COMMENTARY

Examining the Mechanisms Underlying Contextual Preference Reversal: Comment on Trueblood, Brown, and Heathcote (2014)

Konstantinos Tsetsos  
University of Oxford

Marius Usher  
University of Tel Aviv

Nick Chater  
University of Warwick

Trueblood, Brown, and Heathcote (2014) provide a new model of multiattribute choice, which accounts for 3 contextual reversal effects (similarity, attraction and compromise). We review the details of the model and highlight some novel predictions. First, we show that the model works by setting a “fine balance” between 2 opposing factors that influence choice. As a result, small changes in the attributes of choice alternatives can disturb this balance. Second, we show that the model gives a partial account of the compromise effect. We describe a number of experiments that could distinguish the MLBA from other models of multiattribute choice.

Keywords: decoy effects, multiattribute choice, preference reversal

Supplemental materials: http://dx.doi.org/10.1037/a0038953.supp

Almost every nontrivial choice involves balancing different attributes. Do we prefer a dull but cheap vacation or an exciting but alarmingly expensive one? A small flat with a short commute, or a large house with a long commute? Experimental research has revealed that such choices are subject to a puzzling set of contextual reversal of preferences, which, although at odds with rationality principles, has the potential to reveal the cognitive mechanisms that underlie choice. To illustrate, a customer may prefer a salad dessert to a cheesecake; but switch to the cake when the waiter points out the availability of another (less attractive) cake. This is the attraction effect (also labeled the asymmetric dominance effect; Huber, Payne, & Puto, 1982). Two other contextual preference reversals effects, the similarity (Tversky, 1972) and compromise (Simonson, 1989), show that the preference between two multiattribute alternatives (A vs. B) changes systematically when a third alternative is added (see Figure 1 for illustration).

Contextual reversal effects are inconsistent with option-based theories of multiattribute choice (which include almost all economic models of choice; e.g., McFadden, 1974), in which each option is evaluated independently; and hence threaten the very idea that choices arise from stable and context-independent utility functions. Intensive research has explored the psychological mechanisms underlying context-dependent evaluation and preference formation (Bhattacharya, 2013; Dhar & Glazer, 1996; Pettibone & Wedell, 2000; Roe, Bussemeyer, & Townsend, 2001; Soltani, De Martino, & Camerer, 2012; Tsetsos, Usher, & Chater, 2010; Tversky & Kahneman, 1991; Tversky & Simonson, 1993; Usher & McClelland, 2004). However, so far, only a minority of theories have captured simultaneously all three context effects (Bhattacharya, 2013; Roe et al., 2001; Usher & McClelland, 2004; Wollschlager & Diederich, 2012).

Recently, Trueblood and colleagues (Trueblood, Brown, & Heathcote, 2014) provided an excellent review of the relevant literature and presented an innovative new model—the Multiattribute Linear Ballistic Accumulator (MLBA)—that predicts the three contextual prefer-
The structure of the comment is as follows. We briefly summarize the models, deriving contrastive predictions for future experimental tests. The aim of this Comment is twofold. First, we examine the MLBA and show that despite its substantial merits, including analytical tractability, the model works by finely balancing competing forces. As a result, its predictions are sensitive to the exact position of the alternatives in the attribute-space. Second, we discuss the various components that are at play in the MLBA and in a number of other related models, deriving contrastive predictions for future experimental tests. The structure of the comment is as follows. We briefly summarize the MLBA model; relate it to previous computational theories of preference reversal in multiattribute choice; evaluate its stability to stimulus reversal effects. Further, their model was shown to outperform a seminal model of multiattribute choice, decision field theory (DFT; Roe et al., 2001) in fitting experimental data, while also capturing the influence of time pressure on the magnitude of context-dependencies.

The MLBA Model

The model consists of two parts: (a) a front-end component, which maps choice alternatives represented in a 2-dimensional attribute space (e.g., Figure 1) into selection tendencies (or drift rates) and (b) a back-end component, which maps these selection tendencies onto choice probabilities.

