Reflecting Emotional Aspects and Uncertainty in Multi-expert Evaluation: One Step Closer to a Soft Design-Alternative Evaluation Methodology

Stoklasa Jan, Talášek Tomáš, Stoklasová Jana

DOI: 10.1007/978-3-319-78075-7_17

Reflecting emotional aspects and uncertainty in multi-expert evaluation: one step closer to a soft design-alternative evaluation methodology

Jan Stoklasa, Lappeenranta University of Technology, School of Business and Management, Skinnarilankatu 34, 53850 Lappeenranta, Finland and Palacký University Olomouc, Faculty of Arts, Department of Applied Economics, Křížkovského 8, 771 47 Olomouc, Czech Republic, e-mail: jan.stoklasa@lut.fi

Tomáš Talášek, Palacký University Olomouc, Faculty of Arts, Department of Applied Economics, Křížkovského 8, 771 47 Olomouc, Czech Republic, e-mail: tomas.talasek@upol.cz

Jana Stoklasová, Marital and family counseling centre Prostějov, Bezručovo nám. 9, 796 01, Prostějov, Czech Republic, e-mail: jana.stoklasova.cz@gmail.com

Introduction

Multiple-expert evaluation is based on the assessment of the given object (e.g. a decision alternative, project or a design alternative) by several experts. The aim frequently is to obtain an objective assessment and to consider as wide range of points of view as possible, so that no drawback of the alternative is overlooked. This is understandable and well in line with the basic ideas of operations research, mainly with the requirement of multi-disciplinarity. Obviously the amount of expertise, the relevance of the experts for the given purpose, their “decision power” etc. can be reflected in the process i.e. by specifying the weights of their opinions/evaluations. The more diverse the set of experts is, the more comprehensive the overall evaluation obtained from them can be. On the other hand this diversity introduces several issues concerning the aggregation of the evaluations provided by these individuals. Even if we consider criteria that are measurable (or at least quantifiable), we need to make sure that the same measuring instrument is used, the same scales are applied and that all the experts have access to all the relevant information. Even when this is achieved and the weights of the experts (representing the value of their opinion in the particular situation) are determined, the confidence of the experts answers can be variable rendering the overall evaluation difficult to interpret.

When we consider less tangible criteria, the situation becomes even more challenging. With a decreasing ability to measure the values of the criterion, the need for qualitative approach to its assessment increases. Linguistic scales (Zadeh 1975), Likert-type scales (Likert 1932; Stoklasa, Talášek, Kubátová and Seitlová 2017) or semantic-differential-type scales (Osgood 1964; Osgood, Suci and Tannenbaum 1957) anchored with linguistic values are used. Unfortunately the use of linguistic values introduces another "degree of freedom" in the evaluation process. The words (linguistic expressions) used to anchor the scales can be understood differently (or in some cases even not understood at all) and even if there was some level of consensus concerning the denotative meaning of these linguistic terms, their connotations will most probably vary from person to person. A selection of the same linguistic value by two different evaluators thus no longer guarantee that the same evaluation was expressed by them. Unifying the understanding of the meaning of the linguistic terms can be a tedious task (Stoklasa 2014; Stoklasa and Talášek 2015). The uncertainty inherent in the use of linguistic labels and linguistic variables has to be reflected appropriately (see e.g. Stoklasa 2014; Stoklasa, Talášek and Musilová 2014; Talašová, Stoklasa and Holeček 2014) and in many cases this is done by the use of fuzzy modelling (Fiss 2011; Stoklasa and Talášek 2016; Stoklasa, Talašová and Holeček 2011; Stoklasa, Talášek et al. 2014; Stoklasa, Talášek and Luukka 2018; Talašová et al. 2014), interval-valued modelling (Stoklasa, Talášek and Stoklasová 2016; Stoklasa, Talášek and Stoklasová 2017; Stoklasa, Talášek and Stoklasová 2018) or by finding alternative lossless representations of the set of evaluations instead of their direct aggregation (Stoklasa, Talášek, Kubátová et al. 2017).

There is one more (and a rather interesting) perspective we can take on multi-expert evaluation. In practice we frequently need to obtain assessment of alternatives, products and projects that are not only "objective", but that also reflect the "gut feeling" of the evaluators and their emotions triggered by the alternative. This is important, since a "bad feeling" or “fear” from the suggested alternative can indicate that some criteria (potentially relevant only for a subset of the evaluators) might not have been considered, or even might not be consciously known to the evaluators. This is not a new finding in the field of operations research. Brill, Chang and Hopkins (1982) proposed the “modelling to generate alternatives” (MGA) approach and since then it has been frequently applied in various fields (see e.g. Yeomans (2011), Yeomans and Gunalay (2011) for some recent municipal waste management and environmental management applications). The main idea behind MGA is to replace the search for the best solution by searching for a set of sufficiently good solutions, which are comparably good, but which differ as much as possible from each other in their characteristics. This is supposed to provide solutions to the “gut unacceptability” of the best solution by providing comparably good alternatives to it, which are sufficiently different in terms of their characteristics, but not in terms of their outcome. This showcases that “hidden”, unknown or forgotten criteria do exist and their identification can prove to be crucial to a successful multi-expert evaluation. It thus seems that emotions can be a relevant factor in the evaluation - and should be reflected in the evaluation models. The emotional component of evaluation has been stressed in the context of Kansei engineering by Jindo, Hirasago and Nagamachi (1995), Nagamachi 1995 or Kobayashi and Kinumura (2017) and in the field of design (see e.g. Huang, Chen and Khoo 2012) and marketing, where the connections of emotions and products is very relevant.

