British Journal of Mathematical and Statistical Psychology (2015)

The British Psychological Society

www.wileyonlinelibrary.com

A generalized distance function for preferential choices

Nicolas A. J. Berkowitsch*, Benjamin Scheibehenne, Jörg Rieskamp and Max Matthäus

University of Basel, Switzerland

Many cognitive theories of judgement and decision making assume that choice options are evaluated relative to other available options. The extent to which the preference for one option is influenced by other available options will often depend on how similar the options are to each other, where similarity is assumed to be a decreasing function of the distance between options. We examine how the distance between preferential options that are described on multiple attributes can be determined. Previous distance functions do not take into account that attributes differ in their subjective importance, are limited to two attributes, or neglect the preferential relationship between the options. To measure the distance between preferential options it is necessary to take the subjective preferences of the decision maker into account. Accordingly, the multi-attribute space that defines the relationship between options can be stretched or shrunk relative to the attention or importance that a person gives to different attributes describing the options. Here, we propose a generalized distance function for preferential choices that takes subjective attribute importance into account and allows for individual differences according to such subjective preferences. Using a hands-on example, we illustrate the application of the function and compare it to previous distance measures. We conclude with a discussion of the suitability and limitations of the proposed distance function.

1. Introduction

Past work has illustrated repeatedly that when people make choices, they do not evaluate options independently, but rather evaluate options relative to each other (Huber, Payne, & Puto, 1982; Rieskamp, Busemeyer, & Mellers, 2006; Simonson & Tversky, 1992; Slovic & Tversky, 1974). One way to explain these interdependent evaluations is to assume distance-dependent competition of options in the multi-attribute space (Roe, Busemeyer, & Townsend, 2001; Rooderkerk, Van Heerde, & Bijmolt, 2011). Accordingly, options in the multi-attribute space compete with each other based on their perceived similarity, where similarity is defined as a decreasing function of their distance (Nosofsky, 1984; Shepard, 1987). The more similar two options are, the more strongly the evaluation of one option will affect the evaluation of the other option. In the present paper, we propose a generalized distance function (GDF) that defines the relationships between preferential choice options in a multi-attribute space and takes a person's subjective preferences into account.

*Correspondence should be addressed to Nicolas A. J. Berkowitsch, Department of Psychology, University of Basel, Missionsstrasse 62a, 4055 Basel, Switzerland (email: nicolas.berkowitsch@unibas.ch).

DOI:10.1111/bmsp.12048

The psychological distance between options cannot be defined simply by the Euclidean distance in the multi-attribute space, because this ignores the preferential relationship between options and the importance that the decision maker gives to different attributes. Figure 1 illustrates this difference for the hypothetical choice between notebook computers. Figure 1a shows the options on the original attribute space, while Figure 1b provides a visual impression of the psychological distances. In the figure, the labelled dots A, B, and C represent three different notebooks described by their processor speed and their battery life such that higher values are monotonically related to desirability. Although the Euclidean distances between notebooks A and C and notebooks A and B in the multi-attribute coordinate space are exactly the same, we argue that their psychological distances differ. Notebooks A and B are highly competitive, as they both lie on the indifference line, meaning that a change from notebook A to notebook B appears acceptable, because the loss in battery life is compensated by notebook B's higher processor speed. In comparison, notebook C appears completely inferior to notebooks A and B and a change from notebooks A or B to notebook C is not acceptable for the decision maker, because this change would lead to a loss of battery life and a loss of processor speed. In this example, notebook C is *dominated* by notebooks A and B. Consequently, notebooks A and B are perceived as more similar to each other than either is to notebook C, so that the psychological distance is smaller between A and B than between A and C (or B and C).

One way to represent the psychological distance between the choice options is to define a distance along the dominance line and the indifference line, as suggested by Hotaling, Busemeyer, and Li (2010) and Huber *et al.* (1982). The two distances can be

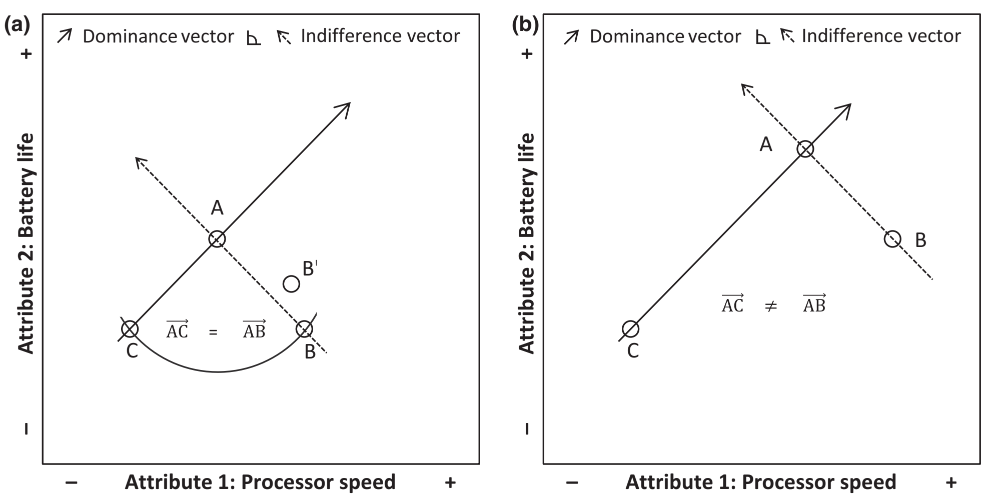

Figure 1. The (a) Euclidean distance and (b) psychological distance for the same set of notebook computers, labelled A–C. (a) The two options B and C have the same Euclidean distance but not the same psychological distance to option A. The psychological distance of option C to option A and option B is relatively large, because option C is dominated by the other two options. A person giving higher importance to the 'battery life' attribute as compared to the 'processor speed' attribute will be indifferent only between option A and option B′ and not between option A and option B, resulting in a rotated indifference vector. (b) The psychological space is stretched in the dominance direction, whereas distance in the indifference direction is the same as in (a).

