A Computational Model for Constructing Preferences for Multiple Choice Options

Lena M. Wollschlaeger (l.wollschlaeger@jacobs-university.de)
Jacobs University Bremen, Psychology & Methods
Campus Ring 1, 28759 Bremen, Germany

Adele Diederich (a.diederich@jacobs-university.de)
Jacobs University Bremen, Life Sciences & Chemistry
Campus Ring 1, 28759 Bremen, Germany

Abstract

When choosing between multiple alternatives, people usually do not have ready-made preferences in their mind but rather construct them on the go. The 2N-ary Choice Tree Model (Wollschlaeger & Diederich, 2012) proposes a preference construction process for N choice options from description, which is based on attribute weights, differences between attribute values, and noise. It is able to produce similarity, attraction, and compromise effects, which have become a benchmark for multi-alternative choice models, but also several other context and reference point effects. Here, we present a new and mathematically tractable version of the model – the Simple Choice Tree Model – which also explains the above mentioned effects and additionally accounts for the positive correlation between the attraction and compromise effect, and the negative correlation between these two and the similarity effect as observed by Berkowitsch, Scheibehenne, and Rieskamp (2014).

Keywords: computational model; multi-alternative choice; choice from description; preference construction; context effects

Introduction

The decision making process involves various steps such as setting and prioritizing objectives, identifying choice alternatives, searching for information, developing preferences, and eventually taking a course of action. Here, we focus on developing preferences in multi-alternative choice situations and use in the following decision making from description as basic paradigm. Given a set of at least three choice alternatives that are described by at least two attributes, which they have in common, how do people choose one of these options? Simon (1955) argues that preferences in this kind of situation are dynamically constructed over time due to limited processing capacities. The decision maker experiences preference uncertainty (cf. Simonson, 1989) and tries to overcome it by gradually integrating the given information (see Payne, Bettman, & Johnson, 1992, for a review on constructive processing in decision making). The resulting preferences are stochastic and highly dependent on the context, i.e., on the alternatives in the choice set and on any external reference points. Naturally, a model describing multi-alternative decision making from description should be a context-sensitive cognitive process model. The recently proposed 2N-ary Choice Tree Model for preference construction for N choice options (2NCT; Wollschlaeger & Diederich, 2012) assumes that the decision maker compares attribute values within attributes and between alternatives in a pairwise manner. Attributes are selected for examination based on attribute weights that reflect salience. Within attributes, pairs of attribute values are selected for comparison based on so-called comparison values. In the 2NCT Model, the comparison values have a "global" component that remains constant over time during preference construction, a "local" component that depends on the outcomes of previous comparisons (reflecting leakage and inhibition, cf. Roe, Busemeyer, & Townsend, 2001; Usher & McClelland, 2004), and a random component. Advantageous and disadvantageous comparison outcomes for each alternative are counted separately and the difference of these counters is compared to two thresholds: a positive choice criterion and a negative elimination criterion. Implementation of an asymmetric value function (emphasizing disadvantageous comparison outcomes, cf. Usher & McClelland, 2004) into the 2NCT Model is possible. Here, we present a revised and simpler version of the 2N-ary Choice Tree Model, the Simple Choice Tree (SCT) Model. Therein, the local component is omitted from the definition of comparison values, making the model mathematically tractable while maintaining its ability to account for similarity, attraction and compromise effects. Furthermore, a new parameter, the focus weight λ, is introduced. It replaces the asymmetric value function and allows the SCT Model to account for correlations between the effects.

Benchmark: Context Effects

Three context effects, demonstrating the influence of choice set composition on preferences, have played a prominent role in the multi-alternative preference construction modeling literature: The similarity effect, the compromise effect, and the attraction effect. All three effects occur when adding a third alternative to a set of two equally attractive yet clearly distinguishable options described by two attributes. Let A_1 and A_2 be two choice alternatives with two common attributes, D_1 and D_2, describing them. We assume that D_1 is the unique strongest attribute for A_1 and D_2 is the unique strongest attribute for A_2, that is, A_1 scores high on D_1 but low on D_2 and vice versa for A_2. One can think of the alternatives as placed in a two-dimensional space with dimensions D_1 and D_2. We further assume that the probability for choosing alternative A_1 from the binary choice set is equal to the probability for choosing alternative A_2, $P(A_1|A_1,A_2) = P(A_2|A_1,A_2)$.