Front-End Component

Assuming three alternatives that differ in two attributes (e.g., price \([P]\) and quality \([Q]\), Figure 1), the front-end component calculates selection tendencies or drift rates \(d_i\) for each alternative \(i\) using the following formulas:

\[
d_1 = V_{12} + V_{13} + I_0
\]

\[
d_2 = V_{21} + V_{23} + I_0
\]

\[
d_3 = V_{31} + V_{32} + I_0
\]

(1)

The term \(V_{ij}\) reflects the result of comparing options \(i\) and \(j\), whereas \(I_0\) is simply a constant ensuring that at least one option is chosen in the back-end process. The comparison terms, \(V_{ij}\), are derived as follows. The first step is to map objective attribute values to their subjective counterparts. Thus, a choice alternative with objective attribute values \((x, y)\) will be perceived as being worth \((U_A, U_E)\) with the subjective values, \(U_A\) and \(U_E\) being both two dimensional functions of \(x\) and \(y\), defined via a projection from the indifference line \(x + y = K\) to a convex curve characterized by the parameter \(m\) (see Figure 2):

\[
\left(\frac{x}{a}\right)^m + \left(\frac{y}{b}\right)^m = 1
\]

(2)

A more detailed description of this mapping is given in the Appendix in Trueblood et al. (2014).

If \(U_i = (U_{pi}, U_{qi})\) and \(U_j = (U_{pj}, U_{qj})\) are the subjective attribute values for options \(i\) and \(j\) then the comparison terms, \(V_{ij}\), described above, are specified as follows:

\[
V_{ij} = U_{pi} - U_{pj}
\]

Figure 2. Projection from objective to subjective attributes \((U_i)\). The linear indifference curves (diagonal lines with negative slope) are transformed into convex curves for \(m = 5\) (solid curves) following Equation 2. Arrows show transformation from objective \((x, y)\) space, to subjective, \((U_i, U_E)\) space. With this transformation, midrange alternatives have higher subjective values (e.g., compare \(U_A\) to \(U_E\)). Alternative \(E\) is constructed so as the subjective utilities of \(A, B, \) and \(E\) are roughly equal (\(U_A \approx U_B \approx U_E\)). We use the \([A, B, E]\) triplet to show that in the Multiattribute Linear Ballistic Accumulator \((MLBA)\) the compromise effect is conditioned on the subjective inferiority of one of the extremes.
The weights $w_{P_i}$ and $w_{Q_i}$ correspond to the attentional weight given to each dimensional comparison. These weights embody exponentially decaying functions of subjective distance, on each dimension:

$$w_{P_i} = \exp\left(-\lambda |U_{P_i} - U_{P}\right)$$
$$w_{Q_i} = \exp\left(-\lambda |U_{Q_i} - U_{Q}\right)$$

Thus, the greater the subjective distance between two options on a given dimension, the less weight this comparison will have. Crucially, the decay parameter $\lambda$ can differ depending on whether the item of interest, $i$, is better or worse than the comparison item, $j$, on a particular dimension—that is, whether that dimension exhibits an advantage or a disadvantage for $i$. In particular, to account for the three context reversal effects simultaneously, the MLBA has to use weights for disadvantages which decay more rapidly than the corresponding weights for advantages: that is, $\lambda_\text{A} < \lambda_\text{D}$, where $\lambda_\text{A}$ is the decay for advantages and $\lambda_\text{D}$ is the decay for disadvantages. As we show below and as discussed in the original paper, this asymmetry is essential for explaining the similarity effect.

**Back-End Component**

The back-end component of the MLBA is a decision process based on the linear ballistic accumulator (LBA) model (Brown & Heathcote, 2008), which generates choices from the drifts, $d_i$, calculated in the front-end component. The LBA choice mechanism has analytical solutions and hence is particularly convenient. Within the MLBA two methods can be used to determine responses (decisions and their reaction times [RTs]). The first one involves a self-terminating response, in which a decision-criterion is used and the choice is won by the decision-accumulator whose activation first reaches the criterion. The second method corresponds to an externally controlled decision-time, in which the accumulator whose activation is highest at the time the response is requested determines choice. Here we follow Trueblood et al. (2014) and use the self-terminating version of the model with LBA parameters: $A = 1$, $\chi = 2$, $s = 1$.

The MLBA explains the attraction, compromise and similarity effects, as applied to the sets of alternatives shown in Figure 1/Table 1, with a single set of parameters (e.g., $m = 5$, $\lambda_\text{A} = 0.2$, $\lambda_\text{D} = 0.4$, $I_0 = 5$ given in Trueblood et al., 2014). Before we evaluate the MLBA qualitatively, we attempt to classify its mechanisms in relation to previous theories of contextual preference reversal.

**Classification of Mechanisms of Contextual Preference Reversal**

Trueblood et al. (2014) compare the MLBA to two alternative computational models: decision-field theory (DFT; Roe et al., 2001) and the leaky-competitor accumulator (LCA; Usher & McClelland, 2004). These theories build on the tradition of sequential sampling models of information integration (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006; Busemeyer & Townsend, 1993; Smith & Ratcliff, 2004; Usher & McClelland, 2001) and are dynamic in nature, predicting not only choice outcomes but also the moment-by-moment fluctuations of the corresponding preference states. A related and more recent model of contextual preference reversal in multiattribute choice (published at, roughly, the same time as the MLBA) is the associative accumulation model (AAM; Bhattacharya, 2013), which we will consider in our discussion below. In the light of the complexity of these theories, it is crucial to focus on the central mechanisms that generate contextual reversal effects and not on peripheral assumptions about implementational or dynamical details of the models.