expressed by an *indifference vector* (dashed lines in Figure 1) and a *dominance vector* (continuous lines in Figure 1). In addition, the psychological distance between options also depends on the subjective importance given to the different attributes. A person giving equal importance to the two attributes might be indifferent about notebooks A and B. However, a person giving higher importance to the attribute 'battery life' as compared to the attribute 'processor speed' will no longer be indifferent between notebooks A and B, but will be indifferent between notebook A and a hypothetical notebook B′ that has a slightly higher battery life than B (see Figure 1a).

This example illustrates that to be broadly applicable, a psychological distance function needs to meet several requirements. It should capture the different preferential relationships between options (i.e., whether one option is dominant or not); it must be flexible enough to account for the varying importance of different attributes; and it should be applicable for choice problems with more than two attributes. Past research on distance functions has addressed some of these requirements (Hotaling *et al.*, 2010; Huber *et al.*, 1982; Nosofsky, 1986; Rooderkerk *et al.*, 2011; Wedell, 1991), but an approach incorporating all requirements simultaneously is missing. Our goal in the present paper is to provide a GDF for preferential choice options and to compare it to alternative approaches. As a starting point, we review approaches that offer partial solutions to the listed requirements. Based on these solutions, we develop a GDF fulfilling all requirements simultaneously. We then illustrate the application of the GDF with a concrete example. We conclude with a discussion of the suitability and limitations of the proposed distance function and of possible ways to test the function empirically.

2. Psychological distance

Past researchers on perceptual categorization emphasized the difference between the objective distance between objects and their psychological distance. For instance, Nosofsky (1986) argued that individuals often do not distribute their attention equally to each dimension describing (perceptual) objects, hence their psychological space differs. To address this aspect, Nosofsky suggested a weighted Minkowski metric, which stretches and shrinks the different dimensions relative to the attention that an individual devotes to each dimension (Carroll & Wish, 1974). More attention to an attribute stretches the dimension in the space, whereas less attention shrinks the dimension in the space. This approach is common in categorization research (Nosofsky & Johansen, 2000; Nosofsky & Zaki, 2002; Zaki, Nosofsky, Stanton, & Cohen, 2003).

To reach a better understanding of judgement and decision making, it is important to take a slightly different perspective on the distance between options. That is because past research indicated that the evaluation of choice options critically depends on the degree to which an available option dominates another. For example, studies investigating the attraction effect (Huber & Puto, 1983; Huber *et al.*, 1982) report that the effect only emerges when the dominated option is close enough to the dominating choice option (Bhatia, 2013; Trueblood, Brown, & Heathcote, 2014; Tsetsos, Usher, & Chater, 2010). Thus, to account for the attraction effect it is important to know not only whether an option dominates another, but also to what extent this option is dominated.

To account for this dimension, Huber *et al.* (1982) advocated treating distance in the dominance and indifference directions differently. They suggested that 'dominated items . . . are represented in the limit as being an infinite distance below those items that dominate them' (p. 92), whereas 'distances among non-dominating pairs . . . must be finite' (p. 92). In other words, they proposed placing a higher weight on the distance in the dominance direction relative to the distance in the indifference direction. Figure 1 illustrates this idea. The distance between options A and B is the same in both graphs, whereas the distance between options A and C is larger in Figure 1b than in Figure 1a, because the psychological space is stretched in the dominance direction.

In his dimensional weight model Wedell (1991) built on this idea and suggested rotating the indifference and dominance vectors to account for the preferential relationship when a dominated option is added to the choice set. The very same logic could be applied to account for varying levels of attribute importance. First, consider an individual who weights equally the two attributes 'processor speed' and 'battery life' in the notebook example. This can be expressed by choosing an indifference dimension and a dominance dimension along the diagonals of the Cartesian coordinate plane (see the solid indifference and dominance vectors **iv** and **dv** in Figure 2). Accordingly, an individual who weights processor speed higher (lower) than battery life would be depicted by a steeper (less steep) indifference vector. For illustration, these vectors are labelled **iv**′ and **iv**″ in Figure 2. In the extreme case, where one attribute weight is infinitely more important than the other attribute, the indifference vector would be parallel to the axes of the coordinate plane.

In their preferential choice model, Rooderkerk *et al.* (2011) discussed an alternative approach to account for individual differences. They suggested the indifference vector is the same for all individuals, whereas the dominance vector (also called the preference vector) is multiplied by an individual weight, yielding dominance vectors of different lengths. The rotation approach (i.e., rotating the indifference vector) from Wedell (1991)