Similarity Effect The similarity effect was named and first studied systematically by Tversky (1972). He observed the effect when comparing the binary choice set $\{A_1,A_2\}$ to the ternary choice set $\{A_1,A_2,A_3\}$ where A_3 is similar to one of the original alternatives, say A_1, in scoring high on attribute D_1 and low on attribute D_2 while overall being similarly attractive (i.e. $P(A_1|A_1,A_3) = P(A_3|A_1,A_3)$). The probability of choosing A_1 over A_2 decreases when the decision maker chooses from the ternary choice set as compared to the binary set: $P(A_1|A_1,A_2)/P(A_2|A_1,A_2) > P(A_1|A_1,A_2,A_3)/P(A_2|A_1,A_2,A_3)$.

Attraction Effect The attraction effect (or decoy effect or asymmetric dominance effect) was introduced by Huber, Payne, and Puto (1982) as consistent violation of the regularity principle. This principle, as presumed for example by the theory of Elimination by Aspects (Tversky, 1972), states that additional alternatives cannot increase the choice probabilities of the original options. However, Huber et al. (1982) claim that the relative probability for choosing alternative, say, A_1 can be increased by adding a third alternative A_3 to the choice set that is similar to but dominated by A_1 (and symmetrically for alternative A_2). A_3 then serves as a decoy for alternative A_1, drawing attention to it and therewith improving its evaluation and increasing its choice probability.

Compromise Effect Originally intended to explain the attraction effect, the theory of Reason-based Choice (Simonson, 1989) predicts an additional context effect, the compromise effect. It occurs when a third alternative A_3, equally attractive as the original alternatives A_1 and A_2, but more extreme with respect to the attribute values, is added to the choice set. If A_3 is more extreme than alternative A_1, that is, if it scores higher than A_1 on attribute D_1 but lower on attribute D_2, then it increases the choice share of A_1 as compared to the binary situation (and vice versa for alternative A_2): $P(A_1|A_1,A_2,A_3)/P(A_2|A_1,A_2,A_3) > P(A_1|A_1,A_2)/P(A_2|A_1,A_2)$. However, note that the more similar the additional extreme alternative A_3 is to its adjacent alternative A_1, the more shares it takes away from A_1 via the similarity effect.

Interrelations of the Effects Recently, several studies have explored similarity, attraction and compromise effects and their interrelations in different choice scenarios. In a within-subject consumer choice design, Berkowitsch et al. (2014) find that the similarity effect is negatively correlated with both the attraction and the compromise effect while the latter two are positively correlated. In a similar vein, Liew, Howe, and Little (2016) criticize that most of the results regarding context effects are based on averages over participants, not taking into account individual differences. Before analyzing the data from their inference and consumer choice experiments, they cluster it according to the observed choice patterns. The differences between clusters are remarkable, some even show negative (reverse) context effects while positive effects are observed in the averaged data. Before explaining how the Simple Choice Tree (SCT) Model accounts for the similarity, attraction and compromise effects and their interrelations, we introduce the basic mechanisms of the model.

The Simple Choice Tree Model

Let n_a be the number of alternatives under consideration, $\{A_1, A_2, \ldots, A_{n_a}\}$, and n_d the number of attributes, $\{D_1, \ldots, D_{n_d}\}$, that characterize them. The decision maker is provided with one attribute value per alternative per attribute, that is, $n_a \cdot n_d$ attribute values in total. Let m_{ij} be the attribute value for alternative A_i with respect to attribute D_j. Attribute values within attributes and between alternatives are repeatedly compared and the resulting evidence is accumulated in two counters S_i^+ and S_i^- for each alternative $A_i, i \in \{1, \ldots, n_a\}$. The positive counter S_i^+ accumulates evidence for choosing alternative A_i and the negative counter S_i^- accumulates evidence for rejecting it. Here, the initial counter states are set to zero, $S_i^+(0) = 0 = S_i^-(0)$. Definition of non-zero initial counter states accounting for prior knowledge about the choice alternatives is possible. However, these additional free parameters make the model less parsimonious and complicate parameter estimation. The counter states at time t, $S_i^+(t)$ and $S_i^-(t)$, are the initial counter states increased by the respective evidence accumulated until t. Their difference defines the momentary preference state for alternative A_i at time t: $Pref(A_i,t) = S_i^+(t) - S_i^-(t)$. We will now answer the following questions: (1) How is attention allocated between choice alternatives and attribute values? (2) How are alternatives evaluated and how is evidence accumulated? (3) When does evidence accumulation stop and which alternative is chosen?

