for more than two equiprobable outcomes or for the alternative memory distributions, however. For all practical purposes, indifference would also result when accuracy deteriorates enough, that is, when γ tends to infinity; in that case, any value is certainty equivalent.

Without specifying further context, risk sensitivity of the above kind cannot be understood solely from rationality arguments based on average gains, because the variable (risky) choice option has the same average gain as the fixed (safe) option. Other definitions of rational choice exist, however. There has been a great deal of theoretical and empirical interest in the extent to which animal decision making can be seen as rational (Fawcett et al., 2014; Houston, 2012; Kacelnik, 2006; McNamara, Trimmer, & Houston, 2014; Monteiro,

Vasconcelos, & Kacelnik, 2013; Rieskamp, Busemeyer, & Mellers, 2006; Schuck-Paim & Kacelnik, 2007; Shafir, Waite, & Smith, 2002; Trimmer, 2013). A theoretical investigation of SUT is still lacking in these regards. We undertake that next.

2.1.2. *SUT and rationality: Transitivity and its equivalents*

Tversky and Russo (1969) showed that several rationality concepts areequivalent to the condition known as the *Strong Stochastic Transitivity* (SST). Some studies on starlings have indicated that the birds uphold such principles of rationality (Schuck-Paim & Kacelnik, 2007; Monteiro et al., 2013), but an overwhelming number of studies suggest that humans do not (Rieskamp et al., 2006). But what SUT predicts regarding rationality remains unclear. We treat here only the case of SST, and list the equivalent conditions in the Appendix D.

Simply put, transitivity of choice preferences means that if an animal prefers an option *b* over an option *a* and also prefers an option *c* over the option *b*, then it should prefer option *c* over option *a* (a strict definition is given below). Now, there are many conceivable cases where such a condition might not hold despite optimal behaviour on behalf of the animal; for example, when the options involve differences in several qualitative dimensions, when the animal's internal or environmental state differs across the pairwise comparisons, or the options have differential availability in the future or their value depends on the overall context of the choice process (Houston, 2012; Houston, McNamara, & Steer, 2007; McNamara et al., 2014; Schuck-Paim, Pompilio, & Kacelnik, 2004). Typically, violations of transitivity have been seen as problematic for decision-making theories, because they suggest that there is no single scale that can be used to assign value (or 'utility') to choices (Houston, 2012; Rieskamp et al., 2006), but SUT is a

theory in which we have a fixed dimension of value and aim to explain effects of variance on the behaviour that arises from it. Thus, we ask a novel question: is it possible to induce violations of strong stochastic transitivity in SUT predictions just by manipulating 'experimental' variance?

The definition of SST translates to the context of SUT as follows:

$$\text{if } P(S_b > S_a) \geq ½ \text{ and } P(S_c > S_b) \geq ½, \text{ then}$$

$$P(S_c > S_a) \geq \max\{P(S_b > S_a), P(S_c > S_b)\}.$$

We prove this general relationship does not hold for SUT by deriving a contradiction; that is, by giving a specific example where it is violated. Although an infinite number of such examples exists, a single illustrative one suffices here, and in it we fix $\gamma = 0.25$. Then, let option *a* be ($m_{a,1} = 3$, $m_{a,2} = 10$, $m_{a,3} = 15$, $\pi_{a,1} = 0.6$, $\pi_{a,2} = 0.3$, $\pi_{a,3} = 0.1$), option *b* be ($m_{b,1} = 5$, $\pi_{b,1} = 1$) and option *c* be ($m_{c,1} = 6$, $\pi_{c,1} = 1$). Now SUT predicts that $P(S_b > S_a) \approx 0.56$, $P(S_c > S_b) \approx 0.70$, but $P(S_c > S_a) \approx 0.60$, which is *less* than 0.70. This contradicts the SST assumption, and thus disproves it for SUT. Since the proof was not constructive, we would like to better understand why this happens.

Figure 4A shows the cumulative distribution functions of each option's SUT representation. Because option *a* is highly skewed, option *c* compared to option *b* takes a much larger share of total probability mass over the decision boundary (the diagonal line) than when option *c* is compared to option *a* despite *b* already being preferred to *a*; this breaks SST [see Figure panels 4B–D; cf. Eq. (1) and Fig. 1]. In effect, this is because of the skew, which causes the heavy tail of the distribution of *a* to only slowly "drag" over the diagonal line as a function of translations in the more concentrated (fixed or safe) member of the joint bivariate distribution. Thus, we conclude this section by observing that strong stochastic transitivity can be broken for

SUT by relatively simple manipulations of skew. Similar argument could be made for both truncated normal and inverse Gaussian base distributions.

As our example shows, it is not difficult to construct experimental settings for which SUT predicts a violation of SST. Just one such example suffices for a mathematical proof by contradiction; however, it is equally easy to construct special cases where SST is predicted to hold. For example, taking a less skewed option *a* such that ($m_{a,1} = 3$, $m_{a,2} = 6$, $\pi_{a,1} = 0.5$, $\pi_{a,2} = 0.5$) and keeping *b* and *c* as they were above, one finds that *b* is preferred to *a* at $P(S_b > S_a) \approx 0.61$, *c* to *b* at $P(S_c > S_b) \approx 0.70$, and as expected in SST, $P(S_c > S_a) \approx 0.73 \geq \max\{0.61, 0.70\}$. Thus, the take-home message here is not just that SST does not generally hold for SUT, but also that experimentalists interested in both rationality and SUT should establish what SUT predicts in their particular experiment.

2.1.3. SUT and rationality: Probability Matching

A much discussed finding in decision-making research is that even though an animal has strong evidence that an option yields a more frequent pay off than its alternative, it does not consistently choose the better option but only in proportion to the reward frequencies of the two options (Erev & Barron, 2005; Houston, Kacelnik, & McNamara, 1982; Koehler & James, 2014; Vulkan, 2000). This finding, known as probability matching, can be seen as a violation of rationality in the sense that the behaviour does not maximize expected rewards. Our general experiment can be turned into a typical probability-matching study by considering two uncertain rewards of equal size but unequal frequency; for example, ($m_{1,1} = 3$, $m_{1,2} = 0$, $\pi_{1,1} = 2/3$, $\pi_{1,2} = 1/3$) and ($m_{2,1} = 3$, $m_{2,2} = 0$, $\pi_{2,1} = 1/3$, $\pi_{2,2} = 2/3$). Given that the expected reward of the first option is 2 and that of

the second option only 1, the animal would maximize its gains by always choosing the first option. The probability-matching behaviour, in contrast, would yield a choice frequency of 2/3. What does SUT predict?

The prediction of SUT in the specific probability-matching experiment is given by integral of $f_1 \times F_2$ over the real line, where the functions $f_1(x) = 1/3\times\delta(x) + 2/3\times N(x; 3, \gamma 3)$ and $F_2(x) = 2/3\times H(x) + 1/3\times\Phi(x; 3, \gamma 3)$. Here, δ is the Dirac's Delta function, H is its cumulative distribution function (i.e., Heaviside step function), N is the SUT's normal density and Φ its cumulative distribution function. According to the derivation of Appendix E, we get a prediction $P(S_1 > S_2) = 1/3\times2/3\times1/2 + 1/3\times1/3\times\Phi(0) + 2/3\times2/3\times(1 - \Phi(0)) + 1/3\times2/3\times\int N(x)\Phi(x)dx$. We suppressed m and γ in the above equation, but by setting $m = 3$ and $\gamma = 0.5$, it follows that $P(S_1 > S_2) \approx 0.659 \approx 2/3$. Thus, we recovered the probability-matching prediction almost exactly using SUT.

2.1.4. SUT and Coefficient-of-Variation models

It has been noticed that risk sensitivity is well predicted by the coefficient of variation (CV) of the "risky" option, but not by variance or standard deviation (Shafir, 2000; Weber, 2010; Weber, Shafir, & Blais, 2004). The CV is defined as the standard deviation divided by the mean, similarly to γ in SUT. A crucial difference, however, is that the CV refers to the properties of the experimental condition in question (fixed over individuals and species, varies across experiments), whereas γ refers to the subject's internal processes (varies over individuals and species, fixed across experiments). Specifically, the meta-analytic findings suggest that risk sensitivity [i.e. |P(choose risky) – 0.5|] is linearly predicted by the CV of the risky option in

experiments where subjects (animals or humans) need to learn the reward values of the choice options by experience instead of being simply told them (as in some experiments on humans) (Shafir, 2000; Weber et al., 2004). Given that both the models, CV-based and SUT, rely on the 'scalar property' and are applied to same data, their inter-relationship should be made explicit. This follows next.

We consider again the "paradigmatic" experiment where subjects choose between fixed/safe and variable/risky option, both having equal means: the "safe" option is ($\pi_{1,1} = 1$, $m_{1,1} = m$) and the "risky" option is ($\pi_{2,1} = 1 - \theta$, $\pi_{2,2} = \theta$, $m_{2,1} = s$, $m_{2,2} = s + (m - s)/\theta$), where $0 < \theta < 1$ and $0 < s < m$. For ease of reference, we call s the small reward, m the mean reward and $b := s + (m - s)/\theta$ the big reward. We now ask: to what extent does the CV of the risky option linearly predict the risk sensitivity implied by SUT? That is, are predictions of SUT and CV models distinguishable? To get a picture of this, we arbitrarily fix $s = 3$ and start varying the difference between big and small rewards of the variable option, $b - s$, and the probability of the big reward, θ, while holding the mean reward equal across the two choice alternatives and comparing predictions of SUT to the ensuing "experimental" CVs. We observe that the standard deviation increases relative to the mean (i.e. CV increases) both when the reward values get further from each other (i.e. $b - s$ increases) and when the probability of the big reward (i.e., θ) decreases (Figure 5A). Figure 5B–D shows how variance (Var), standard deviation (s. d.) and coefficient of variation (CV) change as a function of the mean difference, $b - s$, for three different values of θ (see panel legend). Given this series of "experiments", one can then evaluate the prediction of SUT and see how it relates to the three alternative descriptive statistics of variation.