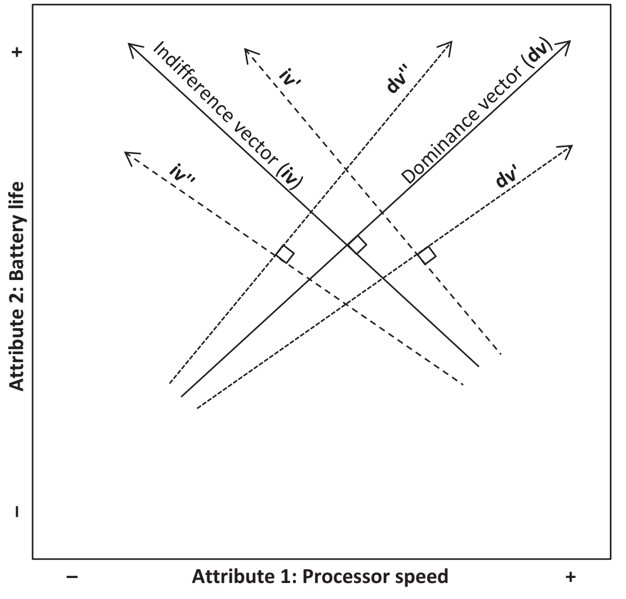

Figure 2. Individual preferences in a two-attribute space illustrated by an individual indifference vector and an orthogonal dominance vector. The solid indifference and dominance vectors indicate a person who weights the two attributes 'processor speed' and 'battery life' equally, whereas the dashed indifference and dominance vectors indicate a person giving a higher weight to either processor speed (**iv**′ and **dv**′) or battery life (**iv**″ and **dv**″).

has some advantages over Rooderkerk *et al.*'s (2011) suggestion of individual dominance vectors: The indifference vector directly expresses how many units an individual is willing to give up to increase the other attribute by one unit without introducing an additional parameter to weight the dominance vector.

This overview of past research reveals that several approaches to distance functions have addressed some of the aforementioned requirements, but none of them fulfil all requirements simultaneously (see Table 1). In the following we outline a distance function for preferential choice that does precisely that, and thus we refer to it as a GDF. We describe the function on a conceptual level and then provide the mathematical details, followed by a concrete example.

3. A generalized distance function

We continue with the notebook example with two attributes (processor speed and battery life) before generalizing it to multiple attributes. As a first step we need to ensure that the attributes are comparable with each other, to avoid distortions due to different scales and ranges of attribute values. This can be achieved by standardizing the attribute values, so that they have the same range, for example between 0 and 10. This additionally increases the visual interpretability of the indifference vector. Next, we have to specify the directions and lengths of the indifference and dominance vectors. Because the dominance vector is orthogonal to the indifference vector (Tversky, Sattah, & Slovic, 1988; Wedell, 1991), the direction of the dominance vector follows from the indifference vector. The direction of the indifference vector can be determined by so-called 'exchange ratios', which indicate how many units of one attribute an individual is willing to give up to increase the other attribute by one unit. That is, the indifference vector contains information about the exchange ratios between the attributes. For options described by more than two attributes, the number of possible exchange ratios increases. These multiple exchange ratios can be captured by *multiple indifference vectors* forming an indifference plane.

For example, all possible exchange ratios of an option described by five attributes can be captured with four exchange ratios, resulting in four indifference vectors – One fewer than the number of attributes (for details, see Section 4). Whereas the number of indifference vectors depends on the number of attributes, the number of distance types remains constant: Options are still described by *distance in the indifference direction* and by *distance in the dominance direction*. An option either dominates another option or it does not, which means that there is still only a single dominance vector even if options are described on several attributes. As a consequence, a single parameter is sufficient to weight distance in the dominance direction more strongly than in the indifference direction.

As is the case for two attributes (i.e., one indifference vector), the dominance vector has to be orthogonal to all indifference vectors in the multi-attribute case. Because of this property, the dominance vector can be derived from the indifference vectors (see Section 4 for details). To obtain indifference vectors and a dominance vector of equal lengths, we simply normalize them to the Euclidean length of 1, by dividing each vector by its Euclidean length.

Next, we express the distance between two options by means of the distance in the indifference and dominance direction. The line connecting the two options is called the distance vector, represented in the standard attribute coordinate plane. Instead of using these attribute dimensions, we now express the distance vector in terms of the

Table 1. Comparison of distance functions

Approach	Preferential relationship between options	Weighting of dominance relative to indifference direction	Importance weighting of attributes	Multiple attributes	Number of free parameters (where n = number of attributes)	Nested within the GDF
Huber *et al.* (1982)	✓	✓	–	–	1 'dominance' parameter[a]	Yes
Nosofsky (1986)	–	–	✓	✓	$n + 1$ (n attention weight parameters and 1 distance metric parameter)	No
Wedell (1991)	✓	–	✓	–	2 attribute weight parameters	Yes
Hotaling *et al.* (2010)	✓	✓	–	–	2 + 1 (2 attribute weight parameters and 1 'dominance' parameter)[b]	Yes
Rooderkerk *et al.* (2011)	✓	–	✓	–	2 + 1 (2 attribute weight parameters and 1 'dominance' parameter)[c]	Yes
GDF	✓	✓	✓	✓	n ($n - 1$ attribute weight parameters and 1 'dominance' parameter)	–

Note. GDF, generalized distance function.
[a]Verbally mentioned.
[b]Both attention weight parameters are fixed to .5.
[c]One attribute weight and the dominance parameter is fixed to 1.

indifference vectors and the dominance vector. This is achieved by a change of basis. The transformed distance vector indicates how many units of the standardized indifference vectors and of the standardized dominance vector are necessary to 'travel' from one to another option.

Finally, we calculate the Euclidean length of the transformed distance vector and multiply distance in the dominance direction by a parameter greater than 1. If this parameter is set to 1, the resulting distance equals the Euclidean distance of the unweighted distance vector and hence dominance would not be considered.