Mainly alternatives that are good enough in terms of the measurable criteria and that do not trigger a defensive emotional reaction in the decision-makers responsible for the final choice, when suggested as solutions, have the potential to be accepted. It is thus reasonable to reflect the emotional component when needed and to use the information concerning the prevailing emotional tone (and its consistency among the experts) as an additional resource in final decision-making. This way a soft emotion-oriented evaluation can be considered either as an alternative to standard multi-expert evaluation methods using measurable criteria and well defined aggregation function, or as an additional approach to the quantitative one providing a qualitative insight, information on less tangible aspects of the evaluation situation and also on the consistency of the understanding of (or feeling about) the linguistic values used in the more qualitative context.

In this chapter we focus on this softer component of evaluation and show, how the uncertainty stemming from a lower understanding of the linguistic labels, their perceived irrelevance or lower confidence concerning the final answers can be combined with the information concerning variable emotional responses of the evaluators to the labels (the effect of the connotative aspects of their meaning) in a multi-expert evaluation methodology. We suggest to substitute crisp (real-number) values by interval values when the uncertainty is present. We also need to keep in mind that as the scales for measurable criteria need to be of the same type and of the same ranges to be meaningfully aggregated, so do the uncertainties - not all the types of uncertainty can be combined into a single overall uncertainty without the loss of meaning. We propose how to deal with these different types of uncertainty.

To have a clear application framework for such a soft multi-expert evaluation methodology, we choose the area of product design, where emotions not only play a crucial role, but where the stimulation of a specific emotion in the user of the product can even be one of the goals. In this area, the emotional design and Kansei engineering (Nagamachi 1995) approaches, introduced to reflect the consumers' needs in the design process, have already justified the focus on the emotional aspects of the evaluation. More specifically we are proposing a generalization of the product classification method in emotional design proposed by Huang et al. (2012). The original method uses Kansei adjectives and semantic differential scales (in their standard, real-numbered version) and introduces an inter-expert “emotional connotation" variability check through the assessment of Kansei tags in terms of their emotional connotation. This method will be summarized in the following section.

In the third section we propose a generalization of the data gathering procedure for the method which reflects the perceived irrelevance of the Kansei tags for the evaluation of a given alternative by introducing an uncertainty into the evaluation - converting the real-number evaluation into an interval one. This step is inspired by Stoklasa, Talášek and Stoklasová (2017) and Stoklasa, Talášek and Stoklasová (2018). Also the confidence of the evaluators' answers concerning the emotional connotation of the Kansei tags is recorded and reflected analogously. A new measure of emotional dissensus on the emotional-loading of the Kansei tags is proposed and its use in product evaluation and classification is discussed. The conclusions section follows.

Key concepts summary box

Modelling to generate alternatives (MGA) – an approach to optimization aiming on providing not one, but more feasible solutions as different from each other as possible, all with the values of the objective function close to the optimal value (see e.g. Brill, Chang and Hopkins 1982 or Yeomans 2011).

Kansei engineering – it is a consumer-oriented approach to product design based on the reflection of less tangible aspects such as feelings concerning the product in the design process. The aim is to inspire specific feelings by the features of the design alternative (see Nagamachi 1995; Jindo, Hirasago and Nagamachi 1995; or Kobayashi and Kinumura 2017).

Kansei adjectives – Kansei words in the form of adjectives, i.e. words describing customers' or consumers' needs, feelings and perceptions concerning the product (see e.g. Jiao, Zhang and Helander (2006) for a Kansei mining system).

Kansei tag – group or cluster of Kansei adjective corresponding to the same concept or basic emotion (Xu and Wunsch (2009) provide an example of a clustering algorithm suitable for the creation of Kansei-adjectives clusters, i.e. Kansei tags).

Likert scale – a psychometric measurement instrument popularized by Likert (see e.g. Likert (1932)) frequently used in questionnaires. Likert scales are discrete scales with linguistic labels on the agree-disagree or similar continuums, which are supposed to be symmetrical with respect to the middle point (either present in the scale itself, or theoretical; e.g. strongly agree, agree, undecided, disagree, strongly disagree) of the scale. Usually the equidistance of the scale-values is assumed.

Semantic differential – a method proposed by Osgood, Suci and Tannenbaum (1957) for the measurement of attitudes. The method utilizes discrete bipolar-adjective scales to get input information and uses factor analysis to define the semantic space and represents the attitude towards a concept (or its connotative meaning) as a point in this *n*-dimensional semantic space.

Basic-emotion based semantic differential method for product classification as proposed by Huang et al. (2012)

As suggested by Jiao, Zhang and Helander (2006) the Kansei adjectives can be used to facilitate the expressing of consumers' needs, emotional states and feelings in connection with the product that is being evaluated. The use of Kansei adjectives (or their clusters represented by Kansei tags; some clustering algorithms suitable for this purpose can be found e.g. in Xu and Wunsch (2009)) is well compatible with semantic-differential-type scales (or Likert-type scales) and as such presents a simple enough combination of tools for obtaining inputs for the evaluation process. It is therefore suggested also in the emotional design semantic-differential method based on basic emotions introduced in Huang et al. (2012).