Experimental variance is a convex function of the mean difference (Fig. 5B), standard deviation is a linear function (Fig. 5C), and coefficient of variation is a concave function (Fig. 5D). As it happens, SUT predicts that risk sensitivity (i.e., |P(choose safe) – 0.5|) is a concave function of the experimental mean difference (Fig. 5E). Unsurprisingly, when a concave function [i.e. P(choose safe)] is plotted against another concave function (CV), the resulting graph looks more linear than when plotting it against convex (Var) or linear (s.d.) functions (Figure 5F–H). This does not mean that predictions of SUT would be linear functions of CV, however. Both the SUT prediction and the CV tend towards a constant function of mean difference for high values of θ (Fig. 5H), but for the low values, the deviations from a linear relationship between SUT and CV are especially prominent for the low values of γ (i.e., high perceptual accuracy; Fig. 5I). Although CV may predicts risk sensitivity well when compared to grossly different kinds of functions, such as variance and standard deviation (Shafir, 2000; Weber et al., 2004), it remains to be tested which of these two "fair" competitors, SUT or CV, better explains the available data.

2.2. A key experiment

A strong test for SUT exists and should be easy to implement experimentally, although to our knowledge, this has not been done so far. According to SUT, if an animal has a preference, represented by a proportion of choices p_s, in an experiment ($\pi_{1,1} = 1$, $m_{1,1} = m$; $\pi_{2,1} = 1$, $m_{2,1} = s$) where $s < m$ and a preference p_b in an experiment ($\pi_{1,1} = 1$, $m_{1,1} = m$; $\pi_{2,1} = 1$, $m_{2,1} = b$) where $b > m$, then SUT predicts that it has a preference $\theta p_s + (1 - \theta)p_b$ in the experiment ($\pi_{1,1} = 1$, $m_{1,1} = m$; $\pi_{2,1} = \theta$, $m_{2,1} = s$, $\pi_{2,2} = 1 - \theta$, $m_{2,2} = b$). This is a direct consequence of the way that SUT defines variable outcome representations as a linear superposition (mixture density) of fixed outcome representations [see equation (2)]. The test of SUT then consists of first conducting the two

experiments that measure the preferences (p_s and p_b) when an animal compares just two fixed rewards (using typical paradigms for the given species), and then conducting the last experiment where the variable reward is compared to the fixed reward; if SUT is a correct model, the outcome of the last experiment should be predictable from the two preceding experiments [i.e. yield the preference $\theta p_s + (1 - \theta)p_b$]. This should hold for all values of θ between 0 and 1. Otherwise, a central assumption of SUT, represented in the equation (2), will be falsified.

The strength of this test derives from the fact that it does not require an estimate for the only unknown parameter of the model, γ. The animal's preference in the third experiment should be $\theta p_s + (1 - \theta)p_b$ for all γ. Moreover, the key experiment is invariant with respect to replacing N in equation (2) with some other distribution; for example, the truncated normal or the inverse Gaussian. This test specifically assesses whether animals and humans encode variable rewards using mixture distributions of fixed rewards and then perform mental sampling from these mixture distributions. These are core assumptions of SUT that seem to warrant most study now, since Weber's law and the scalar property have been roughly confirmed in simpler settings (Akre & Johnsen, 2014, Buhusi & Meck, 2005; Gibbon, 1977; Schuck-Paim & Kacelnik, 2007; Simen et al., 2011).

3. Discussion

This paper has defined basic properties of SUT more rigorously than previous work, and highlighted the model's predictions to a fuller extent than has been available from the previous literature. Contrary to common beliefs, we have shown here that (1) SUT predicts both risk aversion and risk proneness for both reward amounts and delays to reward, (2) SUT implies

violations of different concepts of rationality (it violates strong stochastic transitivity and its equivalents, and leads to probability matching behaviour), and (3) SUT can predict, but does not always predict, a linear relationship between risk sensitivity in choices and coefficient of variation in the choice experiment. These predictions are not sensitive to replacing the model's Normal distributions with truncated normal, or with inverse Gaussians implied by drift-diffusion models of neural processing (Simen et al., 2011). Furthermore, we have provided a straightforward way to test SUT's assumption of mental sampling from mixture-distribution representations. In a broader context, our results contribute to better understanding of proportional processing, which may be a prerequisite for success in the Tinbergian aim to unify proximal and distal accounts of choice behaviour (Akre & Johnsen, 2014; Kacelnik & El Mouden, 2013; McNamara & Houston, 2006; Tinbergen, 1963).

We expect that we have not exhausted the predictions of SUT, but we nevertheless believe this theoretical review to be useful for experimentalists in guiding their hypotheses and in facilitating the testing of SUT. It should also be useful for theoreticians considering extensions of SUT, or wishing to become familiar with the model. Regarding extensions of SUT, there is an obvious need to understand risk preferences in animal and human decision making when the available options vary in several qualitatively different dimensions (Houston, 1991, 2012; Shapiro, Schuck-Paim, & Kacelnik, 2012). The most pressing extension in these regards concerns simultaneous experimental variation in both reward amounts and delays to reward (Shapiro et al., 2012). In addition to many real-world decision-making challenges involving multiple dimensions, sometimes it is also difficult, or nearly impossible, to ensure that the experimental manipulations occur only in a single dimension (cf. section 1.3). A further extension pertains to modelling of risk *versus* ambiguity, and will be discussed in the sub-section

below. Then we will discuss possible mechanisms underlying SUT and our view of immediate future prospects.

3.1. Risk and Ambiguity

An animal's choice preferences under risk can sometimes be dissociated from those under ambiguity (Santos & Rosati, 2015; Trimmer et al., 2011). Here, "risk" refers to a case where probabilistic variation in the reward outcomes is known, whereas "ambiguity" refers to problems where the probabilities are unknown. For example, a decision maker faces "risk" when deciding whether a ball drawn randomly from an urn turns out to be red or black, if it knows the number of red and black balls in the urn. It faces "ambiguity" when the number of balls in the urn is unknown, or only partially known. As we discussed above, SUT pertains only to risk, and in typical experiments where SUT is used to understand the data, the experimenter tries to carefully ascertain that the animal has learned the "relative shares of the balls". Cross-species comparisons suggest that the cognitive systems supporting risk preferences are distinct from those supporting ambiguity preferences (Santos & Rosati, 2015), but clearly these are often difficult to fully dissociate in experiments and thus it would be valuable to be able to use SUT despite variations in the degree of ambiguity.

In reality, learning relative values of m's and π's is likely to be inter-dependent, but a simplifying approximation would be to assume that they are independent. By introducing a distribution for π's, one would then extend SUT to handle ambiguity as well. Since the value of π can make a qualitative, not just quantitative, difference to choice behaviour according to SUT (e.g., θ in Figure 3), we know that a distribution on it has to make a qualitative difference too.

Extensions for modelling ambiguity will be left for future studies. Instead, we will next consider some of the possible mechanistic underpinnings for the descriptive aspects of SUT.

3.2. What does the parameter γ stand for?

Although Kacelnik and Brito e Abreu (1998) state that the coefficient of variation, γ, is not a free parameter of the model, because it reflects internal-processes noise and can be estimated empirically, they did not explicitly specify the mechanisms that it reflects. Thus, γ makes SUT a descriptive rather than mechanistic model. In an analogy, "descriptive" population models frequently introduce a parameter, K, for the carrying capacity of the environment that needs to be estimated empirically; only models where *all* parameters have an interpretation in terms of the behaviour of individuals (constituents of the population) are then called "mechanistic" (Geritz & Kisdi, 2012). By that logic, when analysing individuals instead of populations, all the parameters in a mechanistic model need to have an interpretation in terms of the constituents of an individual. Unless we specify the physiological and mental processes that suffer from the noise γ, SUT remains a descriptive model. Because mechanistic models have many benefits over descriptive ones (Geritz & Kisdi, 2012; Servedio et al., 2014), we now discuss possible ways to explicitly connect γ with internal mechanisms.

The striatal beat-frequency model is an influential mechanistic and neural model of behavioural timing, and related mechanisms could be involved in the processing of quantity and number in addition to time (Buhusi & Meck, 2005; Matell & Meck, 2004; Merchant, Harrington, & Meck, 2013). The model assumes that behavioural timing is neurally implemented as pattern detection, or pattern matching. At the start of a temporal estimation task, a 'start-gun' process

(located in ventral tegmental area and substantia nigra pars compacta) synchronizes multiple cortical oscillatory neurons with different wavelengths that project onto striatal medium spiny neurons, which continuously compare the current pattern of activity to the (learned) pattern that has been detected at the time of reward. Although they did not make a connection with the striatal beat-frequency model, Deco & Rolls (2006) modelled vibrotactile frequency discrimination as a pattern detection (attractor) decision-making neural network, explaining observations of Weber's law in that context (Deco & Rolls, 2006; Deco, Scarano, & Soto-Faraco, 2007). Their model clearly suggested that γ has an inverse relationship with the number of neurons in the attractor network; that is, accuracy increases with the number of neurons (Deco & Rolls, 2006). Similarly, one expects more distinct patterns and more accuracy for superpositions with a greater number of cortical oscillator neurons than with a lesser number, because they allow a better approximation to a full Fourier transform of an impulse function for time of reward. This suggests an inverse relationship between γ and the number of neurons for the striatal beat-frequency model too. Hence, it seems possible that the number of coding neurons is related to accuracy, $1/\gamma$.