To illustrate, the MLBA consists of a front-end (central) process that produces drift-rates for a given choice scenario and a back-end (peripheral) process that converts the drifts to choice probabilities. Because we are mainly concerned with accounting for choice data, we will focus now on the MLBA front-end process, and we compare it with the processes that operate in DFT and LCA and AAM. Back-end predictions concerning RTs are discussed later. Since the model parameters of the MLBA were fitted to the preference-reversal data, we will refer to the MLBA whose parameters are specified by this fit, as the MLBA account or mechanism of the reversal effects.

**Comparisons With All Items in the Choice Set**

A central assumption common to DFT, LCA, and MLBA (as well as in earlier static models; Tversky & Kahneman, 1991) is that the value of an alternative is determined by comparisons with...
all the items in the choice set (Eq. 1). For example, in a three-alternative choice the value of an option will be determined by its comparison with the other two options. For most models that predict context effects, additional nonlinearities in the valuation process are required.\(^1\) By contrast, in the DFT and MLBA the nonlinearities are introduced via distance-dependent interactions (Equations 3, 4). In the LCA the value function is assumed to be asymmetric, having loss-aversion that penalizes options that have large disadvantages (Usher & McClelland, 2004). As we discuss below, the MLBA mechanism also treats advantages and disadvantages in an asymmetric fashion.

**Distance-Dependent Interactions**

This is a similar feature between DFT and MLBA (Equation 4), implemented by distance-dependent inhibition in the former and distant-dependent comparison weights in the latter (see Figure 3). Distance-dependency allows both models to capture the attraction effect as the target alternative (A) that dominates the decoy (e.g., RF), is “compared” more with the inferior decoy, so obtaining a larger net advantage than the competitor (B).

**Asymmetric Weights for Advantages and Disadvantages**

Both the LCA and the MLBA embody an asymmetry in the weighting of advantages and disadvantages but differ on the direction of this asymmetry. In the LCA, following Tversky and Kahneman (1991), the asymmetry reflects loss-aversion (Kahneman & Tversky, 1979): Disadvantages loom larger than advantages. For example, the compromise effect is explained because the extreme options have large disadvantages on their unfavorable dimensions, which are penalized more than the two moderate disadvantages of an average option.

In the MLBA, advantages count more than disadvantages and this asymmetry increases with distance (Figure 3; monotonic curves for the relative weights as a function of the advantage/disadvantage magnitude). This mechanism allows MLBA to predict the similarity effect, because a new option S similar to A contributes more (on the corresponding dimension; i.e., P) to a distant alternative, B, than to the nearer alternative, A. In Figure 3, we show the attentional weight for a given advantage or disadvantage (monotonic curves), as a function of the distance, x, between two alternatives on a given dimension (we assume m = 1, here, but this is not essential). To illustrate the similarity effect we assume comparisons of options A and B relative to S (from Figure 1). Because all three options lie on the same indifference curve, pairwise comparisons between them will always give an advantage on one dimension and an equal disadvantage on the other dimension, before the weighting in Equation 4 takes place. However, for the competitor, B, the addition of S gives a larger advantage/disadvantage than it does for A. The solid non-monotonic curve in Figure 3 shows the net value (now using the weights in Eq. 4) from an equal advantage/disadvantage that S adds to either A or B, as function of their horizontal distance from S. As long as this quantity is increasing a similarity effect will ensue (compare the black circle showing the net value conferred to A by S with the gray circle for the boost conferred to B by S). Thus, for alternative sets such as those in Figure 1, the comparison with S contributes more to the distant alternative B than to the proximal one, A. Note that greater added values to distant alternatives reverses the direction of the loss-aversion asymmetry, which penalizes extreme alternatives (Tversky & Simonson, 1993; Usher & McClelland, 2004). Because loss-aversion is useful in penalizing extreme alternatives, the MLBA requires a different mechanism to explain the compromise effect (see below).

**Nonlinear Subjective Attribute Space**

This assumption is used only in the MLBA and reflects the necessity to set parameter m > 1 (Equation 2) to simultaneously capture all three effects. Accordingly, the equal preference curves in the attribute space are not straight diagonal lines but convex curves (see Figure 2). This implies a bias for alternatives in the middle of the range. Hence, even for binary choice (5,5) will be preferred to (6,4). Using this assumption, the MLBA explains the compromise effect. In choices between the alternatives A = (4,6), B = (6,4) and A’ = (2,8), introducing the latter alternative results in a preference of A over B.