4. Mathematical formalization of the generalized distance function

To specify the GDF, we first rescale the attribute values (v_{old}) of the n attributes to equal ranges, from $\min_{\text{new}}$ to $\max_{\text{new}}$, according to

$$v_{\text{new}} = \min_{\text{new}} + \frac{(v_{\text{old}} - \min_{\text{old}})(\max_{\text{new}} - \min_{\text{new}})}{\max_{\text{old}} - \min_{\text{old}}}, \tag{1}$$

where $\min_{\text{old}}$ refers to the theoretical minimum value of the old attribute range and $\max_{\text{old}}$ refers to the theoretical maximum value of the old attribute range. Next, we define an importance weight vector $\mathbf{W}$ that contains the individual importance weights for the n attributes and scale the weights such that they sum to 1. The number of possible exchange ratios is then given by

$$n_{\text{ExchRat}} = \frac{n(n-1)}{2}, \tag{2}$$

where n is the number of attributes. A strategy to reduce the n_{ExchRat} needed without loss of information is to compare each attribute against an arbitrary attribute, for example, against the first attribute. Therefore, each indifference vector $\{\mathbf{iv}_j\}_{j=1}^{n-1}$ is an n-dimensional vector and can be calculated as

$$\mathbf{iv}_j = \begin{bmatrix} -\frac{w_{j+1}}{w_1} \\ 0 \\ \vdots \\ 0 \\ \frac{w_1}{w_1} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{w_{j+1}}{w_1} \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \text{for all } j = 1, \ldots, n-1, \tag{3}$$

where w_1/w_1 (=1) is at the (j + 1)th position.

Notice that this leads to $n - 1$ indifference vectors. Independent of the number of attributes, each $\mathbf{iv}_j$ has exactly two non-zero entries, the exchange ratios. This is because the other n_{ExchRat} can be recovered from the entries of the $n - 1$ indifference vectors. The first entry in each $\mathbf{iv}_j$ marks the exchange ratio between the (j + 1)th and the first attribute and the algebraic sign indicates that the vector is pointing toward the axes of the (j + 1)th attribute. In other words, $\mathbf{iv}_j$ reveals how many units of the (j + 1)th attribute are gained for giving up on one unit of the first attribute.

Because the n-dimensional dominance vector $\mathbf{dv}$ should be orthogonal to all $n - 1$ indifference vectors, we determine the case in which the dot product of each indifference vector with $\mathbf{dv}$ is zero. In general, this vector fulfils

$$\mathbf{iv}_j \cdot \mathbf{dv} = 0, \quad \text{for all } j = 1, \ldots, n-1, \tag{4}$$

which leads to the generalized form

$$\mathbf{dv} = \begin{bmatrix} \frac{w_1}{w_1} \\ \frac{w_2}{w_1} \\ \vdots \\ \frac{w_j}{w_1} \\ \vdots \\ \frac{w_n}{w_1} \end{bmatrix}. \tag{5}$$

Now we can construct the $n \times n$ matrix $\mathbf{B}^*$, containing the $n - 1$ indifference vectors $\mathbf{iv}_1, \ldots, \mathbf{iv}_{n-1}$ and the dominance vector $\mathbf{dv}$. This is

$$\mathbf{B}^* = [\mathbf{iv}_1, \ldots, \mathbf{iv}_j, \ldots, \mathbf{iv}_{n-1}, \mathbf{dv}]. \tag{6}$$

Observe that $\mathbf{B}^*$ is a basis of the attribute space. To standardize the lengths of the indifference vectors and the dominance vector to 1, each vector is divided by its Euclidean lengths $l_{\mathbf{iv}}$ and $l_{\mathbf{dv}}$, where $\{l_{\mathbf{iv}_j}\}_{j=1}^{n-1}$

$$l_{\mathbf{iv}_j} = \|\mathbf{iv}_j\|_2, \quad \text{for all } j = 1, \ldots, n-1, \tag{7}$$

and

$$l_{\mathbf{dv}} = \|\mathbf{dv}\|_2. \tag{8}$$

Thus, we obtain the basis $\mathbf{B}$, which is

$$\mathbf{B} = \left[\frac{\mathbf{iv}_1}{l_{\mathbf{iv}_1}}, \ldots, \frac{\mathbf{iv}_j}{l_{\mathbf{iv}_j}}, \ldots, \frac{\mathbf{iv}_{n-1}}{l_{\mathbf{iv}_{n-1}}}, \frac{\mathbf{dv}}{l_{\mathbf{dv}}}\right]. \tag{9}$$

$\mathbf{B}$ contains the standardized indifference vectors and the standardized dominance vector. Next, we define the distance of the position of options (i.e., A, B, or C) in the Cartesian coordinate system on the standardized scale (e.g., 0–10 in the example below) as a vector labelled $\mathbf{dist}_{\text{stand}}$. We want to transform this distance vector that connects the two options into the new distance vector $\mathbf{dist}_{\text{trans}}$ that expresses the trajectory path created by the previously introduced indifference vectors and by the dominance vector. This is achieved by a change of basis. Therefore we multiply the inverse of the basis $\mathbf{B}$ by $\mathbf{dist}_{\text{stand}}$:

$$\mathbf{dist}_{\text{trans}} = \mathbf{B}^{-1} \cdot \mathbf{dist}_{\text{stand}}. \tag{10}$$

The first $n - 1$ entries of $\mathbf{dist}_{\text{trans}}$ express the distance in units of each $\mathbf{iv}_j$, whereas the last entry of $\mathbf{dist}_{\text{trans}}$ expresses the distance in units of $\mathbf{dv}$. Now we need to calculate the Euclidean length D^2 of $\mathbf{dist}_{\text{trans}}$ and multiply the distance in the dominance direction by a parameter $wd > 1$. This ensures that the distance in the dominance direction is weighted more strongly than the distance in the indifference directions. This is computed as follows:

$$D^2 = \mathbf{dist}'_{\text{trans}} \cdot \mathbf{A} \cdot \mathbf{dist}_{\text{trans}} \tag{11}$$

where $\mathbf{A}$ is an $n \times n$ diagonal matrix and is constructed in the following way:

$$\mathbf{A}_{j,j} = \begin{cases} 1 & \text{if } j = 1, \ldots, n-1, \\ wd & \text{if } j = n. \end{cases} \tag{12}$$

This ensures that only the difference in dominance direction, the last column of $\mathbf{dist}_{\text{trans}}$, is weighted by wd. By setting $\mathbf{A}$ to the identity matrix (i.e., $wd = 1$), one obtains the standard Euclidean norm.