Attention Allocation

At the beginning of the process, when information about the alternatives and attributes is made available to the decision maker, each attribute D_j, $j \in \{1, \ldots, n_d\}$, is assigned a weight ω_j, $0 \leq \omega_j \leq 1$, reflecting its salience. The attribute weights determine how much attention the decision maker gives to the respective attributes during the preference construction process. Attributes with higher weights get more attention than attributes with lower weights. To allow for at least some of the attention to be allocated randomly between attributes, we define a random component (see below) for which an additional weight $\omega_0, 0 \leq \omega_0 \leq 1$ is designated. Assuming that the weights sum up to one, $\sum_{j=0}^{n_d} \omega_j = 1$, they can be interpreted as attention probabilities for the attributes: At each points of the preference construction process, the decision maker concentrates on attribute $D_j, j \in \{1, \ldots, n_d\}$ with probability ω_j.

Having selected an attribute D_j, the decision maker concentrates on the specific attribute values of two alternatives and compares them. Pairs of attribute values are selected for comparison according to their importance for the decision. The more diagnostic the attribute values are, i.e., the more they discriminate between the alternatives, the more important they become for the decision. Pair selection probabilities within attribute D_j are therefore defined to be proportional to the absolute differences $d_{ikj} = |m_{ij} - m_{kj}|, i \neq k \in \{1, \ldots, n_a\}$. In order to obtain probabilities, we normalize these differences to sum up to one: The probability for selecting the pair

$\{m_{ij}, m_{kj}\}$ for comparison is $p_{ikj} = d_{ikj} / \sum_{\{l,m\}} d_{lmj}$, $l \neq m \in \{1, \dots, n_a\}$. Note that the normalization of absolute differences balances out inequalities between attributes with – on average – bigger or smaller differences. Higher salience of an attribute D_j, $j \in \{1, \dots, n_d\}$, with, for example, higher absolute differences, is thus not hard-wired into the model but is reflected in a higher attribute weight ω_j instead.

Preference Sampling

The actual comparison of the two selected attribute values m_{ij} and m_{kj} is ordinal and directional: Let $m_{ij} > m_{kj}$, then the comparison can be either positively phrased, e.g. "m_{ij} is greater than m_{kj}", or it can be negatively phrased, e.g. "m_{kj} is smaller than m_{ij}". For the positive phrasing, m_{ij} is called *focus value* and m_{kj} is called *reference value*. The focus value determines the counter whose state is increased by $+1$, here S_i^+, since the comparison is advantageous for the associated alternative A_i. For the negative phrasing, m_{kj} is the focus value and m_{ij} is the reference value, leading to an increase by $+1$ of counter S_k^-, since the comparison is disadvantageous for alternative A_k. Which phrasing the decision maker uses for the comparison and therewith which counter is updated might, for example, depend on the wording of the task or the decision maker's attitude (cf. Choplin & Hummel, 2002). It is implemented into the model via the focus weight λ, $0 \leq \lambda \leq 1$. If $\lambda = 1 - \lambda = 0.5$, the decision maker uses the positive and negative phrasing both about equally often. If $\lambda > 0.5$, the decision maker has a tendency towards the negative phrasing and towards updating negative counters. If $\lambda < 0.5$, the decision maker has a tendency towards the positive phrasing and towards updating positive counters. The focus weight λ replaces the asymmetric value function that was applied to the absolute differences between attribute values in the original 2NCT Model (Wollschlaeger & Diederich, 2012). While the asymmetric value function hard-wired a tendency towards updating negative counters into the 2NCT Model, weighting with λ allows for flexible balancing of attention to positive versus negative aspects of the alternatives in the SCT Model. It is therefore especially useful in situations without a loss/gain-framing, e.g., in perceptual or preferential choice. Note that λ is a global weight and independent from the attributes and attribute values. However, it allows us to define counter updating probabilities for the positive and negative counter of alternative $A_i, i \in \{1, \dots, n_a\}$ with respect to attribute $D_j, j \in \{1, \dots, n_d\}$: $p_{ij}^+ = \sum_{k:(m_{ij} > m_{kj})} (1 - \lambda) \cdot p_{ikj}$ for updating S_i^+ and $p_{ij}^- = \sum_{k:(m_{ij} < m_{kj})} \lambda \cdot p_{ikj}$ for updating S_i^-.