Neural pattern detection and attractor models suggest that the 'negative' memories implied by the left tail of SUT's 'mental' normal variates just represent neural convergence errors for which the tail provides a good description. The same process might also explain why the reward-frequency distribution (the mixing distribution π of SUT) is not confounded with the reward-value distributions (the normal distributions of SUT) in the animal's head. This would be because the animal does not 'remember', or even represent to begin with, any normally distributed range of reward values; just the one and correct value (i.e., the pattern), which it only fails to match with a running pattern detection process to a varying degree.

Furthermore, the idea of associative memory (attractor) networks resonates with the single-sampling assumption in SUT, because such a network can simultaneously contain representations of multiple patterns and be able to activate only one at the time, the choice of which is subject to errors, or noise (Hopfield, 1982; Dayan & Abbott, 2001).

In addition to a pattern-detection approach, a drift-diffusion process where the internal clock is based on 'counting' neural spikes is also biologically plausible and in line with experimental findings (Simen et al., 2011). This implies that γ is related to a delicate balance between excitatory and inhibitory networks that achieve the spike counting and to inherent noise in Poisson-spike generating neural units. In another kind of model, γ arises as a by-product of competing neural firing rates and a learning process (Pearson et al., 2010). Pearson et al. (2010) suggested that the decrease in accuracy (here, increase in γ) exists because it is beneficial for reinforcement learning; that is, the associated increase in behavioural sampling ensures better learning outcomes. This offers a theoretical viewpoint in which an apparent decrease in accuracy can have 'benefits'.

Understanding the processes behind perceptual accuracy might ultimately upgrade SUT's status from "descriptive" to "mechanistic". In turn, more rigorous testing of SUT and outlining of its empirically supported predictions could help to distinguish between the mechanistic interpretations and to organise empirical findings. It is also likely to illuminate certain empirical phenomena, such as the "wisdom of crowds within an individual", where researchers have pointed out that an average of individual's guesses for a quantity tends to be more accurate than the individual guesses, despite he or she gaining no further information between the successive guesses (Rauhut & Lorenz, 2011; Vul & Pashler, 2008). Such a phenomenon is expected if

individuals' mental representations are partly equivalent to independent random samples, as assumed by SUT.

3.3. Future prospects

We have outlined a "key experiment" described in the section 2.2 of this paper, and consider it as an important future test of SUT. The key experiment tests the general structure of SUT, irrespective of the estimable parameter γ. Another possible goal for future work is to extend the existing meta-analyses to a new one that explicitly compares the relative merits of SUT and CV models (Kacelnik & Bateson, 1996; Shafir, 2000). We also stressed the importance of finding explicit mechanisms for the mental coefficient of variation, γ, whose inverse captures the choice accuracy. In addition to these important aims, other open theoretical and empirical issues remain for SUT, such as understanding selective pressures and constraints on γ. Regarding extensions of SUT, since experiments often involve several stimulus dimensions despite all the efforts to constrain them to a single one, a multidimensional extension of SUT could be useful in understanding these data (Brunner, Gibbon, & Fairhurst, 1994; Shapiro et al., 2012). Another future extension of SUT would capture the role of ambiguity in addition to the role of risk, because animals' choice preferences under risk can sometimes be dissociated from those under ambiguity (Santos & Rosati, 2015; Trimmer et al., 2011).

Related to the role of ambiguity is the role of learning. Animals invariably need to learn the various reward probabilities before SUT can be applied, and in addition to allowing experimental control over the accuracy of the learned estimates, an explicit model of the learning process might illuminate the still open mechanistic interpretations in SUT. Sequential learning in

humans and animals is an active field of research on its own (Cappé, Garivier, Maillard, Munos, & Stoltz, 2013; Erev & Roth, 2014; Houston et al., 1982; Speekenbrink & Konstantinidis, 2014; Sutton & Barto, 1998; Yechiam, Busemeyer, Stout, & Bechara, 2005), and might be fruitfully married with the study of proportional processing (Pearson et al., 2010).

3.4. Conclusion

Much scientific debate around proportional processing has pertained to 'memory' distributions for delays to reward and for reward amounts/sizes. Here we showed that the exact memory distributions often make little difference to predictions about risk preferences, rationality and external *versus* internal statistics based on proportional processing (Weber's law), and that these predictions are much richer than has been expected. It is our hope that this theoretical investigation will stimulate fruitful experimental work and further theory in the future.

Disclosures and Acknowledgements

The authors declare no conflicting interests. This work was supported by European Research Council Advanced Grant 250209 to A.I.H. We thank Alex Kacelnik for comments on a previous version of this manuscript.

Appendix A Dirac's delta function as a natural model for zero rewards and delays in SUT

Dirac's delta functional δ is not an ordinary function, but a linear mapping from the space $D(\mathbf{R})$ of infinitely differentiable and compactly supported, real-valued 'test' functions onto the real line

R, such that for any φ in $D(\mathbf{R})$, $\delta[\varphi] = \varphi(0)$ (Rudin, 1991). A standard and convenient definition by abuse of notation sets δ to be the function that yields the integral equation

$$\delta[\varphi] = \int_{-\infty}^{\infty} \varphi(x)\delta(x)dx = \varphi(0),$$

but a function that would satisfy this does not exist in the ordinary theory of (Riemann or Lebesgue) integration. Instead, the theory of distributions can be used for a rigorous definition (Rudin, 1991). For present purposes, it suffices to acknowledge that a rigorous definition does exist, and several limiting processes of ordinary functions naturally lead to δ. Many of them are not relevant to SUT, however. Thus, we now prove that $\lim_{\mu\to 0} N(\mu, \gamma\mu) = \delta$.

Without a loss of generality, let $\mu = 1/n$ with n a positive integer and $\gamma = 1$ in the proof, and define a map δ_n from $D(\mathbf{R})$ to **R** such that for any φ

$$\delta_n[\varphi] = \int_{-\infty}^{\infty} \frac{n}{\sqrt{2\pi}} e^{-\frac{(nx-1)^2}{2}} \varphi(x)dx. \qquad \text{(A.1)}$$

The function multiplying $\varphi(x)$ within the integral in (A.1) is just $N(1/n, 1/n)$. The proof is completed by showing that the difference $|\delta_n[\varphi] - \varphi(0)|$ gets arbitrarily small as n grows. Let $\varepsilon > 0$ be an arbitrary, small number. Choose a value $\tau > 0$ such that $|\varphi(x) - \varphi(0)| < \varepsilon/3$ for all x in the interval $[0 - \tau, 0 + \tau]$, which is known to exist due to the continuity of φ. Then we have

$$\delta_n[\varphi] - \varphi(0) = \int_{-\infty}^{\infty} \frac{n}{\sqrt{2\pi}} e^{-\frac{(nx-1)^2}{2}} (\varphi(x) - \varphi(0))dx$$

$$= \int_{-\infty}^{-\tau} \frac{n}{\sqrt{2\pi}} e^{-\frac{(nx-1)^2}{2}} (\varphi(x) - \varphi(0))dx + \int_{-\tau}^{\tau} \frac{n}{\sqrt{2\pi}} e^{-\frac{(nx-1)^2}{2}} (\varphi(x) - \varphi(0))dx$$

$$+\int_{\tau}^{\infty}\frac{n}{\sqrt{2\pi}}e^{-\frac{(nx-1)^2}{2}}(\varphi(x)-\varphi(0))dx\,.$$

The middle integral we chose to be less than $\varepsilon/3$ in absolute value for all n, since $N(1/n, 1/n)$ integrates to 1. Because φ is continuous on a compact support, it is bounded and therefore there exists $M < \infty$ such that $M > |\varphi(x) - \varphi(0)|$ for any x. For the other two integrals, we may notice that $nx/\tau > 1$ when $x > \tau$, and by further assuming that $1/n < \tau$, use approximations like

$$\left|\int_{\tau}^{\infty}\frac{n}{\sqrt{2\pi}}e^{-\frac{n^2(x-1/n)^2}{2}}(\varphi(x)-\varphi(0))dx\right| \le \frac{M}{\sqrt{2\pi}}\int_{\tau}^{\infty}n\frac{nx}{\tau}e^{-\frac{n^2x^2}{2}}dx = \frac{M}{\tau\sqrt{2\pi}}e^{-\frac{n^2\tau^2}{2}}. \qquad \text{(A.2)}$$

Now, by choosing n so large that both right-hand side of (A.2) is smaller than $\varepsilon/3$ and $1/n < \tau$ we have completed the proof. Although there are other ways to derive δ than using $\lim_{\mu\to 0} N(\mu, \gamma\mu)$, this one is important here, and will be specifically needed in Appendix E.