Let us now consider p evaluators need to evaluate n alternatives with respect to m criteria (represented here by Kansei tags). We also consider q basic emotions which will be used to assess the variance in understanding the Kansei tags by different evaluators (the number and list of these basic emotions is dependent on the underlying theory we choose for the purpose). Huang et al. (2012) propose a 7-step procedure consisting of the following steps (here we present just a brief description with comments, see Huang et al. (2012, pp. 571—575) for more details):

1. *Selection of the Kansei adjectives* for the purpose of the evaluation and their grouping into *Kansei tags*. This step also involves specifying the set of alternatives to be evaluated. It is an initial step which in general terms requires the criteria (here represented by clusters of Kansei adjectives grouped under a unifying Kansei tag) and alternatives to be specified. Also the *set of basic emotions* should be specified in this step. We will consider all the Kansei tags to be represented by continuous universes $[-r, r]$, where $r > 0$, i.e. by intervals of the length $2r$. Note, that any other interval of the same length can

be used without any loss of information (just a linear transformation of the values of the interval would be required; e.g. Huang et al. (2012, pp. 573) use intervals $[1,7]$). The basic emotions will also be represented by continuous universes of a length possibly different from the length of the Kansei-tag universe, denoted by $[-d, d], d > 0$, i.e. by intervals of the length $2d$ (Huang et al. (2012, pp. 573) consider intervals $[0,10]$ for this purpose).

2. *Selection of the survey participants*. In other terms this step requires the selection of evaluators - i.e. experts. Different groups of evaluators can be considered (e.g. product users and designers). All the necessary points of view should be represented and the number of the evaluators needs to be reasonable. If needed, weights of the evaluators (i.e. the value of their opinion for the given purpose) can be specified.
3. *Evaluation of the alternatives with respect to the Kansei tags*. A schematic representation of a questionnaire that could be used for this purpose is summarized in the top part of Figure 1. The evaluation of the alternative a_i with respect to the Kansei tag KT_j by the evaluator k is represented by $x_{K_{ijk}} \in [-r, r]$ in further calculations; $i = 1, \dots, n$, $j = 1, \dots, m$ and $k = 1, \dots, p$.

**** Figure 1 about here ****

In essence the use of Kansei adjectives as anchors for the poles of Likert-type or semantic-differential-type scales is an example of simple linguistic modelling. As such it requires a uniform understanding of these adjectives (or Kansei tags) if the information has to be aggregated across the experts/evaluators.

Also the points of view and hence the evaluations of and attitudes towards the object can differ significantly in different subgroups of experts (as confirmed e.g. by Hsu, Chuang and Chang (2000), the importance of criteria can be seen differently and also the emotional connotation of the evaluation can be different. It thus makes sense to at least investigate how consistent the group of evaluators, or its subgroups, are in their interpretation of the criteria or linguistic labels used to represent them. Hence the connotation of the Kansei tags is checked in terms of their association with basic emotions in the next step.

4. *Assessment of the Kansei tags in terms of their emotional associations*. The upper part of Figure 2 again present the questionnaire used for this purpose and its lower part the conversion of the answers into the values $x_{E_{jlk}} \in [-d, d]$, i.e. numerical values representing the assessment of the Kansei tag KT_j by the evaluator k with respect to the basic-emotion BE_l; $j = 1, \dots, m$, $k = 1, \dots, p$ and $l = 1, \dots, q$.

**** Figure 2 about here ****

5. *Calculation of the mean value of each Kansei tag*, which is done by (1). This value $\mu_{K_{ij}}$ is supposed to represent the overall group evaluation of the alternative i with respect to the Kansei tag KT_j.

$$\mu_{K_{ij}} = \frac{\sum_{k=1}^{p} x_{K_{ijk}}}{p} \tag{1}$$

Since the Kansei tags can be interpreted differently by the evaluators, Huang et al. (2012) suggest to investigate the "semantic meaning" of the Kansei tags in terms of basic emotions. The idea behind this being that if the perception of the Kansei tag KT_j is very different among the evaluators (i.e. the variability of its evaluation in terms of the basic emotions is too high), then the aggregated value $\mu_{K_{ij}}$ has very difficult interpretation and needs to be modified to account for this large variability. First, the *mean basic-emotion value* $\mu_{E_{jl}}$ is computed for a Kansei tag KT_j using (2) and then the respective variance V_{jl} is computed using (3).

$$\mu_{E_{jl}} = \frac{\sum_{k=1}^{p} x_{E_{jlk}}}{p} \tag{2}$$

$$V_{jl} = \sum_{k=1}^{p} \frac{(x_{E_{jlk}} - \mu_{E_{jl}})^2}{p} \tag{3}$$

Finally a measure of the total variability V_{KT_j} for each Kansei tag KT_j is calculated using (4). Note, that Huang et al. (2012) compute the total variance as a square root of our V_{KT_j}.

$$V_{KT_j} = \sum_{l=1}^{q} V_{jl} \tag{4}$$

Once the total variability of each Kansei tag is known, it can be interpreted and used in several ways. Generally the higher the value of V_{KT_j}, the larger the inconsistency of understanding and interpreting the j-th Kansei tag is among the evaluators (or to be more specific the more variable emotional associations are triggered by the Kansei tag in the evaluators). One possible way of using these total variance values would be to discard those Kansei tags with the total variability larger than a given threshold, since their aggregated value is almost impossible to interpret correctly. Huang et al. (2012), however, suggest to modify the values of $\mu_{K_{ij}}$ based on the values of V_{KT_j} as described in the following step.

