The Dynamics of Selective Integration during Rapid Experiential Decisions

Konstantinos Tsetsos (k.tsetsos@uke.de)
Department of Neurophysiology & Pathophysiology, University Medical Center Hamburg
D-20246 Hamburg, Germany

Abstract

When making decisions humans often violate the principles of rational choice theory. Recent experiments, involving rapid experiential decisions, uncovered a mechanism that is responsible for various rationality violations. According to this *selective gating* mechanism, incoming value samples are accumulated across time, but prior to their accumulation they are weighted in proportion to their momentary rank-order. Here, using a data-driven approach, I present a dynamic extension of this mechanism, which involves potentially asymmetric inhibition between the inputs. As a result, and contrary to the previous selective gating implementation, the vigour of gating is modulated by the difference between two value samples (a *distance* effect) as well as by the absolute magnitude of the samples (a *magnitude* effect). This extension offers a superior explanation to existing and new data; and links high-level decision phenomena with computational principles previously described in theories of selective attention and visual search.

Keywords: Selective integration; Experiential decisions; Risk-seeking; Intransitivity

Introduction

Behaving organisms update their preferences in response to changes in their internal status and the state of the world; but they also do so in the absence of such changes. For instance, people are found to prefer A (e.g., fresh fish) over B (e.g., steak), when these two alternatives are offered, but B over A when a third – inferior and unchosen– alternative, C (e.g., frozen fish), becomes available (Huber, Payne, & Puto, 1982). Reversals of preference, as in the example above violate the axioms of rational choice theory (Von Neumann & Morgenstern, 2007) and indicate that the valuation of an alternative is *context-sensitive*, not only dependent on the agreement between the goals of the decision-maker and the properties of the judged alternative but also dependent on the properties of other alternatives in the choice set.

Context-sensitive valuation (*hereafter* CSV) phenomena are reported as preference reversals (Tsetsos, Usher, & Chater, 2010) or transitivity violations (Tversky, 1969), elicited in multiattribute choice experiments. Recently, analogues of these CSV phenomena were obtained in a rapid experiential decision task, labelled *value psychophysics* (Tsetsos, Chater, & Usher, 2012; Tsetsos et al., 2016). Obtaining CSV phenomena in a psychophysical task enabled the detailed computational modelling of the involved decision processes, pointing to a selective integration model (*hereafter* SI) that underlies several rationality violations.

According to SI, value information is accumulated across time but, prior to accumulation, higher value samples suppress lower samples via a *selective gating* mechanism. In the extant implementation of the model the vigour of this suppression is constant. For instance, a value sample of 60 will suppress a competing value sample of 59 in the same fashion that it would suppress a competing value sample of 49. Here, I focus exclusively on binary choices and show that this invariant mechanism fails to capture some qualitative patterns in past data. To overcome this limitation, I propose a dynamical mechanism in which the inputs compete against each other via potentially asymmetric inhibition.

In what follows, I describe the extant implementation of the SI model and its trademark behavioural signatures. Then, I use previously published data and outline qualitative patterns that SI in its current form cannot explain. These patterns are replicated in a new experiment. Third, I propose three extensions of the selective integration model and fit these extensions to the new data. Finally, I present results from a new experiment that decisively disentangles these three selective integration extensions.

Selective Integration: description and behavioural signatures

In the SI model for binary choices, value samples in support of the two alternatives arrive simultaneously [1] and are accumulated over time, as in sequential sampling models of perceptual discrimination and categorisation (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006). Importantly, the values that are momentarily higher are passed onto the accumulation layer unaffected but the relatively lower values are truncated, akin to an attentional process that selectively prioritises the accumulation of local winners over local losers. In this section I outline the mathematical details of the extant implementation of selective integration for binary choices and point to key behavioural phenomena predicted by the model.