Appendix B Alternative distributions for certain outcomes

The base density function of SUT is $N(x; \mu, \mu\gamma)$, with $\mu > 0$ and $\gamma > 0$, but this may have positive values for $x < 0$. In other words, negative reward amounts, and even negative delays to reward, are implied for animals' mental representations. If one would like to eliminate the possibility of negative values with only minimal change to the original SUT, the density $N(x; \mu, \gamma\mu)$ can be replaced with another density function

$$N_{trunc}(x;\mu,\gamma\mu) = \begin{cases}\frac{1}{Z}N(x;\mu,\gamma\mu), & x \ge 0\\ 0, & x < 0\end{cases},$$

where $Z = 1 - \Phi(-1/\gamma)$, and $\Phi(\cdot)$ denotes the standard normal cumulative distribution function $\Phi(\cdot; 0, 1)$. N_{trunc} is a truncated normal distribution density with a lower truncation at 0. The mean of this distribution is not μ, however, but $\mu + N(-1/\gamma; \mu, \gamma\mu)/Z$. The standard deviation also differs from $\gamma\mu$. When μ is not close to zero, this distribution is close to a normal distribution, but the difference between the distributions can take any value for small enough μ, which can cause different results in the very low reward amount/delay regime. For example, $N_{\text{trunc}}(0; \mu, \gamma\mu) - N(0; \mu, \gamma\mu) = (1/Z - 1)\exp(-1/2\gamma^2)/ (\gamma\mu(2\pi)^{1/2})$, which can be made arbitrarily large by decreasing μ.

Another distribution that strictly fulfils the scalar property is the Inverse Gaussian distribution

$$IG(x;\mu,\lambda) = \left(\frac{\lambda}{2\pi x^3}\right)^{1/2} \exp\frac{-\lambda(x-\mu)^2}{2\mu^2 x},$$

with the parameter $\lambda = \mu/\gamma^2$. With this parametrization, the mean is μ, the standard deviation is $\mu\gamma$, and skewness is 3γ. Since normal distribution has skewness 0, an Inverse Gaussian differs clearly from a normal distribution for all μ, but it satisfies the scalar property (standard deviation is proportional to mean) and can be a better description of observed response times than the normal distribution (Simen et al., 2011). It is a (difficult) empirical question which one of the three distributions is the best model, but the normal distribution is certainly easiest to analyse theoretically, and in many cases, the predicted behaviour is qualitatively the same for all distributions.

Appendix C Geometric mean is not a certainty-equivalent value in general

Bateson & Kacelnik (1995b) prove that, under the SUT model, a certain reward is equally preferred to an option yielding two equiprobable outcomes when its magnitude equals the geometric mean of magnitudes of the equiprobable outcomes. Their proof uses symmetry of the normal distribution and does not therefore generalize to skewed truncated normal and inverse Gaussian distributions. We now prove by contradiction that it does not hold for more than two outcomes either.

Assume that an animal equally prefers an option that yields m_1, m_2 or m_3, all with a probability 1/3, to an option that yields $m = (m_1 m_2 m_3)^{1/3}$ with certainty (i.e. m is the point of subjective equality between the variable-outcome option and the fixed-outcome option). Now, set $m_1 = m_2$. It follows that the experiment is the same as a two-outcome experiment yielding m_1 with probability 2/3 and m_3 with probability 1/3, and thereby $m = (m_1)^{2/3}(m_3)^{1/3}$ should be a point of subjective equality in such a two non-equiprobable outcome experiment. It is not, however. Take $m_1 = 1$, $m_3 = 5$ and $\gamma = 0.5$, and SUT will predict that P(choose constant) = 0.46 < 0.50. This completes the proof. As a side note, the deviations from the geometric mean are typically not large.

Appendix D Equivalents of Strong Stochastic Transitivity (SST)

Tversky & Russo (1969) showed that SST is equivalent to "simple scalability", "substitutability" and (weak) "independence from irrelevant alternatives". The fact that SUT breaks SST thus means that it breaks all these other requirements of rational choice, and by implication, also

Luce's stronger form of "independence from irrelevant alternatives" (Rieskamp et al., 2006; Luce, 1959). The fact that there are also (at least) two other, non-stochastic, definitions with the same name, "independence of irrelevant alternatives", has been a source of confusion in the literature (Paramesh, 1973; Trimmer, 2013); but these are less relevant for the inherently stochastic SUT. Below, we list the definitions of Tversky & Russo (1969) in the present notation, but do not repeat their proof of equivalence.

Let *a*, *b*, *c* and *d* be choice alternatives in a finite set of alternatives Υ, and let e.g. S_a denote the random variable for a draw from animal's memory representation for *a*. Then SUT expresses the Tversky & Russo (1969) conditions as follows:

Simple scalability: There are real valued functions *G* and *u* such that for all *a* and *b* in Υ

$$P(S_b > S_a) = G(u(b), u(a)),$$

where *G* is strictly increasing in its first argument and strictly decreasing in the second.

Substitutability: For all *a*, *b* and *c* in Υ,

$P(S_b > S_c) \geq P(S_a > S_c)$ if and only if $P(S_b > S_a) \geq ½$.

Independence: For any *a*, *b*, *c* and *d* in Υ,

$P(S_b > S_c) \geq P(S_a > S_c)$ if and only if $P(S_b > S_d) \geq P(S_a > S_d)$.

Whereas SUT violates all the above rationality conditions, it does not violate the condition of "regularity" as can be seen immediately from the argument of Rieskamp et al. (2006) page 644.

Appendix E Details on the use of SUT in probability-matching experiments

This appendix provides details on computations involved when using SUT to predict outcomes of probability-matching experiments. We use the general forms here: $f_1(x) = \pi_{1,1} \times N(x;\, m_{1,1},\, \gamma m_{1,1}) + \pi_{1,2} \times \delta(x)$ and $F_2(x) = \pi_{2,1} \times \Phi(x;\, m_{2,1},\, \gamma m_{2,1}) + \pi_{2,2} \times H(x)$. The aim is to compute $\mathrm{P}(S_1 > S_2) = \int f_1(x) F_2(x) dx$. The Heaviside step function, H, is defined by $H(x) = 1$ when $x > 0$ and $H(x) = 0$, when $x < 0$. The full definition of Dirac's delta functional, δ, requires quite a lot more work (Rudin, 1991), but it is characterized by the properties

$$\begin{cases} \delta(0) = \infty \\ \int_{-\infty}^{\infty} \delta(x)dx = 1 \\ \int_{-\infty}^{s} \delta(x)dx = H(s) \\ \int_{-\infty}^{\infty} \delta(x)\varphi(x)dx = \varphi(0) \end{cases},$$

where φ is a compactly supported continuous function on the real line. The properties of H and δ imply that

$$P(S_1 > S_2) = \pi_{1,1}\pi_{2,1} \int_{-\infty}^{\infty} N(x; m_{1,1}, \gamma m_{1,1}) \Phi(x; m_{2,1}, \gamma m_{2,1}) dx + \pi_{1,1}\pi_{2,2}(1 - \Phi(0; m_{1,1}, \gamma m_{1,1}))$$

$$+ \pi_{1,2}\pi_{2,1}\Phi(0; m_{2,1}, \gamma m_{2,1}) + \pi_{1,2}\pi_{2,2} \int_{-\infty}^{\infty} \delta(x) H(x) dx .$$

Everything in the expression is straightforward to compute, except the last integral over δ times H. In general, it is not uniquely defined, since $H(0)$ is not. But in the case of SUT, the most natural definition is that $\int \delta(x)H(x)dx = ½$, as we now show.

In the Appendix A, we showed that the most natural definition of δ for SUT is the limit of sequence (δ_n) as n tends to infinity, where δ_n was defined in equation (A.1). If we define H_n as the cumulative distribution function of δ_n, then for all $n \geq 1$,

$$\int_{-\infty}^{\infty} \delta_n(x) H_n(x) dx = \frac{1}{2}. \tag{E.1}$$

This is because, by the equation (1), the integral in (E.1) is just $P(X_1 > X_2)$ for two independent and identically distributed (i.i.d.) normal variables X_1 and X_2. By definition of probability, we have $P(X_1 > X_2) + P(X_2 > X_1) = 1$ and the i.i.d. assumption implies that $P(X_1 > X_2) = P(X_2 > X_1)$. Thus, $P(X_1 > X_2) = ½$. Note that this only applies to continuous distributions, as then the set $\{X_1 = X_2\}$ has a measure zero; that is, $P(X_1 = X_2) = 0$ (Klenke, 2008). Now since it holds for SUT that $\int \delta(x)H(x)dx = \lim_{n\to\infty} \int \delta_n(x)H_n(x)$, the integral $\int \delta(x)H(x)dx$ must be equal to ½. Collecting all the arguments thus far, we have fully defined the classical SUT for the general experiment.

References

Akre, K. L., & Johnsen, S. (2014). Psychophysics and the evolution of behavior. *Trends in Ecology & Evolution*, *29*(5), 291–300. http://doi.org/10.1016/j.tree.2014.03.007

Allan, L. G., & Gibbon, J. (1991). Human bisection at the geometric mean. *Learning and Motivation*, *22*(1–2), 39–58. http://doi.org/10.1016/0023-9690(91)90016-2

Barnard, C. J., Brown, C. A. J., Houston, A., I., & McNamara, J. M. (1985). Risk-sensitive foraging in common shrews: an interruption model and the effects of mean and variance in reward rate. *Behavioral Ecology and Sociobiology*, *18*(2), 139–146.