Finally, the random component accounts for times where counter states are updated at random and without any connection to the actual attribute values (for instance due to inattention or misperception, cf. Busemeyer & Townsend, 1993). Technically, it is treated as an additional (phantom) attribute D_0. The counter updating probabilities $p_{i0}^+ = p_{i0}^- = 1/(2 \cdot n_a)$, $i \in \{1, \dots, n_a\}$ with respect to D_0 depend on the number of available choice alternatives and therefore sum up to one: $\sum_{i=1}^{n_a} (p_{i0}^+ + p_{i0}^-) = 1$.

Combining attribute-wise counter updating probabilities $p_{ij}^{\pm}$ with attribute weights ω_j, we can now define weighted counter updating probabilities for the positive and negative counter of alternative A_i:

$$p_i^+ = \sum_{j=0}^{n_d} p_{ij}^+ \cdot w_j \quad \text{and} \quad p_i^- = \sum_{j=0}^{n_d} p_{ij}^- \cdot w_j. \qquad (1)$$

Choice Tree and Stopping Rules

Starting with the presentation of the choice alternatives and their attribute values, the preference construction process consists of a sequence of counter updates. In principle, every possible sequence of counter updates may occur and it is therefore of interest to have them conveniently summarized. For this purpose, we introduce the $(2 \cdot n_a)$-ary choice tree $T = (V, E, r)$ with vertices V, edges $E \subseteq V \times V$ and root $r \in V$, where all vertices are directed away from r and each internal vertex $v \in V$ has $2 \cdot n_a$ children that are associated with the $2 \cdot n_a$ counters. Figure 1 shows an example with three choice alternatives and six counters. The preference construction process is represented by a random walk on T, beginning at the root and passing from there through an edge to another vertex, triggering the update (increase by $+1$) of the associated counter, moving on through another edge and so forth. The next edge to pass through is chosen according to the updating probability of the counter associated with its endpoint. Note that for each vertex the transition probabilities of all outgoing edges sum up to one. An example path of this random walk is pictured in bold in Figure 1.

The preference construction process stops when enough evidence has been accumulated to make the required choice. To this end, the preference states $Pref(A_i, t) = S_i^+(t) - S_i^-(t), i \in \{1, \dots, n_a\}$ are constantly compared to two thresholds, a positive threshold θ^+ and a negative threshold $\theta^- = -\theta^+$. If the preference state for alternative A_i hits the positive threshold, the process stops and A_i is chosen. If, on the other hand, the preference state for alternative A_k hits the negative threshold, A_k is eliminated from the choice set and the process continues with the remaining alternatives until one of them is chosen or until all but one of them have been eliminated. Consider a simple example with three choice alternatives $\{A_1, A_2, A_3\}$ and thresholds $\theta^+ = 2$ and $\theta^- = -2$. The sample path in Figure 1 with its associated sequence of counter updates $S_2^+, S_1^-, S_1^-, S_2^+$, leads to elimination of alternative A_1 after three steps and choice of alternative A_2 after four steps. Other possible sequences resulting in choice of alternative A_2 include $S_3^+, S_1^-, S_2^+, S_2^+$ with direct choice of A_2 after four steps, and $S_1^-, S_3^-, S_3^-, S_1^-$ with elimination of alternatives A_3 after three steps and A_1 after four steps and therewith choice of the only remaining alternative A_2.

Choice Probabilities and Expected Response Times

The probability for walking along a specific path as, for example, shown in Figure 1, is the product of the transition probabilities along the respective edges. The choice probability for alternative $A_i, i \in \{1, \dots, n_a\}$ is equal to the sum