6. *Calculating the adjusted mean values of the Kansei tags* using (5), where F is a linear or nonlinear mapping function. Huang et al. (2012, pp. 574-575)

suggest several possible mapping functions, yet their rationale is not very clear (note that (5) actually moves the average based on the variability).

$$\mu_{K_{ij}}^{adj} = \mu_{K\,ij} - F(V_{KT\,j}) \tag{5}$$

In fact the evaluator, i.e. the decision-maker responsible for the final decision, might not know how to choose one of these functions, since no good practices or lists of "mapping functions of choice for particular problems" exist. Even though the authors claim that the actual choice of the mapping function does not have an effect on the final outcome, we consider this step to be a questionable one and as such it is not supported or further commented by us in this chapter. We, however, acknowledge the value of the information carried in V_{KT_j}.

7. *Presenting the results and drawing conclusions* - final classification or evaluation of the alternatives. Let us for now consider that the adjustment represented by (5) is done in a reasonable and meaningful way. Then either threshold values can be specified to see whether an object (alternative) should be classified under a specific Kansei tag, or simply a profile of Kansei mean values can be provided for each alternative. In the latter case an "ideal" or "desired" evaluation in terms of Kansei values can be specified and the alternative closest to this ideal can be chosen.

Although the method suggested by Huang et al. (2012) provides means for the assessment of consistency of understanding (or feeling about) the Kansei tags by the group of evaluators, it can still be further developed. First of all the modification of mean Kansei values based on their variance is not well justified and might not even be necessary. Second the scale relevance issue (i.e. the possibility that some evaluators might consider some of the Kansei tags less than fully appropriate for the evaluation purposes; or that the emotional assessment might be difficult for the evaluators because they might not be entirely confident about their answers in this step, see e.g. Heise (1969)) as well as the unclear interpretation of values close to the middle one (Kulas and Stachowski 2009) are not dealt with. Hence there are still several possible sources of uncertainty that are not accounted for.

In the next section we therefore suggest a modified data collection procedure in line with e.g. Stoklasa, Talášek and Stoklasová (2016), Stoklasa, Talášek and Stoklasová (2017) and Stoklasa, Talášek and Stoklasová (2018) which can reflect the lower perceived scale relevance and also lower confidence of the evaluators with their answer. We then adopt the emotional-assessment variance perspective and suggest a measure of inconsistency of the perceptions of Kansei tags and its possible use in the evaluation process.

Interval-valued generalization of the basic-emotion based semantic differential method

A Semantic differential scales are a popular tool for data acquisition, mainly due to their simplicity. Unfortunately, this simplicity comes with a price. During the 60 years from the introduction of semantic differential by Osgood et al. (1957), there have been several studies published concerning the problems possibly associated with the use of the bipolar semantic differential scales (both discrete and continuous). The main objections were directed towards

- the inability of the original method to reflect lower scale relevance (i.e. the impossibility of expressing perceived partial or complete irrelevance of the scale for the purpose of evaluation by the evaluators)
- concept-scale interactions (Heise 1969) - i.e. the need of tailoring the semantic differential scale for each purpose/study

- the impossibility of expressing ambivalent attitudes (Kaplan 1972) - note that a single value is required from the evaluator on each scale in the standard version of the semantic differential method.
- and also the problematic interpretability of middle answers as stressed by Kulas and Stachowski (2009) - it is virtually impossible to know, whether a middle value of the scale provided by the evaluator should be interpreted as a "neutral answer", an answer indicating the irrelevance of the scale for the given purpose or the fact that the evaluator does not understand the anchoring linguistic labels well enough in the given context)

**** Figure 3 about here ****

Recently a solution to many of these issues has been suggested in Stoklasa, Talášek and Stoklasová 2016; Stoklasa, Talášek and Stoklasová 2017; Stoklasa, Talášek and Stoklasová 2018 by the enrichment of the data gathering procedure and by a transition to interval-valued answers. This way the uncertainty stemming from lower perceived relevance of the scales and lower confidence of the answers does no longer remain hidden, but is directly transformed into a multi-valued answer. The difference in the data gathering procedure with respect to the original semantic differential method lies in the administration of a second scale with each semantic differential one. This scale is represented by a $[0\%, 100\%]$ universe and is used to obtain the information of the relevance of the scale for the given purpose as perceived by the decision-maker (or it can also be framed as confidence with the answer etc.). The expressed decrease in relevance or confidence is then proportionally transformed into an interval on the original bipolar-adjective semantic differential scale. Figure 3 summarizes the generalized data gathering procedure that would in this case replace the one discussed in the step 4 of the method by Huang et al. (2012) and depicted in Figure 1. Note, that the perceived Kansei-tag relevance $y_{K_{ijk}}$ expressed by the evaluator on the relevance scales r_j is transformed into the values $w_{K_{ijk}} \in [0,2r]$ using (6), $i = 1, \dots, n$, $j = 1, \dots, m$ and $k = 1, \dots, p$. These values represent a part of the Kansei-tag universe proportional in size to the perceived irrelevance of the Kansei tag for the purpose of the evaluation.