The steps required to derive the GDF can be summarized as follows:

1. Standardize all attribute values to the same range (e.g., between 0 and 10) using the theoretical possible range of attribute values.
2. Determine the weights $\mathbf{W}$ to calculate the matrix $\mathbf{B}^*$, containing the $n-1$ indifference vectors and the dominance vector.
3. Normalize the indifference vectors and the dominance vector in $\mathbf{B}^*$ to the Euclidean length of 1 to obtain $\mathbf{B}$.
4. Express the distance vector $\mathbf{dist}_{\text{stand}}$ in terms of the introduced basis $\mathbf{B}$ to get the transformed distance vector $\mathbf{dist}_{\text{trans}}$.
5. To obtain D^2, determine the Euclidean length of $\mathbf{dist}_{\text{trans}}$ and place a higher weight on distance in the dominance direction relative to the indifference direction.

Defined in that way, the GDF has $n - 1$ free parameters for the attribute importance weights and one additional free wd parameter for the weighting of the dominance direction relative to the indifference direction. As shown in Table 1, this number is comparable to existing distance functions. Table 1 further indicates which of the existing distance functions are nested within the GDF that we propose. For instance, in a situation where the choice options are described on only two attributes and the attribute weights are assumed to be equal, the GDF resembles the function suggested by Hotaling *et al.* (2010). Likewise, for a choice situation with two attributes, Rooderkerk *et al.* (2011) estimate the weights of one attribute (while fixing the other to 1) and they do not include a stretching parameter (i.e., wd is fixed to 1).

In the next section, we provide an example in which we apply the GDF to calculate the distances between three notebooks, each described by three attributes, and compare the distances obtained to previous distance functions.

5. Example

As an example, imagine you want to buy a new notebook computer to conduct your research. The notebook should be fast enough to run your simulations smoothly, and the battery life should cover your daily commuting time by train. During your internet search, you came across notebooks with processors up to 5 GHz and batteries that last for 6 h, and you excluded notebooks with display sizes larger than 30 in. Table 2 shows the three notebooks you ended up with.

Let us further assume that you care about having a high processor speed (PS) as much as a long battery life (BL) and that you care less about having a large display size (DS). These

Table 2. Attribute values of three different notebook computers

Attribute	Notebook A	Notebook B	Notebook C
Processor speed in gigahertz [1–5][a]	3.0	2.2	2.2
Battery life in hours [2–6]	3.6	4.0	2.8
Display size in inches [4–30]	11.8	17.0	9.2

Note. [a]The numbers in brackets indicate the possible value range.

preferences can be expressed by the following subjective importance weights: $w_{\mathrm{PS}} = .4$, $w_{\mathrm{BL}} = .4$, and $w_{\mathrm{DS}} = .2$. Therefore, you clearly prefer notebooks A and B over notebook C because they dominate notebook C. The choice between notebooks A and B seems to be more difficult. Going back and forth between the two notebooks, you realize that both have some advantages and disadvantages and you deem them equally attractive. That is, you are indifferent between notebooks A and B. Because notebook A dominates notebook C and competes with notebook B, the perceived distance between notebooks A and C should be larger than between notebooks A and B. However, if we determined the standard Euclidean distance in the rescaled attribute space, the two distances would be the same, as illustrated below.

To apply the GDF to the notebook example, we start by standardizing the ranges of all three attributes ($n = 3$) to equal ranges, for instance, between 0 and 10 according to equation (1). Now, the notebooks can be represented as points in a multi-attribute space with notebook A = (5, 4, 3), notebook B = (3, 5, 5), and notebook C = (3, 2, 2). Next, we determine the basis $\mathbf{B}^*$ containing the two indifference vectors and the dominance vector. The first indifference vector $\mathbf{iv}_1$ is given by

$$\mathbf{iv}_1 = \begin{bmatrix} -\frac{w_{\mathrm{BL}}}{w_{\mathrm{PS}}} \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix},$$

the second indifference vector $\mathbf{iv}_2$ is given by

$$\mathbf{iv}_2 = \begin{bmatrix} -\frac{w_{\mathrm{DS}}}{w_{\mathrm{PS}}} \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -.5 \\ 0 \\ 1 \end{bmatrix},$$

and the dominance vector $\mathbf{dv}$ is given by

$$\mathbf{dv} = \begin{bmatrix} 1 \\ \frac{w_{\mathrm{BL}}}{w_{\mathrm{PS}}} \\ \frac{w_{\mathrm{DS}}}{w_{\mathrm{PS}}} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ .5 \end{bmatrix}.$$

This leads to the basis

$$\mathbf{B}^* = \begin{bmatrix} -1 & -.5 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & .5 \end{bmatrix}.$$

Next, we normalize the indifference vectors and the dominance vector to the length of 1 by dividing them by their Euclidean lengths, where

$$l_{\mathbf{iv}_1} = \sqrt{(-1)^2 + 1^2 + 0} = \sqrt{2},$$
$$l_{\mathbf{iv}_2} = \sqrt{(-.5)^2 + 0 + 1^2} = \sqrt{1.25},$$

and

$$l_{\mathbf{dv}} = \sqrt{1^2 + 1^2 + .5^2} = \sqrt{2.25}.$$

Then the standardized basis is

$$\mathbf{B} = \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{-.5}{\sqrt{1.25}} & \frac{1}{\sqrt{2.25}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2.25}} \\ 0 & \frac{1}{\sqrt{1.25}} & \frac{.5}{\sqrt{2.25}} \end{bmatrix}.$$