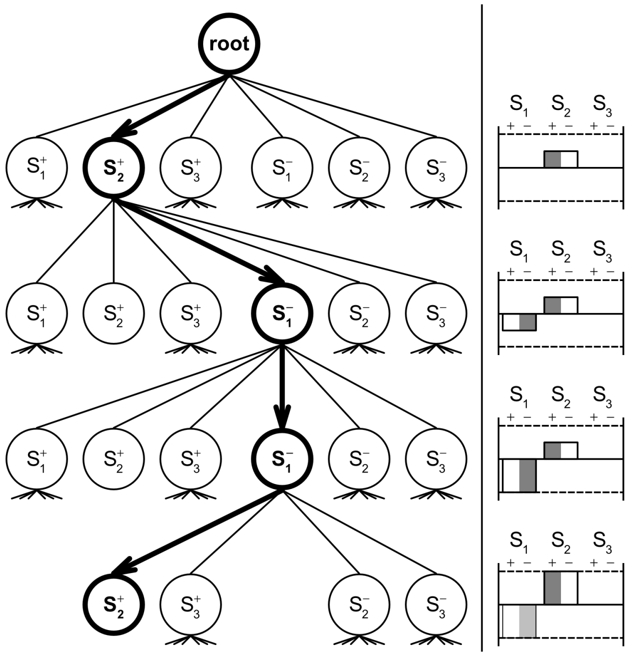

Figure 1: A random walk on the choice tree for three alternatives. The associated sequence of counter updates is $S_2^+, S_1^-, S_1^-, S_2^+$ and the probability for walking along this specific path is $p_2^+ \cdot p_1^- \cdot p_1^- \cdot p_2^+$. Supposing that the rejection threshold θ^- is equal to -2 and the choice threshold θ^+ is equal to 2, this sequence implicates first rejection of alternative A_1 and then choice of alternative A_2. When A_1 is eliminated from the choice set, the vertices associated with its counters no longer appear in the choice tree, as can be seen in the bottom row of vertices here.

of the probabilities for walking along all the specific paths that lead to choice of alternative A_i. Since it is not feasible to calculate probabilities separately for each path and sum them up, we will analyze preference states, choice probabilities and response times instead by interpreting them as independent birth-death Markov chains with absorbing boundaries θ^+ and θ^-. The state space of these birth-death chains $Pref(A_i,t) = S_i^+(t) - S_i^-(t) =: S_i(t), i \in \{1,\ldots,n_a\}$ is $\mathcal{S} := \{\theta^-,\ldots,-1,0,1,\ldots,\theta^+\}$, with $|\mathcal{S}| = \theta^+ - \theta^- + 1$. The transition probabilities are

$$\left.\begin{aligned} p_i(x,x+1) &= p_i^+ > 0 \\ p_i(x,x-1) &= p_i^- > 0 \\ p_i(x,x) = 1 - p_i^+ - p_i^- &= p_i^0 > 0 \end{aligned}\right\} \text{ for } x \in \mathcal{S} - \{-\theta^-, \theta^+\},$$

where $p_i^{\pm}$ is defined in Eq. 1 above; $p_i(x,x+1) = p_i(x,x-1) = 0$, $p_i(x,x) = 1$, for $x \in \{-\theta^-, \theta^+\}$; and zero otherwise. They form a $|\mathcal{S}| \times |\mathcal{S}|$ transition probability matrix $P_i' = (p_{rs}')_{r,s=1,\ldots,|\mathcal{S}|}$, where p_{rs}' is the probability for the birth-death chain to transition from state x_r to state x_s in one step. P_i' can be written in its canonical form P_i by rearranging the rows and columns (changing the indices of the states such that the absorbing states $-\theta^-$ and θ come first). P_i can be decomposed into a 2×2 identity matrix I_2, a $2 \times n_t$ matrix 0 of zeros with $n_t = |\mathcal{S}| - 2$ (the number of transient states in $\mathcal{S}$), a $n_t \times 2$ matrix R_i, containing the probabilities for entering the absorbing states θ^+ and θ^-, that is, for hitting the elimination or choice threshold, and a $n_t \times n_t$ matrix Q_i, containing the transition probabilities between transient states (cf. Diederich, 1997): $P_i = \begin{pmatrix} I_2 & 0 \\ R_i & Q_i \end{pmatrix}$.