Bateson, M., & Kacelnik, A. (1995a). Accuracy of memory for amount in the foraging starling,

Sturnus vulgaris. *Animal Behaviour*, *50*(2), 431–443. http://doi.org/10.1006/anbe.1995.0257

Bateson, M., & Kacelnik, A. (1995b). Preferences for fixed and variable food sources: variability in amount and delay. *Journal of the Experimental Analysis of Behavior*, *63*(3), 313–329. http://doi.org/10.1901/jeab.1995.63-313

Brito e Abreu, F., & Kacelnik, A. (1999). Energy budgets and risk-sensitive foraging in starlings. *Behavioral Ecology*, *10*(3), 338–345. http://doi.org/10.1093/beheco/10.3.338

Brunner, D., Gibbon, J., & Fairhurst, S. (1994). Choice between fixed and variable delays with different reward amounts. *Journal of Experimental Psychology. Animal Behavior Processes*, *20*(4), 331–346.

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience*, *6*(10), 755–765. http://doi.org/10.1038/nrn1764

Cappé, O., Garivier, A., Maillard, O., Munos, R., & Stoltz, G. (2013). Kullback–Leibler upper confidence bounds for optimal sequential allocation. *The Annals of Statistics*, *41*(3), 1516–1541. http://doi.org/10.1214/13-AOS1119

Caraco, T., Martindale, S., & Whittam, T.S. (1980). An empirical demonstration of risk-sensitive foraging preferences. *Animal Behaviour*, 28, 820–830.

Church, R. M., & Deluty, M. Z. (1977). Bisection of temporal intervals. *Journal of Experimental Psychology: Animal Behavior Processes*, *3*(3), 216–228. http://doi.org/10.1037/0097-7403.3.3.216

Dayan, P., & Abbott, L. F. (2001). *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems.* Cambridge, MA: The MIT Press.

Deco, G., & Rolls, E. T. (2006). Decision-making and Weber's law: a neurophysiological model.

The European Journal of Neuroscience, *24*(3), 901–916. http://doi.org/10.1111/j.1460-9568.2006.04940.x

Deco, G., Scarano, L., & Soto-Faraco, S. (2007). Weber's law in decision making: integrating behavioral data in humans with a neurophysiological model. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(42), 11192–11200. http://doi.org/10.1523/JNEUROSCI.1072-07.2007

De Petrillo, F., Ventricelli, M., Ponsi, G., & Addessi, E. (2015). Do tufted capuchin monkeys play the odds? Flexible risk preferences in Sapajus spp. *Animal Cognition*, *18*(1), 119–130. http://doi.org/10.1007/s10071-014-0783-7

Erev, I., & Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Review*, *112*(4), 912–931. http://doi.org/10.1037/0033-295X.112.4.912

Erev, I., & Roth, A. E. (2014). Maximization, learning, and economic behavior. *Proceedings of the National Academy of Sciences*, *111*(Supplement 3), 10818–10825. http://doi.org/10.1073/pnas.1402846111

Fawcett, T. W., Fallenstein, B., Higginson, A. D., Houston, A. I., Mallpress, D. E. W., Trimmer, P. C., & McNamara, J. M. (2014). The evolution of decision rules in complex environments. *Trends in Cognitive Sciences*, *18*(3), 153–161. http://doi.org/10.1016/j.tics.2013.12.012

Geritz, S. A. H., & Kisdi, E. (2012). Mathematical ecology: why mechanistic models? *Journal of Mathematical Biology*, *65*(6-7), 1411–1415. http://doi.org/10.1007/s00285-011-0496-3

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, *84*(3), 279–325. http://doi.org/10.1037/0033-295X.84.3.279

Gibbon, J., Church, R. M., Fairhurst, S., & Kacelnik, A. (1988). Scalar expectancy theory and

choice between delayed rewards. *Psychological Review*, *95*(1), 102–114. http://doi.org/10.1037/0033-295X.95.1.102

Haun, D. B. M., Nawroth, C., & Call, J. (2011). Great apes' risk-taking strategies in a decision making task. *PLoS One*, *6*(12), e28801. http://doi.org/10.1371/journal.pone.0028801

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc Nat Acad Sci USA, 79*(8), 2554–2558.

Houston, A. I. (1991). Violations of stochastic transitivity on concurrent chains: implications for theories of choice. *Journal of the Experimental Analysis of Behavior*, *55*(3), 323–335. http://doi.org/10.1901/jeab.1991.55-323

Houston, A. I. (2012). Natural selection and rational decisions. In S. Okasha & K. Binmore (Eds.), *Evolution and Rationality: Decision, Co-operation and Strategic Behaviour* (pp. 50–66). Cambridge, UK: Cambridge University Press.

Houston, A. I., Kacelnik, A., & McNamara, J. (1982). Some learning rules for acquiring information. *Functional Ontogeny, 1*, 140–191.

Houston, A. I., & McNamara, J. M. (1999). Models of Adaptive Behaviour: An Approach Based on State. Cambridge, UK: Cambridge University Press.

Houston, A. I., McNamara, J. M., & Steer, M. D. (2007). Violations of transitivity under fitness maximization. *Biology Letters*, *3*(4), 365–367. http://doi.org/10.1098/rsbl.2007.0111

Jordan, K. E., & Brannon, E. M. (2006). Weber's Law influences numerical representations in rhesus macaques (*Macaca mulatta*). *Animal Cognition*, *9*(3), 159–172. http://doi.org/10.1007/s10071-006-0017-8

Kacelnik, A. (2006). Meanings of rationality. In S. Hurley & M. Nudds (Eds.), *Rational Animals?* Oxford University Press: Oxford.

Kacelnik, A., & Bateson, M. (1996). Risky theories—the effects of variance on foraging decisions. *American Zoologist*, *36*(4), 402–434. http://doi.org/10.1093/icb/36.4.402

Kacelnik, A., & Bateson, M. (1997). Risk-sensitivity: crossroads for theories of decision-making. *Trends in Cognitive Sciences*, *1*(8), 304–309. http://doi.org/10.1016/S1364-6613(97)01093-0

Kacelnik, A., & Brito e Abreu, F. (1998). Risky choice and Weber's Law. *Journal of Theoretical Biology*, *194*(2), 289–298. http://doi.org/10.1006/jtbi.1998.0763

Kacelnik, A., & El Mouden, C. (2013). Triumphs and trials of the risk paradigm. *Animal Behaviour*, *86*(6), 1117–1129. http://doi.org/10.1016/j.anbehav.2013.09.034

Klenke, A. (2008). *Probability Theory: A Comprehensive Course*. London, England: Springer-Verlag.

Koehler, D. J., & James, G. (2014). Probability matching, fast and slow. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation* (pp. 103–130). Waltham, MA: Academic Press.

Luce, R. D. (1959). *Individual Choice Behavior: A Theoretical Analysis*. New York, NY: Wiley.

Marsh, B., & Kacelnik, A. (2002). Framing effects and risky decisions in starlings. *Proceedings of the National Academy of Sciences*, *99*(5), 3352–3355. http://doi.org/10.1073/pnas.042491999

Matell, M. S., & Meck, W. H. (2004). Cortico-striatal circuits and interval timing: coincidence detection of oscillatory processes. *Cognitive Brain Research*, *21*(2), 139–170. http://doi.org/10.1016/j.cogbrainres.2004.06.012

McNamara, J. M., & Houston, A. I. (1987). Partial preferences and foraging. *Animal Behaviour*, *35*(4), 1084–1099.

McNamara, J. M., & Houston, A. I. (1992). Risk-sensitive foraging: A review of theory. *Bulletin of Mathematical Biology*, *54*(2–3), 355–378.

McNamara, J. M., & Houston, A. I. (2009). Integrating function and mechanism. *Trends in*

Ecology & Evolution, 24(12), 670–675.

McNamara, J. M., Trimmer, P. C., & Houston, A. I. (2014). Natural selection can favour "irrational" behaviour. *Biology Letters*, *10*(1), 20130935. http://doi.org/10.1098/rsbl.2013.0935

Merchant, H., Harrington, D. L., & Meck, W. H. (2013). Neural basis of the perception and estimation of time. *Annual Review of Neuroscience*, *36*(1), 313–336. http://doi.org/10.1146/annurev-neuro-062012-170349

Monteiro, T., Vasconcelos, M., & Kacelnik, A. (2013). Starlings uphold principles of economic rationality for delay and probability of reward. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1756), 20122386. http://doi.org/10.1098/rspb.2012.2386

Paramesh, R. (1973). Independence of irrelevant alternatives. *Econometrica, 41*(5), 987–991.

Pearson, J., Roitman, J. D., Brannon, E. M., Platt, M. L., & Raghavachari, S. (2010). A physiologically-inspired model of numerical classification based on graded stimulus coding. *Frontiers in Behavioral Neuroscience*, *4*, 1. http://doi.org/10.3389/neuro.08.001.2010

Platt, J. R., & Davis, E. R. (1983). Bisection of temporal intervals by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*(2), 160–170. http://doi.org/10.1037/0097-7403.9.2.160

Ratikainen, I. I., Wright, J., & Kazem, A. J. N. (2010). Social class influences degree of variance sensitivity in wild Siberian jays. *Behavioral Ecology*, *21*(5), 1067–1072. http://doi.org/10.1093/beheco/arq106

Rauhut, H., & Lorenz, J. (2011). The wisdom of crowds in one mind: How individuals can simulate the knowledge of diverse societies to reach better decisions. *Journal of Mathematical Psychology*, *55*(2), 191–197. http://doi.org/10.1016/j.jmp.2010.10.002

Reboreda, J. C., & Kacelnik, A. (1991). Risk sensitivity in starlings: variability in food amount

and food delay. *Behavioral Ecology*, *2*(4), 301–308. http://doi.org/10.1093/beheco/2.4.301

Rieskamp, J., Busemeyer, J. R., & Mellers, B. A. (2006). Extending the bounds of rationality: evidence and theories of preferential choice. *Journal of Economic Literature*, *44*(3), 631–661.