Model description

The model described here applies to decisions based on two sequences of inputs, presented simultaneously. The two sequences have thus the same number of *samples* and each pair of samples is presented at a discrete time-step, for a fixed time interval. Here, based on the findings in Tsetsos et al. (2016), I assume that the incoming samples are not corrupted by noise prior to accumulation. The two sequences are labelled S_A and S_B, with $S_A(t)$ indicating the value of sequence A at the (discrete) sample t. Two accumulators (Y_A and Y_B) integrate the values of the sequences across time according to the following difference equations:

[1] It is suggested that the simultaneous processing of two competing samples emulates deliberation over two multiattribute options. This analogy works under the assumtpion that, in real-life multiattribute choices and on each moment, one attribute is considered and the corresponding attribute values serve as inputs into a preference formation process (Tsetsos et al., 2012)

$$Y_A(t) = (1-\lambda)\cdot Y_A(t-1) + I_A(t) + \xi\cdot\zeta_A(t) \quad (1)$$
$$Y_B(t) = (1-\lambda)\cdot Y_B(t-1) + I_B(t) + \xi\cdot\zeta_B(t) \quad (2)$$

In the above t indicates the current discrete time-step (or sample), λ is accumulation leakage, $I_{A,B}(t)$ is the input to the two accumulators on a given time-step, ξ is the standard deviation of the noise at the accumulation level and $\zeta_{A,B}(t)$ are standard Gaussian samples, independent from each other and across time-steps. The accumulators are initialised at 0: $Y_A(t) = Y_B(t) = 0$. At the end of the accumulation period (at $t = T$, with T being the total number of samples presented in each sequence) a decision is made in favour of the accumulator with the higher tally. If both accumulators end up with equal tallies, a decision is made randomly.

The inputs to the two accumulators ($I_{A,B}(t)$) reflect the modified sequence values after the selective integration filter is applied. I refer to this filter as *selective gating*. Selective gating is implemented as follows.

$$I_A(t) = \theta(S_A(t), S_B(t))\cdot S_A(t) \quad (3)$$
$$I_B(t) = \theta(S_B(t), S_A(t))\cdot S_B(t) \quad (4)$$

Function θ returns a value of 1 if the first argument is equal or larger than the second and a value w (selective gating parameter) otherwise:

$$\theta(x,y) = \begin{cases} 1 & \text{if } x >= y \\ w & \text{if } x < y \end{cases} \quad (5)$$

Behavioural signatures of selective integration

Pro variance (PV) effect. Consider two sequences, A and B, with values sampled from normal distributions with means μ_A, μ_B and standard deviations σ_A, σ_B. The *pro variance* effect (PV) occurs when participants choose more often sequence A when $\mu_A = \mu_B$ and $\sigma_A > \sigma_B$. Equivalently, the PV effect is present when accuracy is higher in trials where $\mu_A > \mu_B$ and $\sigma_A > \sigma_B$ (correct answer is A) relative to the accuracy in trials where the means are swapped ($\mu_A < \mu_B$ and $\sigma_A > \sigma_B$; correct answer is B). The PV effect was originally demonstrated in Tsetsos et al. (2012) and it was robustly replicated in Tsetsos et al. (2016). SI explains the PV effect as follows: a losing sample from the high variance distribution will more likely have low value. In the low variance distribution, a losing sample will more likely have a mediocre value. Multiplicatively downweighting a low value results in a smaller loss relative to downweighting a mediocre value (for a value of 30 and for $w = 0.5$ the loss is 15; for a value of 50 the respective loss is 25)(Tsetsos et al., 2012).