Given a row vector Z_i of length n_t which represents the initial preference state (e.g., $\begin{pmatrix} 0 & 0 & 1 & 0 & 0 \end{pmatrix}$) or the initial distribution of preference over the transient states (e.g., $\begin{pmatrix} 0.05 & 0.10 & 0.70 & 0.10 & 0.05 \end{pmatrix}$, cf. Diederich & Busemeyer, 2003) for alternative A_i, the probability that the process is absorbed during the first step can be obtained by multiplying Z_i and R_i, yielding a vector of length 2: $Z_i \cdot R_i = [P(S_i(1) = \theta^+), P(S_i(1) = -\theta^-)]$. In the case that the process was not absorbed during the first step, the distribution of preference over the transient states after the first step is given by $Z_i \cdot Q_i$, a vector of length n_t. Multiplying the result with the matrix R_i yields the probabilities of absorption in the second step: $Z_i \cdot Q_i \cdot R_i = [P(S_i(2) = \theta^+), P(S_i(2) = -\theta^-)]$. The distribution of preference over the transient states is given by $(Z_i \cdot Q_i) \cdot Q_i = Z_i \cdot (Q_i \cdot Q_i) = Z_i \cdot (Q_i)^2$. The entries of the $n_t \times n_t$ matrix $(Q_i)^2$ are 2-step transition probabilities between the transient states, allowing for calculation of absorption in the third step: $Z_i \cdot (Q_i)^2 \cdot R_i = [P(S_i(3) = \theta^+), P(S_i(3) = -\theta^-)]$. Iterating these results indicates that all the relevant probabilities can be obtained from the vector Z_i, the matrix R_i and powers of the matrix Q_i. Since Q_i is a tridiagonal Toeplitz matrix (the entries on the main diagonal are all equal to p_i^0, the entries on the diagonal above the main diagonal are equal to p_i^+ and the entries on the diagonal below the main diagonal are equal to p_i^-), its eigenvalues, eigenvectors and its powers are known and given in closed form (Salkuyeh, 2006), making it easy to compute all the relevant quantities.

We are interested in the conditional probabilities and expected hitting times for each alternative A_i, $i \in \{1,\ldots,n_a\}$, given that A_i is the first alternative to be chosen/eliminated. Therefore, we have to determine the probability that alternative A_k, $k \in \{1,\ldots,n_a\}$ with $k \neq i$, has not been chosen/eliminated until time t. It is given by

$$\begin{aligned} P(-\theta^- < S_k(T) < \theta) &= 1 - \sum_{t=1}^{T} Z_k \cdot (Q_k)^{t-1} \cdot R_k \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= 1 - Z_k \cdot \left(\sum_{t=1}^{T} (Q_k)^{t-1} \right) \cdot R_k \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \end{aligned}$$

The choice and elimination probability for alternative A_i at time T is then equal to

$$\begin{aligned} &[P(S_i(T) = -\theta^-), P(S_i(T) = \theta)] \\ &= \left(Z_i \cdot \sum_{t=1}^{T} (Q_i)^{t-1} \cdot R_i \right) \cdot \prod_{k \neq i} \left(P(-\theta^- < S_k(T) < \theta) \right). \end{aligned}$$

Overall, this yields probabilities

$$[P(chooseA_i), P(eliminateA_i)] = \sum_{T=1}^{\infty} \left([P(S_i(T) = -\theta^-), P(S_i(T) = \theta)] \right)$$

and expected response times

$$[E(T_i|chooseA_i), E(T_i|eliminateA_i)] = \sum_{T=1}^{\infty} T \cdot \left([P(S_i(T) = -\theta^-), P(S_i(T) = \theta)] \right).$$

Note that the infinite sums over T have only a finite number of nonzero addends, since $P(N_i < \infty) = 1$ for all $i \in \{1, \ldots, n_a\}$, thus the choice/elimination probabilities and expected response times can be easily computed.