$$w_{K_{ijk}} = 2r - y_{K_{ijk}} \tag{6}$$

Based on these values, the resulting interval-valued evaluation $\left[x^L_{K_{ijk}}, x^R_{K_{ijk}}\right]$ is computed. The procedure first checks if an uncertainty interval of the width $w_{K_{ijk}}$ can be defined symmetrically around $x_{K_{ijk}}$ and still fit into the $[-r, r]$ interval (i.e. if $\left[x_{K_{ijk}} - \frac{w_{K_{ijk}}}{2}, x_{K_{ijk}} + \frac{w_{K_{ijk}}}{2}\right] \subseteq [-r, r]$). If this is possible, then $\left[x^L_{K_{ijk}}, x^R_{K_{ijk}}\right]$ is defined symmetrically around $x_{K_{ijk}}$, otherwise the uncertainty interval is shifted in such a way that it remains a subset of $[-r, r]$ and retains its width $w_{K_{ijk}}$. This is summarized in formula (7).

$$\left[x^L_{K_{ijk}}, x^R_{K_{ijk}}\right] = \begin{cases} \left[-r, -r + w_{K_{ijk}}\right] & \text{for } \left(x_{K_{ijk}} - \frac{w_{K_{ijk}}}{2}\right) < -r, \\ \left[r - w_{K_{ijk}}, r\right] & \text{for } \left(x_{K_{ijk}} + \frac{w_{K_{ijk}}}{2}\right) > r, \text{and} \\ \left[x_{K_{ijk}} - \frac{w_{K_{ijk}}}{2}, x_{K_{ijk}} + \frac{w_{K_{ijk}}}{2}\right] & \text{otherwise} \end{cases} \tag{7}$$

Analogously, the interval-valued assessments of the Kansei tags with respect to the basic-emotions $\left[x^L_{E_{jlk}}, x^R_{E_{jlk}}\right]$ can be computed using formula (8), where $w_{E_{jlk}}$ is computed using (9), and the uncertainty stems from a lower confidence of the answer concerning a basic emotion BE_l expressed on the scale ca_l represented by the $[0\%, 100\%]$ interval. An example of the data input form along with the necessary notation is summarized in Figure 4.

$$\left[x^L_{E_{jlk}}, x^R_{E_{jlk}}\right] = \begin{cases} \left[-d, -d + w_{E_{jlk}}\right] & \text{for } \left(x_{E_{jlk}} - \frac{w_{E_{jlk}}}{2}\right) < -d, \\ \left[d - w_{E_{jlk}}, d\right] & \text{for } \left(x_{E_{jlk}} + \frac{w_{E_{jlk}}}{2}\right) > d, \text{and} \\ \left[x_{E_{jlk}} - \frac{w_{E_{jlk}}}{2}, x_{E_{jlk}} + \frac{w_{E_{jlk}}}{2}\right] & \text{otherwise} \end{cases} \tag{8}$$

$$w_{E_{jkl}} = 2d - y_{E_{jlk}} \tag{9}$$

**** Figure 4 about here ****

Let us now consider we apply the extended data gathering procedure as summarized in Figures 3 and 4, i.e. that we obtain the interval values $\left[x^L_{K_{ijk}}, x^R_{K_{ijk}}\right]$ instead of $x_{K_{ijk}}$ and $\left[x^L_{E_{jlk}}, x^R_{E_{jlk}}\right]$ instead of $x_{E_{jlk}}$ for all $i = 1, \dots, n$, $j = 1, \dots, m$, $k = 1, \dots, p$ and $l = 1, \dots, q$. Note, that in any case the original crisp values expressed on the Kansei tag scales always lie in the uncertainty intervals, i.e. $x_{K_{ijk}} \in \left[x^L_{K_{ijk}}, x^R_{K_{ijk}}\right]$. The same holds analogously for the basic-emotion assessment of the Kansei tags, i.e. $x_{E_{jlk}} \in \left[x^L_{E_{jlk}}, x^R_{E_{jlk}}\right]$. The proposed modification is so far a direct generalization of the original method, since if there is no uncertainty, i.e. when $w_{K_{ijk}} = 0$ for some i, j and k, we get the interval-valued evaluation computed using (7) in the form $\left[x^L_{K_{ijk}}, x^R_{K_{ijk}}\right] = \left[x_{K_{ijk}}, x_{K_{ijk}}\right]$, which is in fact nothing else than an interval representation of the real number $x_{K_{ijk}}$. The same holds for (8) and the assessment of Kansei tags with respect to the basic-emotions. The *interval-valued basic-emotion based semantic differential method* can now be summarized in the following steps:

1. *Selection of the Kansei adjectives* for the purpose of the evaluation and their grouping into *m Kansei tags*, specification the *set of q basic-emotions*. There is no difference in this step with respect to the method proposed in Huang et al. (2012). We will again consider all the Kansei tags to be represented by continuous universes $[-r, r]$, where $r > 0$ and the basic emotions to be represented by continuous universes of a possibly different length, denoted by $[-d, d]$, $d > 0$.
2. *Selection of the survey participants*. Again, no change with respect to Huang et al. (2012).