To express the distance vector $\mathbf{dist}_{\text{stand}}$ between the notebooks in terms of the indifference vector and the dominance vector, we next make a change of basis. The standard distance vectors between notebooks A and C and between notebooks A and B are

$$\mathbf{dist}_{\text{standAC}} = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$$

and

$$\mathbf{dist}_{\text{standAB}} = \begin{bmatrix} 2 \\ -1 \\ -2 \end{bmatrix}.$$

To yield the transformed distance vector we apply equation (10), so that

$$\mathbf{dist}_{\text{transAC}} = \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{-.5}{\sqrt{1.25}} & \frac{1}{\sqrt{2.25}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2.25}} \\ 0 & \frac{1}{\sqrt{1.25}} & \frac{.5}{\sqrt{2.25}} \end{bmatrix}^{-1} \cdot \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$

and

$$\mathbf{dist}_{\text{transAB}} = \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{-.5}{\sqrt{1.25}} & \frac{1}{\sqrt{2.25}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2.25}} \\ 0 & \frac{1}{\sqrt{1.25}} & \frac{.5}{\sqrt{2.25}} \end{bmatrix}^{-1} \cdot \begin{bmatrix} 2 \\ -1 \\ -2 \end{bmatrix} = \begin{bmatrix} -\sqrt{2} \\ -\sqrt{5} \\ 0 \end{bmatrix}.$$

The distance $\mathbf{dist}_{\text{transAC}}$ indicates that to reach point A from point C, we need to move three units along the dominance vector and none along the two indifference vectors. To reach point A from point B, we need to move by $-\sqrt{2}$ and $-\sqrt{5}$ units along the first and second indifference vectors, and no units in the dominance direction, which means that we are moving on the indifference plane.

By setting $wd = 10$, we assume that distance in the dominance direction is perceived 10 times more strongly than in the indifference direction, resulting in

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 10 \end{bmatrix}.$$

Finally, we can calculate the Euclidean distances D^2_{AC} between options A and C and D^2_{AB} between options A and B by applying equation (11). We obtain $D^2_{\mathrm{AC}} = 90$ and $D^2_{\mathrm{AB}} = 7$. If we had treated distances in the indifference and dominance directions equally (i.e., by setting wd to 1), the distance between options A and C would have decreased to $D^2_{\mathrm{AC}} = 9$, whereas D^2_{AB} would have remained 7, thus suggesting that options B and C have similar distances to option A. Note that if we had also ignored the weights given to the attributes and had determined the Euclidean distance in the standardized attribute space, both notebook C and notebook B would have had the same distance to notebook A of $D^2_{\mathrm{AC\ Eucl}} = 9(= 2^2 + 2^2 + 1^2)$ and $D^2_{\mathrm{AB\ Eucl}} = 9(= 2^2 + (-1)^2 + (-2)^2)$, incorrectly suggesting that notebooks A and C are perceived as being as similar to each other, as notebooks A and B are to each other.

6. Discussion

People often evaluate options relative to each other (Huber *et al.*, 1982; Rieskamp *et al.*, 2006; Simonson & Tversky, 1992; Slovic & Tversky, 1974). Therefore, many cognitive models of decision making take the similarity between options into account (Roe *et al.*, 2001; Rooderkerk *et al.*, 2011). The similarity between options is generally expressed as a decreasing function of their distance to each other (Shepard, 1987). To define the distance between options for preferential decision-making problems, past research has addressed important aspects such as the different preferential relationships between options, the varying subjective importance of different attributes, and the applicability to choice problems with more than two attributes. However, as shown in Table 1, thus far no approach has addressed all of these aspects simultaneously (cf. Hotaling *et al.*, 2010; Huber *et al.*, 1982; Nosofsky, 1986; Rooderkerk *et al.*, 2011; Wedell, 1991). To overcome this limitation, we have developed a new GDF for preferential choices that distinguishes the preferential relationship between options in a multi-attribute space, assigns different weights depending on the type of distance, and accounts for the individual degree of importance of different attributes. In the following, we will further discuss important prerequisites and underlying theoretical assumptions of the proposed distance function.

One important prerequisite is the precise measurement of subjective attribute weights that the function takes as input. To obtain these weights, past research on judgement and decision making has proposed several approaches. One approach is to elicit people's subjective attribute weights by means of established methods from decision analysis (for a review, see Riabacke, Danielson, & Ekenberg, 2012). For example, according to the 'trade-off procedure' participants are asked to create indifferent choice pairs from which the weights are derived following multi-attribute value theory (Keeney & Raiffa, 1976). As an alternative approach, people's subjective attribute weights can also be estimated indirectly based on repeated choices between multiple choice options by means of a modelling approach. By using choice models that incorporate the attribute weight as free parameters (Roe *et al.*, 2001), researchers can estimate the weights using maximum likelihood methods (Lewandowsky & Farrell, 2011). This second approach can also be applied when the weights are estimated based on discrete choices. When presenting similar choices repeatedly, this approach can also be used to assess the reliability of the measures obtained. Choice data also provide the basis for many mathematical decision models that take the distance between alternative options into account, and thus can be easily extended with the GDF. This includes the multi-alternative decision field theory (Roe *et al.*, 2001), the leaky competing accumulator model (Usher & McClelland, 2001) and the extended contextual random utility model (Rooderkerk *et al.*, 2011) that all rely on quantitative information about the distances between available options. For a concrete implementation of this approach and its application to empirical choice data, see Berkowitsch, Scheibehenne, and Rieskamp (2014).