Frequent-winner (FW) effect. Consider two sequences A and B, consisting of the same three low (L), medium (M) and high (H) value samples such that $H - M = M - L$. The order of appearance of these three samples differs in the two sequences: $A \rightarrow LMH$ and $B \rightarrow HLM$. Thus, when presented simultaneously in that order, sequence A wins locally twice by a small margin (M vs. L and H vs. M) and loses once by a larger margin (L vs. H). According to the SI model ($w < 1$) a choice bias in favour of sequence A is expected since $M + H + L\cdot w > L\cdot w + M\cdot w + H$ or $M > M\cdot w$. This *frequent-winner* effect (*hereafter* FW) can also appear when accuracy is higher in trials in which all values in A are augmented by a small constant c (such that A has a higher total value and dominates B in 2 out of 3 samples), relative to the accuracy in trials in which all values of B are augmented by a small constant c (in that case B has a higher value but A still dominates B in 2 out of 3 samples). Importantly, SI can lead to intransitive preference cycles when a third sequence C with values MHL is considered. In that case, in the respective binary choices, A will be preferred over B, B will be preferred over C and C will be preferred over A. The FW effect, and the corresponding weak stochastic transitivity violations, were robustly obtained across 4 experiments, in which participants had to choose between bars of different length, presented sequentially (Tsetsos et al., 2016).

Challenges for Selective Integration

Under the selective integration framework the PV and FW effects occur due to the same mechanism, controlled by the selective gating parameter. It is therefore expected that the two effects will be strongly correlated across participants. However, a re-examination of the 4 experiments reported in Tsetsos et al. (2016) ($N = 93$) reveals no correlation between the two effects ($r = 0.000$). On the contrary, the effects predicted by fitting the SI model show indeed significant positive correlation ($r = 0.323, p = 0.002$). Additionally, for the same parametrisation, the SI model predicts a much stronger FW than PV effect (difference in the predicted effects: $M = 0.128, SE = 0.014, t(92) = 9.150, p < 0.001$). However, in the observed data this difference does not occur (difference in the observed effects: $M = 0.006, SE = 0.023, t(92) = 0.240, p = 0.811$). The model predicts well the magnitude of the FW effect but, although the predicted PV effect is significant ($M = 0.036, SE = 0.004, t(92) = 9.177, p < 0.001$), it is much smaller than the observed one (observed and predicted difference for the PV effect: $M = 0.102, SE = 0.014, t(92) = 7.274, p < 0.001$).

The lack of correlation between the PV and FW effects and the underestimation of the PV effect challenge the existing implementation of SI. It is conceivable, however, that these patterns are specific to the design and stimuli used in Tsetsos et al. (2016) (i.e. accumulation of lengths) and do not reflect generalisable limitations of the model. I examine next whether this is the case by characterising the PV and FW effects in an experiment that involves the accumulation of numerical values (c.f. Tsetsos et al. (2012)).

Experiment 1

Participants. 25 participants ($M_{age} = 28.1, SD_{age} = 6.4$, 14 female) with normal or corrected-to-normal vision and no

history of neurological or psychiatric impairment were recruited from Birkbeck's (University of London) participants pool. All participants gave informed consent to participated and all procedures were approved by the local ethics committee.

Task & Procedure. On each trial, participants observed pairs of black 2-digit numerical values presented rapidly and sequentially, to the left and right of a central fixation point and against gray background. The viewing distance was 60 cm and each numerical character was 0.93°wide and 1.5°long. After the presentation of 9 pairs of numbers, the central fixation point turned blue and participants were asked to choose which stream had on average the higher value. After giving a response the blue dot turned green (red) to indicate a correct (incorrect) response. The presentation rate of the numbers was 800 ms and 1 second gap was left between trials. Overall there were 4 blocks with 65 trials each. At the end of each block participants could see their accuracy so far. At the end of the experiment participants received £7 and a £2 bonus if their accuracy exceeded 75%.