**Sequential Stochastic Samplings of the Attributes**

The broad assumptions concerning the sequential stochastic sampling of attributes (SSSA) are common to the DFT, LCA, and...
AAM, and the original elimination by aspects explanation of the similarity effect (Tversky, 1972). When the SSSA is combined with leaky integration of momentary values into preferences, both the DFT and the LCA, produce the similarity effect. In the LCA, for example, the SSSA counteracts the large disadvantages penalty factor that applies on the dissimilar alternative, \( B \) (in the set, \( A, S, B \)). Thus despite its overall weaker net value, the distant alternative, \( B \), is chosen with a probability of more than 1/3, because the instantaneous preference for similar options (\( A, S \)) are temporarily correlated rising and falling together. This temporal correlation reflects the chooser’s attention switching between the attributes. Similar options, therefore, tend to “split their wins” when attention is centered on attributes where both excel (see Figure 4). On the other hand, the dissimilar alternative \( B \) in Figure 4) gets a relative choice advantage because when its strong attribute is scanned, it tends to dominate alone. Because the MLBA mechanism does not rely on SSSA, it cannot use temporal correlations to explain the similarity effect; instead, the similarity effect in the MLBA account depends on advantages being weighted more heavily than disadvantages, as discussed above.

### Asymmetric Attribute Weights

In DFT and LCA, it is assumed that the two dimensions characterizing the three alternatives are equally important. In MLBA the weights do not represent the importance of each attribute but rather capture the attention allocated to a given (signed) pairwise comparison. A different approach was taken in the AAM model. There, the dimensional weights change in a bottom up fashion as a function of the choice problem. In particular, the weight given to any given dimension is proportional to the sum of the attribute values on that dimension. For instance, in an attraction effect scenario \( \{A, RF, B\} \), Figure 1), dimension \( Q \) will be more heavily weighted because it is associated with higher values overall \( Q_A + Q_{RF} + Q_B \). Using this mechanism, together with SSSA, AAM explains the three effects considered here as well as a range of other effects encountered in multiattribute choice. Note that the attribute weights can change, exogenously, in another recent model of multiattribute choice (Wollschläger & Diederich, 2012). Because this model does not explain all three effects using single parameter sets (but by allowing attribute weights to be different for each effect), we do not discuss it here.

Overall, we saw that in three models (LCA, DFT, and MLBA) preference formation relies on the estimation of the total dimension-wise advantages and disadvantages of a given option relative to the other available options. In LCA, AAM, and DFT, there is a sampling process (SSSA), whereas in MLBA the outcome of each pairwise comparison is weighted differentially depending (a) on its sign (advantage/disadvantage) and (b) on its magnitude. In all models temporal integration gives rise to preference states. Nonlinearities come into play in all models (other than AAM, see Footnote 1), at different processing stages. In LCA, disadvantages loom larger than advantages, in DFT similar options interact more via lateral inhibition, whereas in MLBA large and negative differences are less impactful than small and positive ones. A further nonlinearity is introduced in the MLBA model by assuming that the choice space is curved such that extreme alternatives worth less than more average ones. Finally, in AAM the weights assigned to the two dimensions change as a function of the choice problem at hand. Having described the main mechanisms of the MLBA model with respect to LCA, DFT, and AAM we turn next in evaluating its behavioral characteristics.

### Sensitivity to Stimulus Changes in the MLBA Model

The MLBA is able to explain the three contextual reversal effects as arising from a fine balance of opposing forces. On the one hand, allowing the weights to decay with distance increases the attraction effect because the close comparisons between the target and the decoy are magnified. On the other hand, this mechanism undermines the similarity effect because the distant alternative will receive less input. To counterbalance this, the MLBA assumes that advantages decay more slowly than disadvantages. This mechanism now boosts more “isolated” alternatives (that have large advantages and disadvantages), increasing the similarity effect but working against the attraction and compromise effects. Opposing mechanisms exist also in DFT and LCA; for example increasing the local inhibition in the former and the loss-aversion asymmetry in the latter increase the magnitude of the attraction/compromise effects at the expense of the similarity effect. Tension between mechanisms and effects raises the question of how sensitive the model is to changes of the values of the alternatives in the two-dimensional attribute space. As we have previously shown (Tsetsos et al., 2010) DFT and LCA are in a position to produce all three effects regardless of the relative location of the alternatives in the choice space, given fixed parameter sets. In this section, we evaluate value-sensitivity for the MLBA distant-advantage mechanism, which, as we show, drives the MLBA explanation for the similarity effect, but comes into tension with the compromise and the frequency-attraction effects (because in both of these cases the distant option(s) need to be chosen less).