3. *Evaluation of the alternatives with respect to the Kansei tags*. The enhanced questionnaires depicted in Figure 3 will be used. Although these questionnaires double the number of answers needed from the evaluators, the information concerning the relevance of the Kansei tag is not a difficult one to provide. In fact a 100% relevance can be considered to be a default value and only if the perceived relevance is lower, an input from the evaluator specifying how low it is would be required. The evaluations of the alternative a_i with respect to the Kansei tag KT_j by the evaluator k are now represented by $\left[x_{K_{ijk}}^{L}, x_{K_{ijk}}^{R}\right] \subseteq [-r, r], i = 1, \dots, n, j = 1, \dots, m$and $k = 1, \dots, p$.
4. *Assessment of the Kansei tags in terms of their emotional associations*. Again an enhanced questionnaire (see Figure 4) will be used to obtain inputs for this purpose. The intervals representing the assessment of the Kansei tag KT_j by the evaluator k with respect to the basic-emotion BE_l are thus obtained in the form of $\left[x_{E_{jlk}}^{L}, x_{E_{jlk}}^{R}\right] \subseteq [-d, d], j = 1, \dots, m, k = 1, \dots, p$ and $l = 1, \dots, q$.
5. *Calculation step - determination of the Kansei-tag means and assessment of the consistence of understanding of the Kansei tag by the evaluators.* The Kansei tag mean values, represented again by intervals, are now calculated by (10).

$$\mu_{K_{ij}}^{I} = \frac{\sum_{k=1}^{p}\left[x_{K_{ijk}}^{L}, x_{K_{ijk}}^{R}\right]}{p} = \left[\sum_{k=1}^{p} \frac{x_{K_{ijk}}^{L}}{p}, \sum_{k=1}^{p} \frac{x_{K_{ijk}}^{R}}{p}\right] \qquad (10)$$

This value $\mu_{K_{ij}}^{I}$ is supposed to represent the overall group evaluation of the alternative i with respect to the Kansei tag KT_j. It can, however, be properly interpreted only if the understanding of the Kansei tags was identical (or very similar) for all the evaluators. If some of the evaluators have different emotional associations with the Kansei tags than the others, the aggregated value $\mu^{\mathrm{I}}_{\mathrm{K_{ij}}}$ might be difficult to interpret. In line with Huang et al. (2012), we will therefore investigate the consistency of the emotional associations triggered by the Kansei tags in the group of evaluators. First we calculate the *mean basic-emotion value* $\mu^{\mathrm{I}}_{\mathrm{E_{jl}}}$ for each Kansei tag KT_j and each basic-emotion BE_l using (11).

$$\mu^{I}_{E_{jl}} = \frac{\sum_{k=1}^{p}\left[x^{L}_{E_{jlk}}, x^{R}_{E_{jlk}}\right]}{p} = \left[\sum_{k=1}^{p} \frac{x^{L}_{E_{jlk}}}{p}, \sum_{k=1}^{p} \frac{x^{R}_{E_{jlk}}}{p}\right] \quad (11)$$

Now we need to assess to what extent the intervals $\mu^{I}_{E_{jl}}$ differ among the evaluators. To do so, we will now apply the concept of strong consensus in the BE_{jl} dimension as introduced in Stoklasa, Talášek and Stoklasová (2018). A set of p interval-valued evaluations $\{I_1, \dots, I_p\}$ with their associated crisp values $x_{I_k} \in I_k$, for all $k = 1, \dots, p$, is considered to represent a strong consensus in the given evaluation dimension if and only if $I_1 \cap \dots \cap I_p \neq \emptyset$ and $x_{I_k} \in (I_1 \cap \dots \cap I_p)$ for all $k = 1, \dots, p$. A strong consensus is thus present if there exists an interval of values on which all the evaluators agree and that comprises all the crisp evaluations. Having introduced this concept, we now define a measure of inconsistency (variability) of the emotional assessment of the Kansei tag KT_j with respect to the basic-emotion BE_l as V^{I}_{jl}, which is computed using (12), where $CI_{jl} = \bigcap_{k=1}^{p} [x^{L}_{E_{jlk}}, x^{R}_{E_{jlk}}] = [CI^{L}_{jl}, CI^{R}_{jl}]$, $MI_{jl} = [\min_{k} x^{R}_{E_{jlk}}, \max_{k} x^{L}_{E_{jlk}}]$ and $d(x_{K_{ijk}}, CI_{jl})$ is defined by (13).

$$V^{I}_{jl} = \begin{cases} 0 & \text{if } CI_{jl} \neq \emptyset \text{ and } x_{E_{jlk}} \in CI_{jl}\ \forall k, \\ \dfrac{\sum_{k=1}^{p} d(x_{E_{jlk}}, CI_{jl})}{p} & \text{if } CI_{jl} \neq \emptyset \text{ and } x_{E_{jlk}} \notin CI_{jl} \text{ for some } k, \\ \left|MI_{jl}\right| + \dfrac{\sum_{k=1}^{p} d(x_{E_{jlk}}, MI_{jl})}{p} & \text{if } CI_{jl} = \emptyset \end{cases} \quad (12)$$

$$d\left(x_{E_{jlk}}, CI_{jl}\right) = \begin{cases} \left\| \left[x_{E_{jlk}}, CI_{jl}^{L}\right] \right\| & \text{if } x_{E_{jlk}} < CI_{jl}^{L}, \\ \left\| \left[CI_{jl}^{R}, x_{E_{jlk}}\right] \right\| & \text{if } CI_{jl}^{R} < x_{E_{jlk}}, \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

The idea behind (12) is, that if a strong consensus of the interval evaluations of the Kansei tag KT_j with respect to the basic-emotion BE_l exists, then it is possible to find a consensual evaluation and as such the evaluations can be considered consistent (hence the zero value of V_{jl}^{I} in this case). If there is no strong consensus, but if the interval CI_{jl} is nonempty (i.e. if at least *weak consensus* exists), the variability is calculated as the average of the distances of the crisp evaluations from this interval defined by (13). If there is even no weak consensus, we define the smallest interval that would represent a weak consensus, MI_{jl}, calculate the average distance of the crisp evaluations from this interval (again by (13)) and add it to the length of MI_{jl} to obtain the value of the variability.