Design. There were 4 types of trials (65 trials per type) in the experiment, presented in random order. In all trials there was a correct answer, with the sum of the higher sequence differing from the sum of the lower sequence by 72 units. Two types of trials were associated with the PV effect while the other two with the FW effect. In the PV trials the sequences were generated from Gaussian distributions, with the mean of the higher sequence (μ_H) sampled from $\mu_H \sim U(45,65)$. The mean of the lower sequence was $\mu_L = mu_h - 8$. In one type of PV trials, referred to as PV_1 trials, the standard deviation of the higher sequence was $\sigma_H = 20$ while the standard deviation of the lower sequence was $\sigma_L = 10$. In PV_2 trials the standard deviations changed with $\sigma_H = 10$ and $\sigma_L = 20$. The accuracy difference between PV_1 and PV_2 trials quantifies the PV effect. In the FW trials, the mean values of the higher and lower sequences were matched to those in the PV trials. However, the temporal distribution of the sequences was manipulated such that one alternative always dominated the other in 6 out of 9 samples (see also Tsetsos et al. (2016)). In FW_1 trials the higher sequence dominated the lower sequence more often. When the higher sequence dominated, it did so by $U(23,28)$ units and in the less often cases when the lower sequence dominated it did so by $U(22,32)$ units. In FW_1 trials the higher sequence dominated the lower sequence in 3 out of 9 samples only. When the higher sequence dominated it did so by $U(38,48)$ units and when the lower sequence dominated it did so by $U(7,13)$ units. The accuracy difference between FW_1 and FW_2 trials quantifies the FW effect. In all trials the generated sequences were constrained so as to involve only 2-digit numbers ranging from 10 to 90.

Results. Participants performed above chance in all trials (accuracy: $M = 0.774, SE = 0.023$). Accuracy in PV_1 trials was higher than in PV_2 trials ($M = 0.167, SE = 0.032, t(24) = 5.170, p < 0.001, d = 1.034$) replicating thus the PV effect obtained elsewhere (Tsetsos et al., 2012, 2016). Accuracy in FW_1 trials was higher than in FW_2 trials ($M = 0.050, SE = 0.017, t(24) = 2.897, p = 0.008, d = 0.579$). This finding replicates with different stimuli (i.e. numbers) the effect reported in Tsetsos et al. (2016). Contrary to what SI predicts, there was no correlation between the two effects ($r = -0.014, p = 0.949$) and the FW effect was weaker than the PV effect.

Extensions of selective integration

The challenges that were identified for SI using the datasets from Tsetsos et al. (2016) persist in Experiment 1: the two critical effects were not correlated and the PV effect was larger than the FW effect. To address these challenges I here propose 3 extensions of the SI model. One extension is static, with selective gating invariance as in the original model, while the other two are dynamic and biologically inspired. The primary aim of these extensions is to decorrelate the PV and FW effects. Hereafter, the baseline SI model described earlier will be referred to as $\mathcal{M}_{S0}$.

A static extension

The first SI extension involves a transducer function that transforms objective values into their subjective counterparts. The model is thus identical to the one described in Eq. 1-5, with the exception that Eq. 3-4 take the form:

$$I_A(t) = \theta(S_A(t)^\alpha, S_B(t)^\alpha) \cdot S_A(t)^\alpha \quad (6)$$

$$I_B(t) = \theta(S_B(t)^\alpha, S_A(t)^\alpha) \cdot S_B(t)^\alpha \quad (7)$$

Exponentiating the inputs allows the PV effect to occur independent of the selective gating parameter when $\alpha > 1$. In such cases the value function is convex resulting in a risk-seeking bias. Although convex low-level representation is undocumented in numbers (Feigenson, Dehaene, & Spelke, 2004) or lengths (Stevens, 1957)), it is possible that, at a higher processing level, large quantities stand out. This model will be referred to as $\mathcal{M}_{S1}$.