#### The Attraction Effect

Let us consider first the frequency decoy (preferring \( A \) over \( B \) in the choice-set \( \{A, B, F\} \); Figure 1). Although a frequency decoy, \( F \), is less effective in triggering a preference bias for the target, \( A \), (Huber et al., 1982), significant frequency-attraction effects have been reported (Trueblood, 2012; Wedell, 1991; Wedell & Pettibone, 1996). Refer-
ring to Figure 1, we denote the distance between A and F with \( z \) and the advantage of B relative to F with \( d \), while the disadvantage of B to F with \( s = d - z \). In the case of binary choice A and B, being symmetric, have equal total values. We will here describe the extra value that is added to the two options via their comparison to the inferior F, which potentially breaks the tie between A and B in the ternary choice case. Following Equations 1–4, we see that F conveys an advantage to the proximal target, A, whose weighted magnitude, \( V_{AF}^- = z \cdot \exp(-\lambda_A \cdot d) \) while the corresponding weighted disadvantage (on dimension \( Q \)) will be \( \Delta V_{BF}^- = (d-z) \cdot \exp(-\lambda_B \cdot (d-z)) \). Thus the net input to alternative A will be \( V_{AF}^- = V_{AF}^+ \cdot [\exp(-\lambda_A \cdot d) \cdot \exp(-\lambda_B \cdot (d-z))] \) while the corresponding weighted disadvantage (on dimension \( Q \)) will be \( \Delta V_{BF}^- = \Delta V_{BF}^+ \cdot [\exp(-\lambda_B \cdot (d-z))] \). Since \( z < d \) it follows that \( V_{AF}^- < \Delta V_{BF}^- \). Because the advantages-weights (\( \lambda_A > \lambda_B \)), the competitor (B) can get a stronger overall input. For example, a large disadvantage that has fully dissipated (\( \Delta V_{BF}^- = 0 \)) can result in a null or negative attraction effect (where the preference for the target is reduced after the addition of an inferior decoy).

To illustrate this, for the parameter set reported in Trueblood et al. (2014) (parameter set 1; \( m = 5, \lambda_A = 0.2, \lambda_D = 0.3, l_0 = 5 \)), the frequency decay effect is negative for the scenario considered by the authors (Table 1, Effect-1). With a modified parameter set (parameter set 2; \( m = 3, \lambda_A = 0.2, \lambda_D = 0.3, l_0 = 5 \)) of our knowledge, has not been reported in the literature, and which, if verified, will provide considerable support for the MLBA.

The Compromise Effect

We next consider the compromise effect using again the successful parameter set reported in Trueblood et al. (2014) (\( m = 5, \lambda_A = 0.2, \lambda_D = 0.4, l_0 = 5 \)). Clearly, the compromise effect poses a special challenge to the distance-advantage mechanism, which should favor extreme alternatives that have larger advantages and disadvantages than less extreme options. The key ingredient that allows the MLBA to counteract the distance-advantage mechanism and capture the compromise effect in conjunction with the other two effects, is the convex subjective equal-preference curve, when \( m > 1 \) (Equation 2, Figure 2). It was possible to find MLBA parameters that account for the compromise with sets of equal subjective values (\( m = 1 \)). However, in these parameter sets advantages decay faster than disadvantages (\( \lambda_A > \lambda_D \)) and therefore the similarity effect is not predicted.\(^2\)

Crucially, for \( m > 1 \) the model predicts that the compromise effect arises because the extreme decoy option, A’, is subjectively inferior to the middle item. Using this mechanism, with a sufficiently large \( m \)-value (\( m = 5 \)), the model produces a compromise effect, as well as the attraction and similarity effects, for options on the set (A, B, A’). Alternative A’ under the subjective value transformation is inferior to A (transformed values are given as

\(^2\) The same holds for parameter sets with \( m < 1 \).
This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Overall least chosen (Table 2, effect-1: 32% for characteristically predicts that the added extreme decoy will be the compromise). On the contrary, the MLBA mechanism (844 TSETSOS, CHATER, AND USHER)

To illustrate this, we change the extreme option (Figure 2 for a schematic depiction), the compromise effect reverses dramatically.