Context Effects Explained

Three interacting mechanisms produce similarity, attraction, and compromise effects in the Simple Choice Tree Model: (1) selection of pairs of attribute values for comparison based on normalized differences, (2) the possibility to eliminate unwanted alternatives from the choice set, and (3) weighting of attributes based on salience. The first mechanism leads to a higher impact of dissimilar alternatives on the updating probabilities and thus faster evidence accumulation for alternatives with more distant competitors. In the similarity and attraction settings, this applies to the dissimilar alternative A_2, and in the compromise situation to the extreme alternatives A_2 and A_3. The second mechanism and the related focus weight λ determine whether choices are more likely to be based on eliminations or to be made directly. The greater λ, the more likely are the choices based on eliminations. In the similarity situation, greater λ leads to faster elimination of the dissimilar alternative A_2 and subsequent choice or elimination of either alternative A_1 or A_3, that is, a small or even negative similarity effect. On the other hand, smaller λ leads to more direct choices of alternative A_2 and thus a higher similarity effect. Regarding the dissimilar alternative A_2, the same is true in the attraction situation. Greater λ leads to faster elimination of A_2 while smaller λ leads to more direct choices of alternative A_2. However, the attraction effect is higher for greater λ, since after elimination of alternative A_2, either the dominating option A_1 is chosen directly or the dominated option A_3 is eliminated first. In the compromise setting, greater λ increases the probability for the extreme options to be eliminated from the choice set, leaving the decision maker with the compromise option. Smaller λ on the other hand more likely leads to choice of an extreme option and thus a smaller or even negative compromise effect. Attribute weights further moderate the strengths of the context effects, but as long as they are more or less balanced, they play a minor role in the explanation of the similarity, attraction, and compromise effects. However, a high attribute weight is able to bias choice towards the alternative that scores highest on that attribute, covering any context effect.

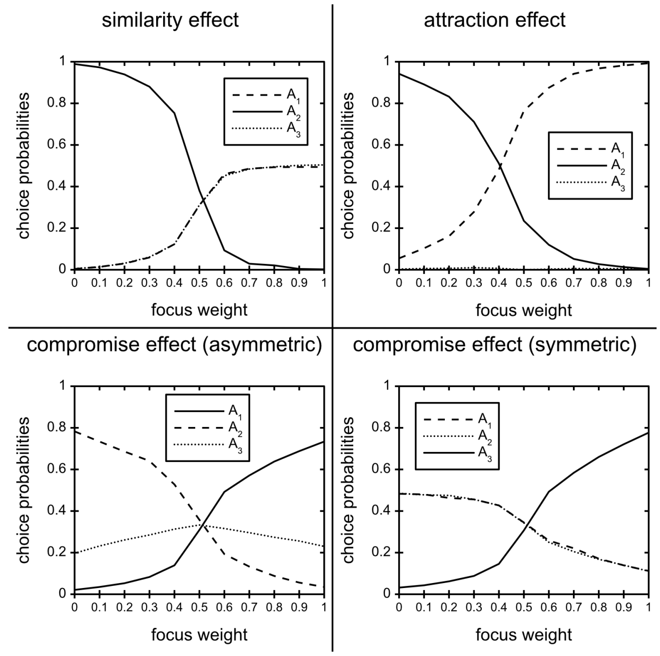

Figure 2: Simulations of choice probabilities for changing focus weight λ in the similarity, attraction, and compromise situation. There is a positive similarity effect for smaller λ and a negative similarity effect for larger λ (upper left) and vice versa for the attraction effect (upper right). The compromise effect (lower left and right) shows for larger λ and is reversed for smaller λ.

We ran several simulations to illustrate these mechanisms. The available choice alternatives were $A_1 = (70, 30)$, $A_2 = (30, 70)$ and $A_3 = (70, 30)$ for the similarity effect, $A_3 = (65, 25)$ for the attraction effect, $A_3 = (90, 10)$ for the asymmetric compromise effect, or $A_3 = (50, 50)$ for the symmetric compromise effect. The attribute weights were $\omega_0 = 0.1$ and $\omega_1 = \omega_2 = 0.45$, and the focus weight λ varied between 0 and 1 in steps of 0.1. For each data point we ran 10000 simulations and the resulting choice probabilities are presented in figure 2. According to the simulations, the similarity effect is opposed to the attraction and the compromise effect. The similarity effect is strongest for low λ, whereas the attraction and the compromise effect are strongest for high λ. This prediction is consistent with the finding that the attraction and the compromise effect are positively correlated with each other and negatively correlated with the similarity effect (Berkowitsch et al., 2014). Note that λ is assumed to be a global weight that does not change between trials but may vary between participants.