Dynamic extensions

Implementations $\mathcal{M}_{S0,1}$ are static in nature. Here I explore dynamical implementations, in which selective gating falls out from continuous competition between the input units. The competition is mediated by inhibition as in models of visual attention (Lee, Itti, Koch, & Braun, 1999). To illustrate the basic idea, the incoming values compete against each other and the activation states of the input units feed continuously, in cascade (McClelland, 1979), to the accumulation level (variables Y in Eq. 2). Almost equivalently, and to maintain comparability between static and dynamic SI implementations, I assume that a given accumulator receives discrete updates. These updates are equal to the temporal average of the activation in the corresponding input unit, for the period during which the stimulus was presented. The first dynamic model, labelled $\mathcal{M}_{D1}$, involves mutual inhibition between the

input units (Usher & McClelland, 2001). Eq. 1-2 remain intact but Eq. 3-5, which implement the selective gating filtering, are replaced as follows:

$$I_A(t) = \frac{\int_0^P X_A(x)dx}{P} \quad (8)$$

$$I_B(t) = \frac{\int_0^P X_B(x)dx}{P} \quad (9)$$

Variables $X_{A,B}$ reflect the input units. Variable P is the duration (in units of time) that a given pair of samples is presented for and dx is a small time interval (set in simulations to $dx = 0.001$ seconds). The input units are initiated at 0 and their dynamics are governed by the following coupled differential equations:

$$dX_A = (-\kappa X_A - \beta f(X_B) + S_A(t))dx \quad (10)$$

$$dX_B = (-\kappa X_B - \beta f(X_A) + S_B(t))dx \quad (11)$$

In the above, κ is a leak parameter (set to 1 throughout this paper), f is the identity function ($f(x) = x$) and β is the strength of mutual inhibition. The two input units thus reflect sustained input, corresponding to the presented values, $S_{A,B}(t)$ at time-step t. For simplicity, the above equations are deterministic, consistent with the finding that during value psychophysics noise at the representation level is negligible (Tsetsos et al., 2016). The input units are subject to a reflecting boundary at 0 that prevents activation states from being negative:

$$X_A = max(X_A, 0) \quad (12)$$

$$X_B = max(X_B, 0) \quad (13)$$

When $\beta = 0$, the two input units quickly converge to their nominal values. When inhibition is present, the larger value suppresses the smaller value, implementing that way a form of selective gating. Figure 1a shows examples of the evolution of activation states in the input units.

The second dynamic extension ($\mathcal{M}_{D2}$) is identical to $\mathcal{M}_{D1}$ with the exception that, following Brown and Holmes (2001), each unit inhibits the other via a sigmoid activation function.Thus, f, which was the identity function in $\mathcal{M}_{D1}$, now becomes:

$$f(x) = \frac{1}{1 + e^{(-g(x-b))}} \quad (14)$$

In the above, g is the slope of the activation function (here set to 1) and b the inflection point of the sigmoid. In other words, b controls when selective gating will kick in while the inhibition strength (β) controls the strength of selective gating. The difference between ($\mathcal{M}_{D1}$) and ($\mathcal{M}_{D2}$) is that, in the former, inhibition is mutual while in the latter inhibition can be non-reciprocal (Figure 1b) and inactive from inputs of low value.

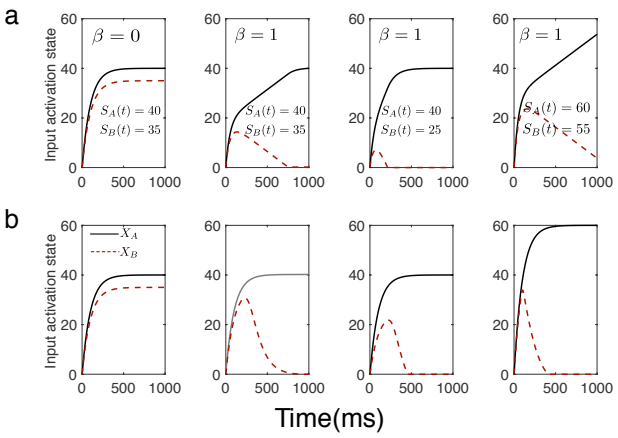

Figure 1: (a) Input activation trajectories for $\mathcal{M}_{D1}$ for different value samples and for $\kappa = 1$. Leftmost panel depicts the special case without inhibition. (b) Same as (a) but for $\mathcal{M}_{D2}$ and $b = 38$, $g = 1$ and $\kappa = 1$.