Now that we have a measure of the variability of the assessment of each Kansei tag with respect to a single basic-emotion (expressed in fact as the measure of dissensus of the evaluations), we need to define an overall variability measure for the Kansei tag across all the basic-emotions. Since only basic-emotions are considered (and not complex ones derived or composed from them), we can consider them to be independent and to constitute q dimensions of a emotional-assessment Cartesian space. In this space we can define the up-to-q dimensional area of variability VA_j as an up-to-q dimensional box with edges of the lengths V_{jl}^{I} by (14).

$$VA_j = \left[0, V_{j1}^{I}\right] \times \dots \times \left[0, V_{jq}^{I}\right] \tag{14}$$

The larger the area VA_j is, the more substantial change of evaluations would have to take place for a strong consensus to be reached. Note, that VA_j represents a up-to-q dimensional block in the basic-emotion space. A measure of its size could therefore be an applicable measure of the overall variability of the emotional assessment of KT_j by the evaluators. We suggest the length of the body diagonal as the overall variability measure $V^I_{KT_j}$. More specifically we use a normalized body diagonal length computed by (15), i.e. $V^I_{KT_j} \in [0,1]$.

$$V^I_{KT_j} = \frac{\sqrt{\left(V^I_{j1}\right)^2 + \left(V^I_{j2}\right)^2 + \cdots + \left(V^I_{jq}\right)^2}}{\sqrt{q(2d)^2}} \tag{15}$$

6. *Reflection of the variability of understanding of Kansei tags in terms of basic-emotions*. The variability of the emotional-interpretation of the Kansei tags and the uncertainty thus introduced in the evaluation model is stemming from a different source than the uncertainty defining the intervals $[x^L_{K_{ijk}}, x^R_{K_{ijk}}]$. It is therefore difficult to combine these two different pieces of information into one and to modify the Kansei-tag mean values based on their variability. There are, however, several methodologically safer ways of using the information concerning the variability of understanding of Kansei tags by the evaluators:
   - the simplest way being just presenting the variability of Kansei tags along with the interval-valued Kansei tag means in the evaluation profile. This way no information is lost or distorted. On the other hand it requires more competencies and skills from the decision-maker responsible for the final decision.
   - or a threshold for acceptable inconsistency can be specified and those Kansei tags that do not meet this minimum consistency requirement might be discarded from the evaluation. This way, however, we are

potentially loosing information and it may render a significant part of the data we have gathered useless, if many Kansei tags have higher overall variability.

- since $V^{I}_{KT_j} \in [0,1]$, these values can be used as weights of the Kansei tags in further aggregation of the results, or even as weights of fuzzy rules using the Kansei tag mean values for classification and/or interpretation purposes.

This way we have covered the step 7 of the original method as well.

The interval-valued version of the basic-emotion based semantic differential method suggested in this chapter can be used as a soft-counterpart to standard multi-expert multiple-criteria decision-making methods, offering both means for the assessment of less tangible criteria and also tools for consistency checking of the connotative meanings of the linguistic labels used in the model. Although the basic-emotion perspective might not be applicable in all problems, it constitutes a blueprint for analogous assessments of the connotative component of the meaning of linguistic terms used as anchors in the semantic differential method.

Conclusion

This chapter suggests a soft design-alternative evaluation methodology using the basic-emotion based semantic differential method by Huang et al. (2012) as its basis and utilizing the interval-valued extension of the semantic differential proposed by Stoklasa, Talášek and Stoklasová (2017) and the concepts of strong and weak consensus introduced by Stoklasa, Talášek and Stoklasová (2018). The combination of these approaches introduces two new possible sources of uncertainty in the method originally proposed by Huang et al. (2012) and offers means for dealing with the low scale relevance issue as well as with some other commonly identified drawbacks of semantic differential scales. It presents a generalization of the method by Huang et al., but does not perform the Kansei tag mean modification step. Instead, alternative uses of the variability of Kansei tags are suggested. The proposed emotion-based linguistic multi-expert evaluation method constitutes a tool for the evaluation of less tangible (and difficult to measure) aspects of the alternatives in multi-expert evaluation problems not restricted to the area of design (or consumer product) evaluation, but generally to every problem where qualitative criteria need to be reflected by a group of experts. As such it might be an interesting source of inspiration also for social sciences and humanities, i.e. areas of research dealing with difficult to measure concepts.

Acknowledgments

This research was partially supported by the grant IGA_FF_2018_002 of the internal grant agency of Palacký University Olomouc.

References

Brill, E. Downey, Shoou-Yuh Chang, and Lewis D. Hopkins. 1982. "Modeling to Generate Alternatives: The HSJ Approach and an Illustration Using a Problem in Land Use Planning." *Management Science* 28 (3): 221–35. doi:10.1287/mnsc.28.3.221.

Fiss, P. C. 2011. "Building Better Causal Theories: A Fuzzy Set Approach to Typologies in Organization Research." *Academy of Management Journal* 54 (2): 393–420. doi:10.5465/AMJ.2011.60263120.