Conclusion

We propose a revised and simpler version of the $2N$-ary Choice Tree Model (Wollschlaeger & Diederich, 2012), the Simple Choice Tree (SCT) Model. It predicts choice probabilities and response times in multi-alternative multi-attribute preferential choice from description and accounts for several effects observed in these situations, including the similarity, attraction, and compromise effect. The SCT Model shares several aspects with existing models: Like Decision by Sampling (DbS; Stewart, Chater, & Brown, 2006), it proposes binary ordinal comparisons and frequency accumulation as basic mechanisms. In DbS, however, pairs of attribute values are chosen at random and reference values may be sampled from long-term memory as well as from the given context. Only advantageous comparisons are counted and the model is not able to account for the above mentioned context effects, nor does it provide solutions for choice probabilities or choice response times. Multi-alternative Decision Field Theory (MDFT; Roe et al., 2001) and the Leaky Competing Accumulator (LCA) Model (Usher & McClelland, 2001, 2004) provide such solutions only for fixed stopping times. Both models, like the SCT Model, are based on pairwise differences of attribute values. To account for the similarity, attraction, and compromise effect simultaneously, however, additional non-linear mechanisms (among others leakage and inhibition, cf. the original 2NCT Model) are required, preventing the models from providing mathematically tractable solutions for optional stopping times. Elimination by Aspects (EBA; Tversky, 1972) proposes "a covert elimination process based on sequential selection of aspects" (p. 296). As an early example for a cognitive process model, it does not make any predictions about choice response times and accounts only for the similarity effect. The SCT model mimics EBA for high values of the focus weight λ, where mostly disadvantageous comparison outcomes are considered and decisions are based on the elimination of choice alternatives. The Multiattribute Linear Ballistic Accumulator Model (MLBA; Trueblood, Brown, & Heathcote, 2014), basically a deterministic version of MDFT, provides analytic solutions for expected response times and choice probabilities like the SCT Model. However, it is unclear if and how the response times are related to the actual integration of information. Furthermore, the model has mostly been applied with fixed stopping times until now. Additional mechanisms allow the MLBA model to account for the compromise effect (a curved subjective value function) and the similarity effect (a higher weight on supportive information as compared to disconfirmatory evidence). The latter is comparable to low values of the focus weight λ in the SCT Model. To summarize, the SCT Model combines aspects of competing models in a new way, yielding qualitatively new explanations for the context effects and additionally predicting correlation patterns amongst the effects. It provides mathematically tractable solutions for both choice probabilities and expected choice response times for optional stopping times, by that outperforming existing models.

Acknowledgement

This research was supported by grant DFG DI506/16-1.

References

Berkowitsch, N. A. J., Scheibehenne, B., & Rieskamp, J. (2014). Rigorously testing multialternative decision field theory against random utility models. *Journal of Experimental Psychology: General*, *143*(3), 1331–1348.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*(3), 432–459.

Choplin, J. M., & Hummel, J. E. (2002). Magnitude comparisons distort mental representations of magnitude. *Journal of Experimental Psychology: General*, *131*(2), 270–286.

Diederich, A. (1997). Dynamic stochastic models for decision making under time constraints. *Journal of Mathematical Psychology*, *41*(3), 260–274.

Diederich, A., & Busemeyer, J. R. (2003). Simple matrix methods for analyzing diffusion models of choice probability, choice response time, and simple response time. *Journal of Mathematical Psychology*, *47*(3), 304–322.

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

Liew, S. X., Howe, P. D. L., & Little, D. R. (2016). The appropriacy of averaging in the study of context effects. *Psychonomic Bulletin & Review*, *23*(5), 1639–1646.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1992). Behavioral Decision Research - a Constructive Processing Perspective. *Annual Review of Psychology*, *43*, 87–131.

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionst model of decision making. *Psychological Review*, *108*(2), 370–392.

Salkuyeh, D. K. (2006). Positive integer powers of the tridiagonal Toeplitz matrices. *International Mathematical Forum*, *1*(22), 1061–1065.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*(1), 99–118.

Simonson, I. (1989). Choice Based on Reasons: The Case of Attraction and Compromise Effects. *Journal of Consumer Research*, *16*(2), 158–174.

Stewart, N., Chater, N., & Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology*, *53*(1), 1–26.

Trueblood, J. S., Brown, S. D., & Heathcote, A. (2014). The multiattribute linear ballistic accumulator model of context effects in multialternative choice. *Psychological Review*, *121*(2), 179–205.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, *79*(4), 281–299.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*(3), 550–592.

Usher, M., & McClelland, J. L. (2004). Loss aversion and inhibition in dynamical models of multialternative choice. *Psychological Review*, *111*(3), 757–769.

Wollschlaeger, L. M., & Diederich, A. (2012). The 2N-ary choice tree model for N-alternative preferential choice. *Frontiers in psychology*, *3*(189).