For choice models that require information about the options' similarity, researchers also need to define the relationship between similarity and distance, for example by assuming that similarity is a decreasing (exponential) function of the distance. In addition to this, for random utility models (McFadden, 2001), which are frequently applied in many areas of research in judgement and decision making and economics, one could replace the individual parameters in the variance–covariance matrix with the GDF. Implementing the GDF in the models mentioned above also allows researchers to empirically estimate the parameter *wd* which weights the distance in dominance direction more strongly than the distance in the indifference directions.

The GDF that we outlined further assumes that people's subjective importance or weighting of one attribute relative to another attribute can be expressed by exchange ratios. Towards a precise understanding of this weighting, it is important to estimate the underlying attribute weights reliably. However, inaccurate exchange ratios cannot turn a dominated choice option into a dominating option. This is because in the extreme case, where one attribute weight is wrongly estimated as infinitely more important than the other attribute(s), the formerly decreasing indifference vector would change into an indifference vector parallel to the standard unit vector, but cannot change into an increasing indifference vector.

The assumption of linear exchange ratios further implies that a disadvantage on one unit for one attribute can always be compensated for by a specific number of units of an advantage on the other attribute. This assumption might sometimes be violated when people make decisions. For instance, a decision maker might require a minimum battery life for her new notebook. In this case, she would do better to apply nonlinear exchange ratios. Here, for instance, the closer a value of an attribute gets to a decision maker's required minimum value (e.g., minimum hours of battery life), the more units of the other attributes are required to compensate a further decrease of the first attribute. In the extreme, the exchange ratio becomes indefinitely small or big, which can be captured by asymptotic indifference curves. This idea is supported by findings from Chernev (2004). He suggested that due to people's extremeness aversion, they prefer options that are closer to the so-called 'attribute-balance line,' which he defined as the line connecting 'all potential options with identical values on both attributes' (p. 251). This can be captured by a convex indifference curve. However, for simplicity we have applied linear exchange ratios to express the importance of different attributes. This approach should be suitable as long as the attribute range does not include very extreme values. It remains to be tested under what circumstances nonlinear exchange ratios lead to a substantial advantage in describing the psychological distance for multi-attribute options in preferential decision making and ultimately to better prediction of preferential choices.

When normalizing the indifference vectors and the dominance vector to equal lengths of 1, and later when we calculated the length of the transformed distance vectors, we assumed the 2-norm distance (i.e., the Euclidean distance). Although the 2-norm distance is widely applied for preferential choices (Hotaling *et al.*, 2010; Rooderkerk *et al.*, 2011), the 1-norm (i.e., the city block) distance is frequently applied in categorization research for stimuli with highly separable dimensions (Nosofsky & Zaki, 2002). For example, one hypothesis following from this is that with an increasing number of attributes describing a preferential choice option, the dimensions become less separable and the 2-norm distance can outperform the 1-norm. This can also be tested – similarly to the Minkowski r metric – by introducing the norm as a free parameter r to see whether the 1-norm (i.e., $r = 1$) or the 2-norm (i.e., $r = 2$) leads to more accurate predictions of preferential choices.

The approach we have taken to define a GDF shares some aspects with the multidimensional scaling approach (MDS; Kruskal, 1964a,b). MDS is often applied in research on judgement and decision making to optimally visualize the perceived similarity between options in the attribute space (for a review, see Carroll & Arabie, 1980). MDS also allows the definition of a new attribute space to express the similarities between options. However, in contrast to MDS, our approach takes the specific preferential relationship between options into account, and it can be readily applied to mathematical models of perception and decision making for single individuals (e.g., Berkowitsch *et al.*, 2014).

Besides fulfilling its general purpose of measuring the perceived distance between preferential choice options, the GDF is particularly useful for studying so-called context effects (Huber & Puto, 1983; Huber *et al.*, 1982; Simonson & Tversky, 1992; Slovic & Tversky, 1974). Context effects refer to choice situations in which preferential choice options are evaluated relative to each other. According to the similarity effect – a well-known context effect – similar choice options compete with each other more strongly than dissimilar options. Therefore adding an option that is similar to one but dissimilar to another option increases the preference for the dissimilar relative to the similar option (Tversky, 1972a,b). To better understand the similarity effect, it is therefore important to determine how the similarity or dissimilarity between two options is perceived. Whereas for one person two given options might be perceived as similar, for another person one of the two options might dominate the other. This can be captured by the GDF that we proposed, because it accounts for individual differences and distinguishes the preferential relationship between options.

Research on context effects has focused mainly on options described by two attributes. With an increasing number of attributes, it becomes harder to tell dominated options apart from indifferent options. Because the GDF allows for multi-attribute options, it can be applied to examine context effects in the multi-attribute space. Furthermore, the GDF can be fruitfully applied to research on individual differences, because it allows the indifference vectors to rotate through the multi-attribute space and hence quantifies individuals' exchange ratios between the attributes. This property also lends itself to a combination with existing choice models that try to estimate the subjective importance that people give to different attributes.

Incorporating the function into a mathematical choice model also provides the basis to empirically test and compare the function relative to alternative approaches in terms of predictive accuracy and complexity (Forster, 2000). Although providing such an empirical test was not the goal of the present theoretical paper, it would be an important next step to better understand how people perceive the similarity and distance between options that are often evaluated relative to each other.