Rudin, W. (1991). *Functional Analysis* (2nd ed.). Singapore: McGraw-Hill, Inc.

Santos, L. R., & Rosati, A. G. (2015). The evolutionary roots of human decision making. *Annual Review of Psychology*, *66*(1), 321–347. http://doi.org/10.1146/annurev-psych-010814-015310

Schuck-Paim, C., & Kacelnik, A. (2007). Choice processes in multialternative decision making. *Behavioral Ecology*, *18*(3), 541–550. http://doi.org/10.1093/beheco/arm005

Schuck-Paim, C., Pompilio, L., & Kacelnik, A. (2004). State-dependent decisions cause apparent violations of rationality in animal choice. *PLoS Biology*, 2, 2305–2315. http://doi.org/10.1371/journal.pbio.0020402

Servedio, M. R., Brandvain, Y., Dhole, S., Fitzpatrick, C. L., Goldberg, E. E., Stern, C. A., … Yeh, D. J. (2014). Not just a theory—the utility of mathematical models in evolutionary biology. *PLoS Biology*, *12*(12), e1002017. http://doi.org/10.1371/journal.pbio.1002017

Shafir, S. (2000). Risk-sensitive foraging: the effect of relative variability. *Oikos*, *88*(3), 663–669. http://doi.org/10.1034/j.1600-0706.2000.880323.x

Shafir, S., Reich, T., Tsur, E., Erev, I., & Lotem, A. (2008). Perceptual accuracy and conflicting effects of certainty on risk-taking behaviour. *Nature*, *453*(7197), 917–920. http://doi.org/10.1038/nature06841

Shafir, S., Waite, T. A., & Smith, B. H. (2002). Context-dependent violations of rational choice in honeybees (*Apis mellifera*) and gray jays (*Perisoreus canadensis*). *Behavioral Ecology and Sociobiology*, *51*(2), 180–187. http://doi.org/10.1007/s00265-001-0420-8

Shapiro, M. S., Schuck-Paim, C., & Kacelnik, A. (2012). Risk sensitivity for amounts of and

delay to rewards: adaptation for uncertainty or by-product of reward rate maximising? *Behavioural Processes*, *89*(2), 104–114. http://doi.org/10.1016/j.beproc.2011.08.016

Shapiro, M. S., Siller, S., & Kacelnik, A. (2008). Simultaneous and sequential choice as a function of reward delay and magnitude: Normative, descriptive and process-based models tested in the European Starling (*Sturnus vulgaris*). *Journal of Experimental Psychology – Animal Behavior Processes*, *34*(1), 75–93. http://doi.org/10.1037/0097-7403.34.1.75

Simen, P., Balci, F., de Souza, L., Cohen, J. D., & Holmes, P. (2011). A model of interval timing by neural integration. *The Journal of Neuroscience*, *31*(25), 9238–9253. http://doi.org/10.1523/JNEUROSCI.3121-10.2011

Speekenbrink, M., & Konstantinidis, E. (2014). Uncertainty and exploration in a restless bandit task. In P. Bello, M. Gaurini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 1491 – 1496). Austin, TX: Cognitive Science Society.

Stubbs, D. A. (1976). Scaling of stimulus duration by pigeons. *Journal of the Experimental Analysis of Behavior*, *26*(1), 15–25.

Sutton, R. S., & Barto, A. G. (1998). *Introduction to Reinforcement Learning* (1st ed.). Cambridge, MA, USA: MIT Press.

Todd, I. A., & Kacelnik, A. (1993). Psychological Mechanisms and the Marginal Value Theorem: dynamics of scalar memory for travel time. *Animal Behaviour*, *46*, 765–775. http://doi.org/10.1006/anbe.1993.1254

Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie*, 20:410–433.

Trimmer, P. C. (2013). Optimal behaviour can violate the principle of regularity. *Proceedings of*

the Royal Society of London B: Biological Sciences, *280*(1763), 20130858. http://doi.org/10.1098/rspb.2013.0858

Trimmer, P. C., Houston, A. I., Marshall, J. A. R., Mendl, M. T., Paul, E. S., & McNamara, J. M. (2011). Decision-making under uncertainty: biases and Bayesians. *Animal Cognition*, *14*(4), 465–476. http://doi.org/10.1007/s10071-011-0387-4

Tversky, A., & Russo, J. E. (1969). Substitutability and similarity in binary choices. *Journal of Mathematical Psychology*, *6*(1), 1–12. http://doi.org/10.1016/0022-2496(69)90027-3

Vul, E., & Pashler, H. (2008). Measuring the crowd within: probabilistic representations within individuals. *Psychological Science*, *19*(7), 645–647. http://doi.org/10.1111/j.1467-9280.2008.02136.x

Vulkan, N. (2000). An economist's perspective on probability matching. *Journal of Economic Surveys*, *14*(1), 101–118. http://doi.org/10.1111/1467-6419.00106

Weber, E. U. (2010). On the coefficient of variation as a predictor of risk sensitivity: behavioral and neural evidence for the relative encoding of outcome variability. *Journal of Mathematical Psychology*, *54*(4), 395–399. http://doi.org/10.1016/j.jmp.2010.03.003

Weber, E. U., Shafir, S., & Blais, A. (2004). Predicting risk sensitivity in humans and lower animals: risk as variance or coefficient of variation. *Psychological Review*, *111*(2), 430–445. http://doi.org/10.1037/0033-295X.111.2.430

Xu, E. R., & Kralik, J. D. (2014). Risky business: rhesus monkeys exhibit persistent preferences for risky options. *Frontiers in Psychology*, *5*. http://doi.org/10.3389/fpsyg.2014.00258

Yechiam, E., Busemeyer, J. R., Stout, J. C., & Bechara, A. (2005). Using cognitive models to map relations between neuropsychological disorders and human decision-making deficits. *Psychological Science*, *16*(12), 973–978. http://doi.org/10.1111/j.1467-9280.2005.01646.x

Figure captions

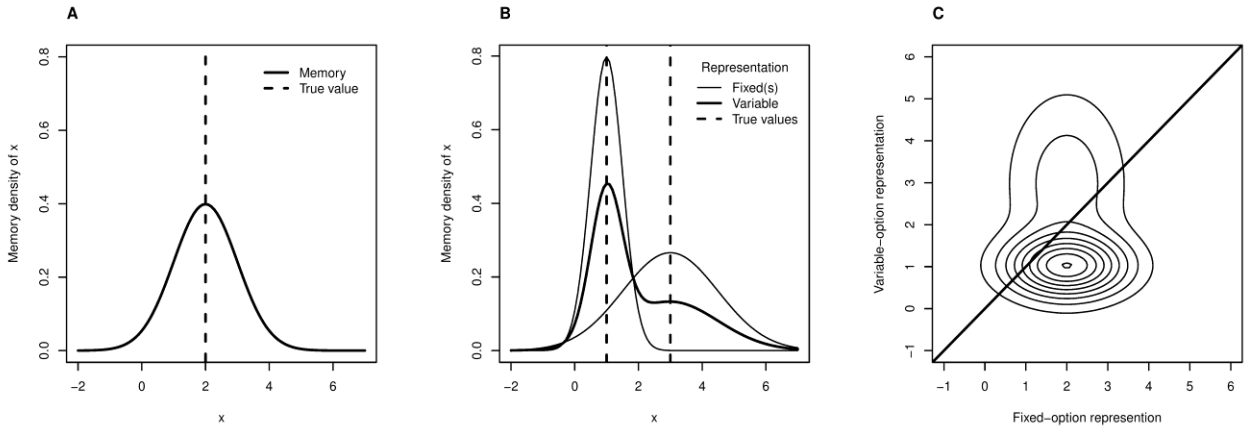

Figure 1. Typical application of Scalar Utility Theory (SUT). **A**) *If an animal has been trained in an experiment where a choice option always delivers a reward of magnitude 2 (vertical dashed line), the animal's operational (memory) representation is approximated by a normal distribution with mean 2 and standard deviation γ2 (solid line for probability density function; γ = 0.5 here).* **B**) *When a choice option yields a reward of 1 with probability of ½ and a reward of 3 with probability of ½ (mean is again equal to 2), then this option's representation is assumed to be a mixture of the representation (thick line) for the equivalent certain rewards (thin lines), with the mixing weights equal to the probabilities of respective cases of reward in the experiment: (½,½). Because the standard deviation of mental representation of a reward is assumed to be proportional to the reward value, a more dispersed representation for the bigger outcome is superposed on a less dispersed one for the lower outcome, which induces a skewed representation for the variable reward.* **C**) *Contours of the joint distribution of the independent fixed (horizontal axis) and variable (vertical axis) choice options. Because of the skew for the variable option and equal means for the fixed and variable option (i.e., 2), a larger share of the probability mass resides below the diagonal; i.e., in the set {fixed option > variable option}. Thus, in independent random samples the event {fixed option > variable option} occurs more often, implying risk-averse behaviour for variable reward amount.*