The question of whether the compromise effect is contingent on the extreme decoy being inferior relative to the middle option is thus critical to assessing the MLBA mechanism. Although many demonstrations of the compromise effect used alternatives whose values are arranged in a linear configuration on the objective attribute space (such as $A$, $A'$, $B$; e.g., Pettibone, 2012), others have attempted to equalize the subjective values of the targets and decoys. For example, Tversky and Simonson (1993) asked participants to choose among 35 mm cameras varying in quality and price. “One group ($n = 106$) was given a choice between a Minolta X-370 priced at $170$ and a Minolta 3000i priced at $240$. A second group ($n = 115$) was given an additional option, the Minolta 7000i priced at $470$. Participants in the first group were split evenly between the two options, yet 57% of the subjects in the second group chose the middle option (Minolta 3000i), with the remaining divided about equally between the two extreme options.” (Tversky & Simonson, 1993: p. 1123). In this case, we do not have a metric for the distance between Minolta 370, 3000, and Minolta 7000. However, together with their prices the cameras appear to be equally attractive: A (Minolta X-370) and B (Minolta 3000i), were of equal subjective preference, since the first group was evenly split between them (choice = 50% ± 4.9%). A and C (Minolta 7000i) also appear to have been of equal preference, because the participants in the second group, who did not choose the compromise option, were evenly divided between them.

Other studies using target and extreme decoy in a linear configuration have shown that the compromise effect does not appear to depend on the extreme decoy being subjectively inferior. Pettibone (2012), for example, found that the choice of the extreme decoy and the competitor ($A'$ and $B$) have rather equal choice probabilities (both around, 30%, compared with more than 40% for the compromise). On the contrary, the MLBA mechanism characteristically predicts that the added extreme decoy will be the overall least chosen (Table 2, effect-1: 32% for $B$ and only 16% for $A'$). It is possible that some compromise studies used linear configurations in the objective attribute space with subjectively inferior extreme decoys and the MLBA mechanism could account for such cases. However, further experiments will be needed to determine the precise boundary conditions of the compromise effect and hence help distinguish different accounts of the cognitive mechanisms from which it arises.

Back-End Predictions

Although we mainly scrutinized the front-end MLBA component, it is also worthwhile to analyze the back-end component. This component is a race model—LBA (Brown & Heathcote, 2008)—with absolute decision termination. As recently discussed by Teodorescu and Usher (2013), race models have fundamentally different characteristics than competitive models (like diffusion, DFT, or LCA). In the MLBA model, this difference interacts with the context sensitivity leading to a set of very particular predictions. We report these predictions below using parameter set 1 ($m = 5$, $\lambda_A = 0.2$, $\lambda_B = 0.4$, $I_0 = 5$) but because $m = 5$ is not essential here, we examine the same parameter set with $m = 1$ (the results do not change qualitatively for different values of $m$). The back-end parameters are the same as in the original paper ($A = 1$, $\chi = 2, s = 1$).

Consider two alternatives, such as $A$ and $B$ in Figure 6A and let us introduce a dominated decoy ($C$ at $P = Q = 0.01$). The mere addition of this decoy increases the drift to both the $A$ and $B$ accumulators relative to the binary case (Equation it provided earlier, e.g., at $C$ (4,4) the drift-rates of $A$ and $B$ have nearly doubled relative to the binary case, Figure 6B). As the decision-termination is not competitive, the two parallel drift changes are not subtracted. As a result, the model predicts a dramatic speed–accuracy trade-off when $C$ is added: a reduction in accuracy (Figure 6C) and a large decrease in the RT (Figure 6D). As $C$ keeps moving beyond $y = 5$ and closer toward $A$ and $B$ a context effect comes in play: now the drift-rates of $A$ and $B$ decrease and their difference starts to increase. Thus a relative slow down is observed and accuracy starts increasing more steeply. Nevertheless, the predicted decision times are maintained way below those of the binary $A$-$B$ decision.

Thus the MLBA predicts that although $C$ is never chosen it influences the choice between $A$ and $B$ (Figure 6B). The predictions about mean choice are consistent with one recent study that showed accuracy improvement as the value of the decoy increases (Chau, Kolling, Hunt, Walton, & Rushworth, 2014) but are in the opposite direction to another study that showed the contrary (Louie, Khaw, & Glimcher, 2013). The most dramatic impact of

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Subjective Attribute Values and Compromise Effect for $M = 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Effect</strong></td>
<td><strong>Choice set</strong></td>
</tr>
<tr>
<td><strong>Effect-1</strong></td>
<td>$A = (4, 6)$</td>
</tr>
<tr>
<td>Transformed values</td>
<td>$U_A = (6.5, 9.8) = 16.3$</td>
</tr>
<tr>
<td>Choice</td>
<td>52%</td>
</tr>
<tr>
<td><strong>Effect-2</strong></td>
<td>$A = (4, 6)$</td>
</tr>
<tr>
<td>Transformed values</td>
<td>$U_A = (6.5, 9.8) = 16.3$</td>
</tr>
<tr>
<td>Choice</td>
<td>26%</td>
</tr>
</tbody>
</table>