Comparison of selective integration extensions

Quantitative comparison of the models

Here I fit the models to the choice data of each participant from Experiment 1. Model predictions for each trial are derived numerically. The negative log likelihood of each parametrisation is calculated on a trial-by-trial basis and summed across trials. $\mathcal{M}_{S0}$ has three free parameters (w, ξ, λ) and $\mathcal{M}_{S1}$ has one extra free parameter (α). For the dynamic extensions, some parameters are set to fixed values ($\kappa = 1, g = 1, P = 500ms, dx = 1ms$), which leaves $\mathcal{M}_{D1}$ with 3 free parameters (σ, λ, β) and $\mathcal{M}_{D2}$ with a fourth parameter (b). The models are compared based on their BIC values, aggregated across participants, and also on a participant-by-participant basis (i.e. the proportion of participants for which a model has the lowest BIC score). Additionally, the correlation between the PV and FW predicted effects is examined. The results are summarised in Table 1.

Table 1: Model comparison in Experiment 1.

Model	Total BIC	% BIC lowest	r
$\mathcal{M}_{S0}$	8,458	28%	0.736
$\mathcal{M}_{S1}$	8,554	0%	0.254
$\mathcal{M}_{D1}$	8,441	24%	0.779
$\mathcal{M}_{D2}$	8,398	48%	0.087

$\mathcal{M}_{D2}$ explains the data best and succeeds to decorrelate the PV and FW effects (last column in Table 1). It is also the only model that does not underestimate dramatically the PV effect and predicts that it will have higher magnitude than the FW effect (Figure 2). In this model, the PV effect is partly driven by parameter b. If this parameter is set above the middle of the value range, selective gating will be inactive for comparisons between mediocre and low values, further exaggerating

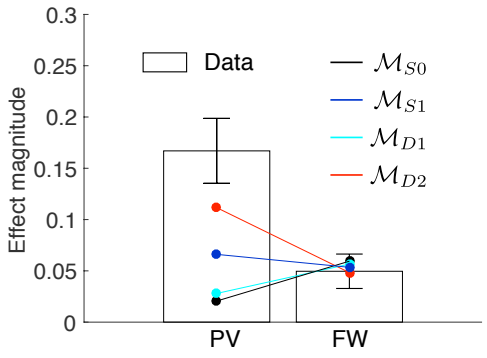

Figure 2: Data and model fits in Experiment 1. Error bars correspond to 2SE.

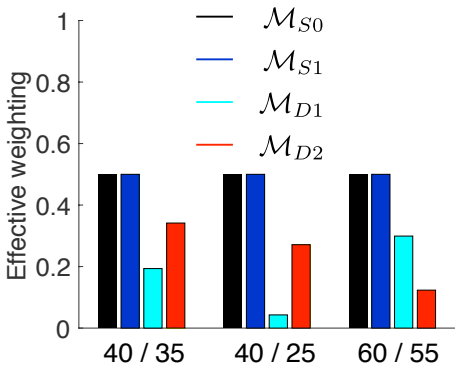

Figure 3: Effective weighting is the ratio between the filtered lower sample and the corresponding unfiltered values. The unfiltered values in static models are nominal values (without selective gating, $w = 1$) and in dynamic models values, as per Eq.8-9, in the $\beta = 0$ instantiation . The values of the different pairs are given in the x-axis.

the choice bias for a high variance sequence. The FW effect is independent of the b value (as long as it is not too high, deactivating altogether selective gating). $\mathcal{M}_{S1}$ also succeeds to decorrelate the effects but provides a poor fit. Importantly, $\mathcal{M}_{D1}$ appears to suffer from the same limitation as $\mathcal{M}_{S0}$: selective gating is controlled by one parameter (β) and the PV and FW are strongly correlated.