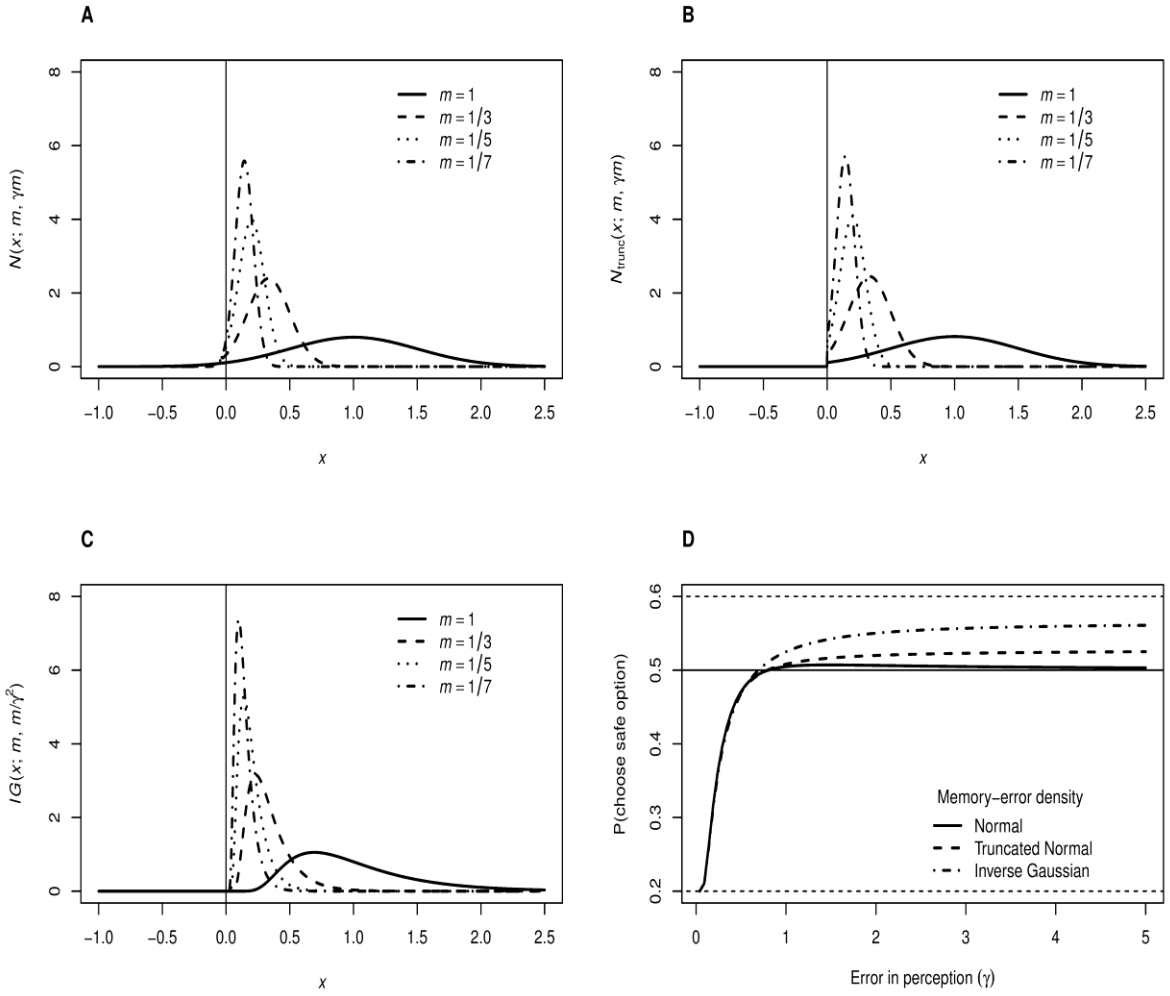

Figure 2. Zero rewards in SUT. **A)** *Irrespective of the (constant) value of* γ*, the standard deviation of SUT's memory representations,* γm*, tend towards zero as the encoded value,* m*, does so. A small probability of a negative memory draw exists, however.* **B**) *With a truncated Normal distribution, negative memories are avoided by setting the likelihood of negative values to zero and normalising the positive likelihoods so as to integrate to 1.* **C**) *The inverse Gaussian distribution is also subject to the scalar property, although defined only for positive rewards.* **D**) *SUT-predicted probability of choosing the safe option in the experiment of Shafir et al. (2008), as a function of* γ*. Solid line gives the usual SUT prediction, dashed line the one where SUT's normal densities have been replaced by zero-truncated normal densities, and dash-dotted line the prediction using Inverse Gaussian distribution (zero reward is represented by Dirac's delta distribution).*

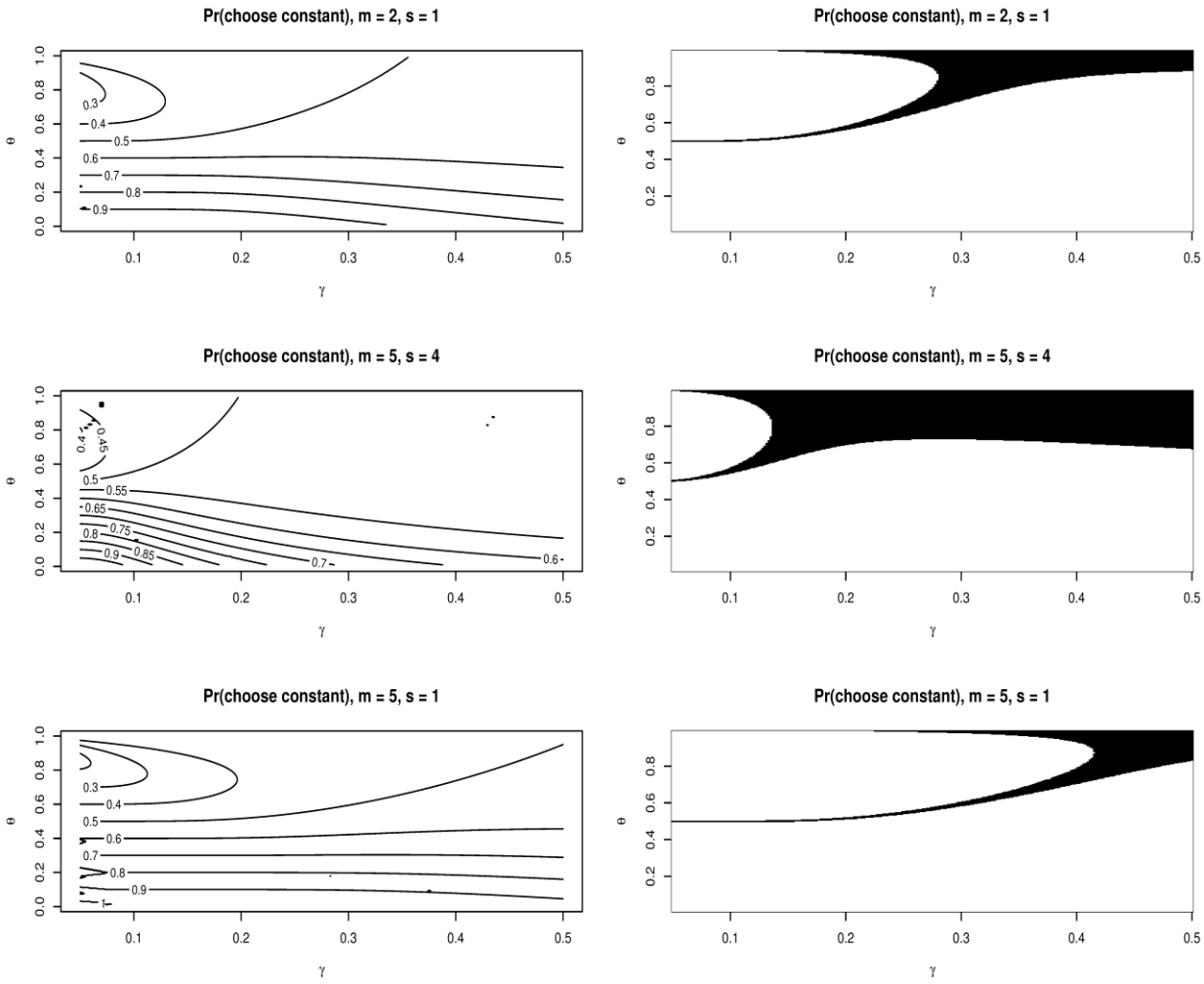

Figure 3. SUT prediction for two equal-mean outcomes as a function of γ *and* θ*. Left column of panels shows contour plots and right column indifference regions (black areas), where* $0.495 < P(S_2 > S_1) < 0.505$*, implying that the animal has no readily observable preference. Rows show the same for different values of mean (*m*) and small outcome of the variable option (*s*). The findings are almost exactly the same when replacing normal base distribution of the model with truncated normal or inverse Gaussian.*