Note. The values in italics correspond to the summed subjective (transformed) utility.
The back-end effects of varying an inferior alternative (C) on the choice between two alternatives, whose preference is not ambiguous (A > B), A: Option A is located at (8.8). Option B at (7.5, 7.5). Option C moves from (0.01,0.01) to (7.2, 7.2). B: Drift rates in the ternary scenarios (solid lines) as a function of the x,y coordinates of option C. Binary baseline drift rates for A and B are plotted with dashed lines. Blacked dotted line shows the drift rate difference between A and B as a function of C position. C: Probability of choice for the three options (solid lines) as C moves closer to A and B; dash lines indicate baseline choice in the binary case. D: Mean decision times as C moves closer to A and B (solid line). The dashed horizontal line shows the mean decision time in the binary scenario.

Discussion

The MLBA explains contextual preference reversal in multiattribute choice without the need for nonlinear value functions, providing a unified account of these phenomena in both hedonic and nonhedonic (perceptual) domains. The model is based on well-motivated principles and due to its reliance on the LBA mechanism it allows analytical solutions that facilitate data fitting. Although we believe these are important contributions, our exploration of the model behavior has uncovered a number of challenges, which we summarize below.

First, the MLBA account for the frequency-attraction effect works only for specific choice scenarios. For every parameter set that produces all three effects simultaneously, it is possible to find decoys for which the model predicts a null or negative attraction effect. Second, for parameter sets that explain all three effects, the account for the compromise effect is limited to situations in which the target is subjectively superior to the extreme decoy. Although extremeness-aversion may be a feature of multiattribute binary choice, one can create choice-sets that compensate for this (e.g., E, A, B, in Figure 2). Thus, a systematic parametric investigation of all three context effects (across the attribute space) will be important to distinguish between the predictions of MLBA and other choice models. Such experimental investigations should also examine decision times, which as shown in Figure 6, can be highly distinctive for MLBA (see the next section).

These two challenges for MLBA result from the fact that advantages are more heavily weighted than disadvantages (especially at large distances), an assumption necessary to explain the similarity effect. This mechanism is not present in the LCA, DFT, or AAM models, as they use a different mechanism—the stochastic sampling of attributes—to explain the similarity effect. The MLBA differential weighting of advantages and disadvantages is reminiscent of the LCA loss-aversion assumption, although it works in the opposite direction. The application of loss-aversion to perceptual domains was criticized by Trueblood et al. (2014), as in such nonhedonic domains, loss-aversion seems less natural. We agree, though we note that the magnitude of the preference reversal effects is significantly reduced in perceptual decision-making (Trueblood, Brown, Heathcote, & Busemeyer, 2013).

One important consideration in the evaluation of all multiattribute choice models is whether their computational complexity and demand is consistent with human decision-making capacities. This issue becomes more critical when the set-choice increases. For example, both the MLBA and the LCA make the assumption that for each pair of choice units, one needs two processing variables to compute differences for each option on each attribute (or four processing variables if we need to keep the positive and negative differences separate); with increasing choice-set the number of such units increases exponentially (Soltani et al., 2012). Because decisions often involve novel situations, all these dynamic units need to be created online and connected with the relevant inputs. In addition, nonlinear value functions, as well as distant dependent interactions, contribute to the model complexity.

In light of these computationally demanding assumptions, we highlight three recent computational approaches, proposed to account for contextual preference reversal and motivated by a reduction of task demands. The AAM model (Bhatia, 2013), mentioned above, assumes neither a comparison between alternatives (no exponential explosion for large-N) nor asymmetric value functions (or distance dependent interactions). Instead it accounts for the contextual reversal effects through the variable and adaptive weighting of the dimensions, with more heavily represented dimensions carrying higher weight. In turn, the highly represented dimensions are those that are more strongly associated with the alternatives. For example, introducing a decoy (D) to A, makes the dominant dimension of A more strongly represented, as it is associated with both A and D. Using this principle, the AAM was shown to explain the attraction, compromise, and similarity effect under single parameter sets as well as additional behavioral phenomena in multiattribute choice.

The second model is the range-normalization (RN) model (Soltani et al., 2012), inspired by neurophysiological constraints. The model was shown to account for the attraction and similarity effects, but not for the compromise effect. In the RN model, the contextual reversals are due to changes in the range of value representation as more alternatives are added to the choice set. The model relies on the notion of divisive normalization with the divisive factors being determined by the range of values on each...
dimension. For instance, the addition of a range-decoy (D) to A will not change the divisive normalization (compared to the binary A-B scenario) in the dimension where A is better than B. However, it will increase the divisive factor on the other dimension, where B has a much larger value than A. As a result, B will pay a larger penalty and its overall value will be reduced.