Qualitative differences among models

The distinction between the four selective integration implementations is the way selective gating is implemented. In the two static implementations ($\mathcal{M}_{S0,1}$) the weight applied on the local loser is invariant to the difference between the winning and losing value samples. On the contrary, in the dynamic extensions, this difference matters. In Figure 3, I show the effective weight applied on the losing sample for several combinations of pairs of values (c.f. Figure 1). Static implementations predict, by definition, invariance of weighting (here for $w = 0.5$). $\mathcal{M}_{D1,2}$ both predict that as the difference between the two inputs increases (c.f. leftmost and middle bars), suppression of the loser increases too. This is reminiscent of the *distance* effect encountered in numerical cognition (Moyer & Landauer, 1967).

Adding a constant to both input samples (c.f. leftmost and rightmost bars) results in opposing predictions in the two dynamic models. $\mathcal{M}_{D1}$ predicts weaker suppression for increased values, since the competition between the two inputs takes longer to resolve. $\mathcal{M}_{D2}$ predicts a stronger suppression of the loser, since the winning unit will breach first the b barrier and will start inhibiting the other unit strongly, resulting in enhanced winner-take-all dynamics (Figure 1, rightmost panels). All models thus make distinctive qualitative predictions regarding a *magnitude* effect, which I exploit in the next experiment.

Experiment 2

In this experiment I examine how the PV effect changes when all sequence values increase or decrease by the same constant amount (Figure 4a). According to static models no change is expected, since selective gating is invariant to the magnitude of the values. $\mathcal{M}_{D1}$ predicts weaker selective gating for high values, while $\mathcal{M}_{D2}$ predicts a stronger PV effect for high values.

Participants. 18 participants ($M_{age} = 26.4, SD_{age} = 5.2$, 11 female) took part. The rest of details are as per Experiment 1.

Task & Procedure. The task and procedure was identical to Experiment 1 except that there were 12 samples in each stream and that the presentation rate was 0.5 seconds. Overall participants did 6 blocks with 50 trials each.

Design. The PV effect was elicited as in Experiment 1, using two trial types ($\text{PV}_{1,2}$) and examining the accuracy difference between them. Here, there were 3 conditions giving rise to 6 trial types (50 trials for each type, randomly presented). In the baseline condition ($\text{PV}_{B_{1,2}}$), the correct sequence had always a mean of 50 and the incorrect a mean of 42. In the *negative offset* condition ($\text{PV}_{-_{1,2}}$), for a given trial, 6 pairs of values were created as per the baseline condition. The remaining 6 pairs were created by subtracting from the mean of both Gaussians a constant ($c = 15$). The regular and lower pairs were presented in random order in a given trial. Equivalently, in the positive offset condition ($\text{PV}_{+_{1,2}}$) a constant ($c = 15$) was added to the values of 6 pairs.

Results. The PV effect increased as both sequences increased in absolute values as indicated by a repeated measures ANOVA ($F(2,34) = 18.74, p < 0.001, n^2 = 0.524$). Tukey post-hoc tests revealed that the PV effect was lower in the negative offset condition relative to the baseline ($p < 0.001$) and the positive offset ($p < 0.001$) conditions. The difference between the baseline and the positive offset condition was not significant ($p = 0.171$). As predicted, this PV increasing pattern was solely captured by $\mathcal{M}_{D2}$. The advan-