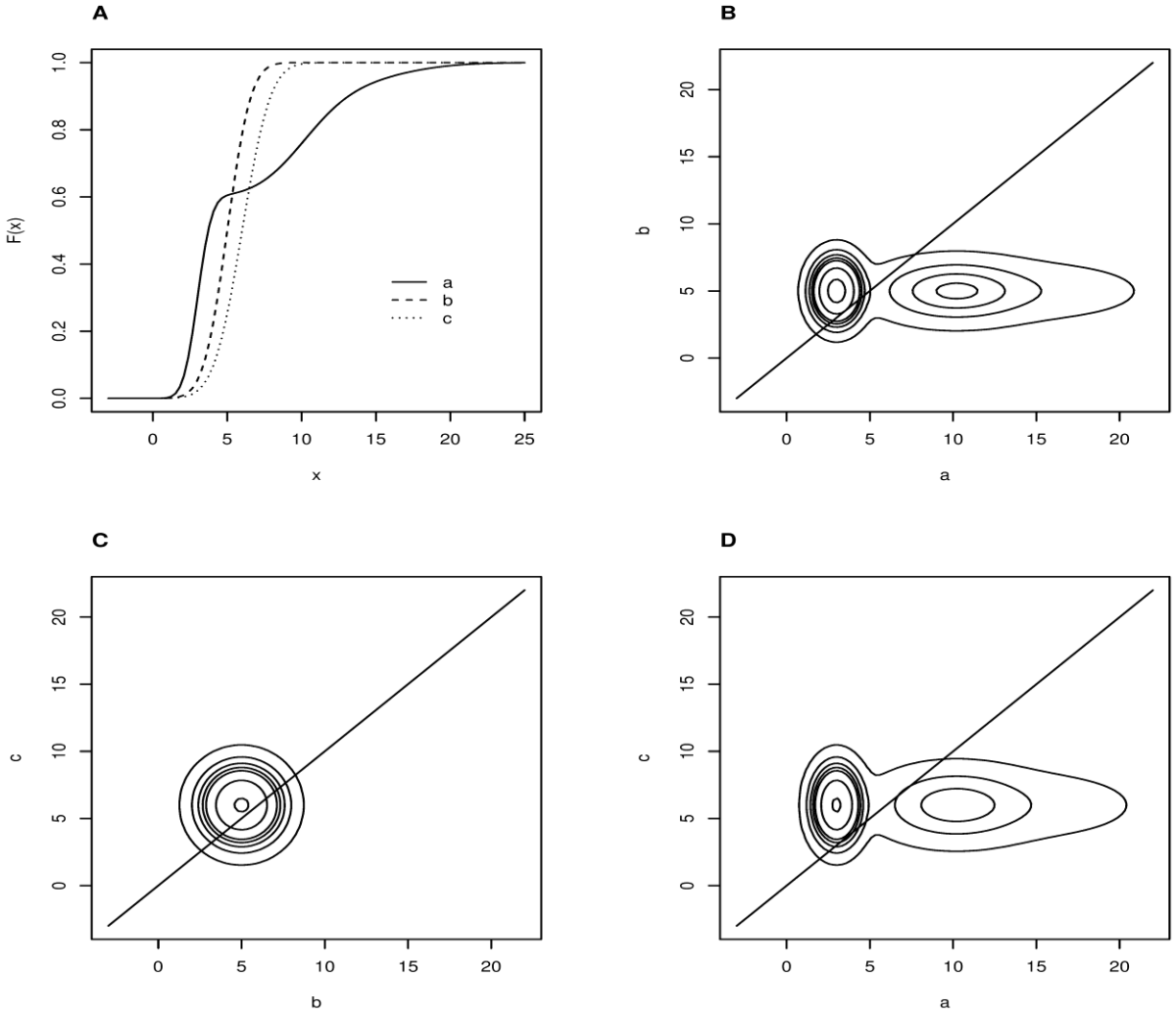

Figure 4. Strong Stochastic Transitivity (SST). An example where SST is violated is illustrated. **A**) *Cumulative distribution functions of alternative* a *(solid line),* b *(dashed line), and* c *(dotted line).* **B–D**) *Joint probability densities of the options. The higher mean of* c *compared to* b *takes more probability mass over the diagonal for the joint distribution of* c *and* b *than for the joint distributions of* c *and* a *or* b *and* a. *Because* b *was already preferred to* a *and* c *preferred to* b, *strong stochastic transitivity would require* c *to be preferred to* a *even more than it is preferred to* b*; but this fails to happen in the present SUT parametrisation.*

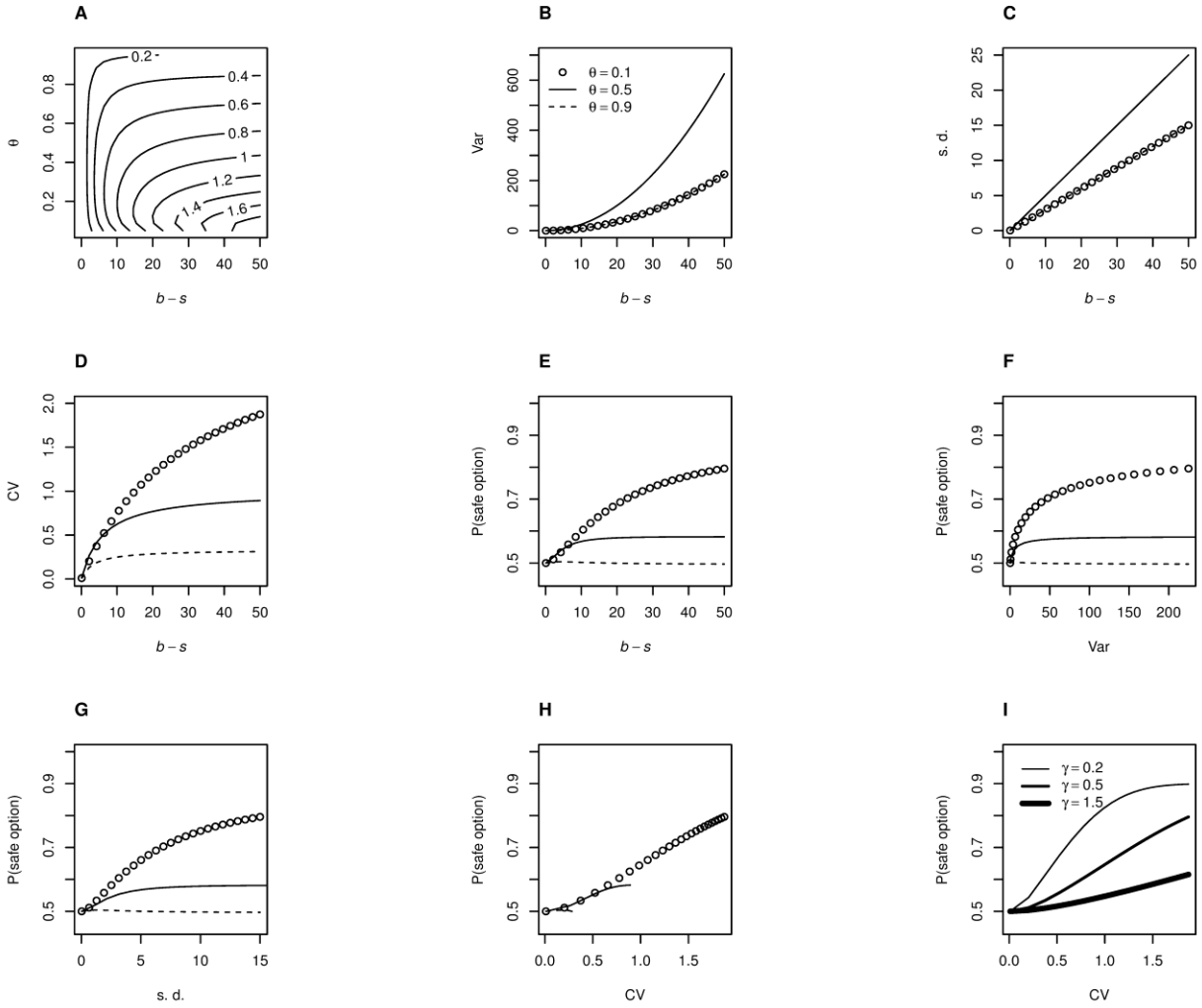

Figure 5. Statistics of variation compared with SUT prediction. Panels **A** *to* **D** *statistically characterize the risky option of the experiment, whereas panels* **E** *to* **I** *compare those statistics with corresponding predictions of SUT; the safe option has a fixed mean equal to that of the risky option.* **A**) *Contours plot for coefficient of variation (CV) of the two-outcome risky, or variable option as a function of the difference of outcomes (*b – s*) and the probability* θ *of getting the big reward* b *when choosing this option (*s *= 3). CV has been used to predict risk sensitivity, which is defined as absolute difference between probability of choosing the risky option instead of a safe option with equal mean and the indifference-value ½.* **B–E**) *Variance (Var), standard deviation (s. d.),and CV of the risky option, and associated SUT prediction as a function of* b – s, *for three different values of* θ. **F–H**) *SUT prediction compared to the variance, standard deviation, and CV.* **I**) *SUT prediction compared to CV when* θ *=0.1, for three different values of the scalar* γ.

Highlights

- Scalar Utility Theory (SUT) is a model used to predict animal and human choice behaviour in the context of reward amount, delay to reward, and variability in these quantities (risk preferences). Here it is shown that, contrary to previous claims:
- SUT can predict both risk averse and risk prone behaviour for both reward amounts and delays to reward depending on experimental parameters
- SUT implies violations of several concepts of rational behaviour
- SUT can predict, but does not always predict, a linear relationship between risk sensitivity in choices and coefficient of variation in the decision-making experiment
- In addition, a straightforward way to test the key assumptions of SUT is suggested.