In summary, the GDF that we have proposed accounts for several requirements that previous approaches have only addressed in isolation. This allows researchers to study preferential choices and context effects in more depth and to investigate individual differences in the multi-attribute space. Ultimately, this should lead to advancement of decision theory by taking the similarity between choice options into account for providing better explanations and predictions of human preferential choices.

Acknowledgment

This study was funded by Swiss National Science Foundation (100014_130149).

References

Berkowitsch, N. A. J., Scheibehenne, B., & Rieskamp, J. (2014). Testing multialternative decision field theory rigorously against random utility models. *Journal of Experimental Psychology: General*, *143*, 1331–1348. doi:10.1037/a0035159

Bhatia, S. (2013). Associations and the accumulation of preference. *Psychological Review*, *120*, 522–543. doi:10.1037/a0032457

Carroll, J. D., & Arabie, P. (1980). Multidimensional scaling. *Annual Review of Psychology*, *31*, 607–649. doi:10.1146/annurev.ps.31.020180.003135

Carroll, J. D., & Wish, M. (1974). Models and methods for three-way multidimensional scaling. In D. H. Krantz, R. C. Atkinson, R. D. Luce & P. Suppes (Eds.), *Contemporary developments in mathematical psychology* (Vol. 2, pp. 57–105). San Francisco, CA: Freeman.

Chernev, A. (2004). Goal orientation and consumer preference for the status quo. *Journal of Consumer Research*, *31*, 557–565. doi:10.1086/425090

Forster, M. R. (2000). Key concepts in model selection: Performance and generalizability. *Journal of Mathematical Psychology*, *44*, 205–231. doi:10.1006/jmps.1999.1284

Hotaling, J. M., Busemeyer, J. R., & Li, J. (2010). Theoretical developments in decision field theory: Comment on Tsetsos, Usher, and Chater. *Psychological Review*, *117*, 1294–1298. doi:10.1037/a0020401

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

Huber, J., & Puto, C. (1983). Market boundaries and product choice: Illustrating attraction and substitution effects. *Journal of Consumer Research*, *10*, 31–44.

Keeney, R. L., & Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value tradeoffs*. New York, NY: Cambridge University Press.

Kruskal, J. B. (1964a). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, *29*, 1–27. doi:10.1007/BF02289565

Kruskal, J. B. (1964b). Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, *29*, 115–129. doi:10.1007/BF02289694

Lewandowsky, S., & Farrell, S. (2011). *Computational modeling in cognition*. Thousand Oaks, CA: Sage.

McFadden, D. (2001). Economic choices. *American Economic Review*, *91*, 351–378. doi:10.1257/aer.91.3.351

Nosofsky, R. M. (1984). Choice similarity and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 104–114. doi:10.1037/0278-7393.10.1.104

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39–57. doi:10.1037/0096-3445.115.1.39

Nosofsky, R. M., & Johansen, M. K. (2000). Exemplar-based accounts of 'multiple-system' phenomena in perceptual categorization. *Psychonomic Bulletin & Review*, *7*, 375–402.

Nosofsky, R. M., & Zaki, S. R. (2002). Exemplar and prototype models revisited: Response strategies, selective attention, and stimulus generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 924–940. doi:10.1037/0278-7393.28.5.924

Riabacke, M., Danielson, M., & Ekenberg, L. (2012). State-of-the-art prescriptive criteria weight elicitation. *Advances in Decision Sciences*, *46*, 1–24. doi:10.1155/2012/276584

Rieskamp, J., Busemeyer, J. R., & Mellers, B. A. (2006). Extending the bounds of rationality: Evidence and theories of preferential choice. *Journal of Economic Literature*, *44*, 631–661. doi:10.1257/jel.44.3.631

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological Review*, *108*, 370–392. doi:10.1037/0033-295X.108.2.370

Rooderkerk, R. P., Van Heerde, H. J., & Bijmolt, T. H. A. (2011). Incorporating context effects into a choice model. *Journal of Marketing Research*, *48*, 767–780. doi:10.1509/jmkr.48.4.767

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, *237*, 1317–1323. doi:10.1509/jmkr.48.4.767

Simonson, I., & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research*, *29*, 281–295. doi:10.2307/3172740

Slovic, P., & Tversky, A. (1974). Who accepts Savage's axiom? *Behavioral Science*, *19*, 368–373. doi:10.1002/bs.3830190603

Trueblood, J. S., Brown, S. D., & Heathcote, A. (2014). The multiattribute linear ballistic accumulator model of context effects in multialternative choice. *Psychological Review*, *121*, 179–205. doi:10.1037/a0036137

Tsetsos, K., Usher, M., & Chater, N. (2010). Preference reversal in multiattribute choice. *Psychological Review*, *117*, 1275–1291. doi:10.1037/a0020580

Tversky, A. (1972a). Choice by elimination. *Journal of Mathematical Psychology*, *9*, 341–367. doi:10.1016/0022-2496(72)90011-9

Tversky, A. (1972b). Elimination by aspects: A theory of choice. *Psychological Review*, *79*, 281–299. doi:10.1037/h0032955

Tversky, A., Sattah, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review*, *95*, 371–384. doi:10.1037/0033-295X.95.3.371

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*, 550–592. doi:10.1037/0033-295X.108.3.550

Wedell, D. H. (1991). Distinguishing among models of contextually induced preference reversal. *Journal of Experimental Psychology*, *17*, 767–778. doi:10.1037/0278-7393.17.4.767

Zaki, S. R., Nosofsky, R. M., Stanton, R. D., & Cohen, A. L. (2003). Prototype and exemplar accounts of category learning and attentional allocation: A reassessment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 1160–1173. doi:10.1037/0278-7393.29.6.1160

Received 12 December 2013; revised version received 26 October 2014