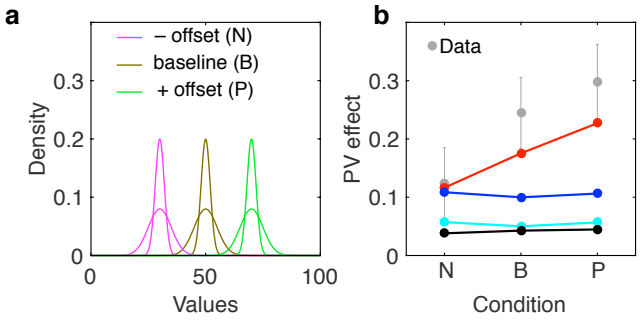

Figure 4: (a) Outline of experimental design. (b) PV effect as a function of negative offset (N), baseline (B) and positive offset (P) conditions for data and models (colour code for models as per Figures 2-3). Error bars correspond to 95% CI.

tage of $\mathcal{M}_{D2}$ in this experiment was clear also in quantitative comparisons (data fitted for each participant separately) using BIC (Table 2).

Table 2: Model comparison in Experiment 2.

Model	Total BIC	% BIC lowest
$\mathcal{M}_{S0}$	5,976	6%
$\mathcal{M}_{S1}$	5,466	0%
$\mathcal{M}_{D1}$	5,448	6%
$\mathcal{M}_{D2}$	5,253	88%

Conclusion

Selective integration is a decision making model that has successfully explained several rationality violations. One potential criticism against this model is that its applicability is limited to rapid decisions from experience, in which attentional demands are increased. However, it has been previously shown that selective gating increases under lower attentional demands (Tsetsos et al., 2016). Additionally, behavioural signatures that are routinely obtained in high-level decisions are also obtained in rapid experiential decisions, implying that the latter can offer a window to more complex choice mechanisms.

Here I presented challenges for the extant implementation of selective integration. These challenges were successfully addressed by a dynamic extension of the model, in which the inputs compete for accumulation via inhibiting each other, as in models of selective attention and visual search. This dynamic extension predicts that the vigour of selective integration increases both when the distance and the absolute magnitudes of the two inputs increase. This prediction was experimentally confirmed. Oveall, this dynamic and biologically inspired (Usher & McClelland, 2001) extension presented here, significantly improves the descriptive adequacy of the selective integration framework and facilitates its validation at the neurophysiological level.

Acknowledgments

I thank Artemis Maipa for assistance with data collection. This work was funded by a Wellcome Career Development fellowship, a British Academy/ Leverhulme Trust awarded and a Marie Curie Individual Fellowship awarded to the author.

References

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological review*, *113*(4), 700.

Brown, E., & Holmes, P. (2001). Modeling a simple choice task: stochastic dynamics of mutually inhibitory neural groups. *Stochastics and dynamics*, *1*(02), 159–191.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in cognitive sciences*, *8*(7), 307–314.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of consumer research*, *9*(1), 90–98.

Lee, D. K., Itti, L., Koch, C., & Braun, J. (1999). Attention activates winner-take-all competition among visual filters. *Nature neuroscience*, *2*(4), 375–381.

McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological review*, *86*(4), 287.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature*, *215*(5109), 1519–1520.

Stevens, S. S. (1957). On the psychophysical law. *Psychological review*, *64*(3), 153.

Tsetsos, K., Chater, N., & Usher, M. (2012). Salience driven value integration explains decision biases and preference reversal. *Proceedings of the National Academy of Sciences*, *109*(24), 9659–9664.

Tsetsos, K., Moran, R., Moreland, J., Chater, N., Usher, M., & Summerfield, C. (2016). Economic irrationality is optimal during noisy decision making. *Proceedings of the National Academy of Sciences*, *113*(11), 3102-3107. doi: 10.1073/pnas.1519157113

Tsetsos, K., Usher, M., & Chater, N. (2010). Preference reversal in multiattribute choice. *Psychological review*, *117*(4), 1275.

Tversky, A. (1969). Intransitivity of preferences. *Psychological review*, *76*(1), 31.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychological review*, *108*(3), 550.

Von Neumann, J., & Morgenstern, O. (2007). *Theory of games and economic behavior*. Princeton university press.