Third, the selective-integration model (Tsetsos, Chater, & Usher, 2012) replaces the explicit value-difference computation with a weighting mechanism that prioritises the integration of more salient samples. Following DFT and LCA, the main assumption is that one dimension is sampled at a time but each value-sample is distorted by a multiplicative weight that is proportional to its relative rank within the active dimension. This model does not rely on nonlinear value functions and was able to account for attraction and similarity effects in a simple value integration task with nonstationary dynamic evidence (i.e., temporally correlated following the stochastic scanning of dimensions principle) (Tsetsos, Chater, & Usher, 2012). One way in which this model can be developed is to replace the rank-dependent weighting with a capacity limited attentional process (e.g., only the first two of three values on a dimension are integrated), which may provide an alternative mechanism to explain the compromise effect. One attractive aspect of this approach is that it links decoy effects to well-described processes with known neural substrates such as selective attention.

Future Experimental Studies

Preference reversal in multiattribute choice has been traditionally examined with experimental materials, such as consumer products, which do not permit precise control of the decision-relevant variables and make within participants tests difficult. Until recently, the attraction, compromise, and similarity effects had not been reported in the same experiment or for the same participant. It was thus unclear whether a complete computational model of multiattribute choice had to capture all three effects under a single parameter set. However, recent experiments in Trueblood (2012) and Berkowitsch, Scheibehenne, and Rieskamp (2013) have showed that the three effects can be obtained using the same paradigm and within single participants, highlighting the need for a unified computational account. We believe that theoretical attempts could benefit by targeted experiments elucidating the mechanisms underlying multiattribute decisions. Below we briefly outline such experimental directions.

Parametric manipulation of the decoys in the attribute space. Computational models of preference reversal make different predictions about the magnitude of the effects as a function of the position of the added alternative in the choice space. For example, as shown in Figure 5 of Trueblood et al. (2014), for the attraction effect the MLBA predicts a nonmonotonic distance-dependent influence of the decoy on the probability of choosing the target. Experimental evidence about the distance-dependency of the attraction effect is currently inconclusive with Soltani et al. (2012) showing that distant decoys increase the magnitude of the effect but with Wedell (1991) showing the opposite pattern. A systematic parametric investigation (equivalent to Figure 7 in Trueblood et al., 2014) is likely to provide important constraints on theories of contextual preference reversal. As discussed above, future experimental studies should also examine decision times, which may provide highly diagnostic evidence for model comparison.

Individual differences and correlation of effects within subjects. Most models (LCA, DFT, and MLBA as well as the selective integration model) use a number of functional components to produce all three reversal effects simultaneously. The difficulty to capture all three effects under a single account might coincide with a difficulty to empirically obtain the effects within the same participant. In support to this possibility, Berkowitsch et al. (2013) obtained all three effects in 19% of their participants (corresponding to nine individuals). Thus, one important objective for models of multiattribute choice is to also account for individual differences in the form of correlations in the magnitude and direction of the effects. More generally, although the focus so far has been in obtaining all three effects under a single parameter set, future work might usefully be directed toward capturing the variability in human multiattribute choice behavior under a single theoretical framework (e.g., allowing different parameters to explain different strategies).

Presentation mode (simultaneous/sequential) and preference reversal. One process that is likely to affect the degree of contextual reversal is the amount of within-attribute versus within-alternatives processing. As the DFT, LCA, MLBA, and the selective-integration model show, within-attribute comparisons are necessary to trigger reversal effects. Different stimuli format and presentation (sequential vs. simultaneous) can encourage the employment of one strategy over the other. For instance, in Tsetsos et al. (2011), we obtained a strong similarity effect presumably because the presentation of the stimulus precluded holistic (compensatory) processing. Alternatively, presenting each option separately (e.g., different pages or computer frames) could encourage holistic compensatory processes. Detailed manipulation of the presentation mode might thus shed light on whether decoy effects are dependent on the way information is processed.

Conclusion

Recent research has made progress in the characterization of the choice mechanism that underlies contextual reversal effects in multiattribute choice. Most models assume that these effects are driven by computationally demanding processes of pairwise comparisons among the alternatives, where evidence is accumulated gradually before a decision is made. These models differ concerning the mechanisms of evidence accumulation and choice. We have claimed that new experimental studies should help distinguish between rival accounts.

References


TSETSOS, CHATER, AND USHER