Heise, D. R. 1969. "Some Methodological Issues in Semantic Differential Research." *Psychological Bulletin* 72 (6): 406–22. doi:10.1037/h0028448.

Hsu, S. H., M. C. Chuang, and C. C. Chang. 2000. "A Semantic Differential Study of Designers' and Users' Product Form Perception." *International Journal of Industrial Ergonomics* 25 (4): 375–91. doi:10.1016/S0169-8141(99)00026-8.

Huang, Y., C. H. Chen, and L. P. Khoo. 2012. "Products Classification in Emotional Design Using a Basic-Emotion Based Semantic Differential Method." *International Journal of Industrial Ergonomics* 42 (6). Elsevier Ltd: 569–80. doi:10.1016/j.ergon.2012.09.002.

Jiao, J., Y. Zhang, and M. Helander. 2006. "A Kansei Mining System for Affective Design." *Expert Systems with Applications* 30 (4): 658–73. doi:10.1016/j.eswa.2005.07.020.

Jindo, T., K. Hirasago, and M. Nagamachi. 1995. "Ergonomics Development of a Design Support System for Office Chairs Using 3-D Graphics." *International Journal of Industrial Ergonomics* 15: 49–62.

Kaplan, K. J. 1972. "On the Ambivalence-Indifference Problem in Attitude Theory and Measurement: A Suggested Modification of the Semantic Differential Technique." *Psychological Bulletin* 77 (5): 361–72. doi:10.1037/h0032590.

Kobayashi, Ma., and T. Kinumura. 2017. "A Method of Gathering, Selecting and Hierarchizing Kansei Words for a Hierarchized Kansei Model." *Computer-Aided Design and Applications* 14 (4). Taylor & Francis: 464–71. doi:10.1080/16864360.2016.1257188.

Kulas, J. T., and A. A. Stachowski. 2009. "Middle Category Endorsement in Odd-Numbered Likert Response Scales: Associated Item Characteristics, Cognitive Demands, and Preferred Meanings." *Journal of Research in Personality* 43 (3). Elsevier Inc.: 489–93. doi:10.1016/j.jrp.2008.12.005.

Likert, R. 1932. “A Technique for the Measurement of Attitudes.” *Archives of Psychology* 22 (140): 5–55.

Nagamachi, M. 1995. “Kansei Engineering: A New Consumer-Oriented Technology for Product Development.” *International Journal of Industrial Ergonomics* 15: 3–11.

Osgood, C. E. 1964. “Semantic Differential Technique in the Comparative Study of Cultures.” *American Anthropologist* 66 (3): 171–200. doi:10.1515/9783110215687.109.

Osgood, C. E., G. J. Suci, and P. H. Tannenbaum. 1957. *The Measurement of Meaning*. Chicago: University of Illinois Press.

Stoklasa, J. 2014. *Linguistic Models for Decision Support*. Lappeenranta: Lappeenranta University of Technology.

Stoklasa, J., and T. Talášek. 2016. “On the Use of Linguistic Labels in AHP: Calibration, Consistency and Related Issues.” In *Proceedings of the 34th International Conference on Mathematical Methods in Economics,* 785–90. Liberec: Technical University of Liberec.

Stoklasa, J., and T. Talášek. 2015. “Linguistic Modelling in Economics and Management Practice - Some Open Issues.” In *Proceedings of the International Scientific Conference Knowledge for Market Use 2015*, 959–69.

Stoklasa, J., T. Talášek, J. Kubátová, and K. Seitlová. 2017. “Likert Scales in Group Multiple-Criteria Evaluation.” *Journal of Multiple-Valued Logic and Soft Computing* 29 (5): 425–40.

Stoklasa, J., T. Talášek, and P. Luukka. 2018. “Fuzzified Likert Scales in Group Multiple-Criteria Evaluation.” In *Soft Computing Applications for Group Decision-Making and Consensus Modeling*, edited by M. Collan and J. Kacprzyk, 357:165–85. doi:10.1007/978-3-319-60207-3_11.

Stoklasa, J., T. Talášek, and J. Musilová. 2014. “Fuzzy Approach - a New Chapter in the Methodology of Psychology?” *Human Affairs* 24 (2): 189–203. doi:10.2478/s13374-014-0219-8.

Stoklasa, J., T. Talášek, and J. Stoklasová. 2018. “Attitude-Based Multi-Expert Evaluation of Design.” (unpublished).

Stoklasa, J., T. Talášek, and J. Stoklasová. 2017. “Reviving the Semantic Differential - Scale Relevance and Uncertainty Entering the Semantic Space.” (submitted).

Stoklasa, J., T. Talášek, and J. Stoklasová. 2016. “Semantic Differential and Linguistic Approximation - Identification of a Possible Common Ground for Research in Social Sciences.” In *Proceedings of the International Scientific Conference Knowledge for Market Use 2016*, 495–501. Olomouc: Societas Scientiarum Olomucensis II.

Stoklasa, J., J. Talašová, and P. Holeček. 2011. “Academic Staff Performance Evaluation - Variants of Models.” *Acta Polytechnica Hungarica* 8 (3): 91–111.

Talašová, J., J. Stoklasa, and P. Holeček. 2014. “HR Management through Linguistic Fuzzy Rule Bases - a Versatile and Safe Tool?” In *Proceedings of the 32nd International Conference on Mathematical Methods in Economics*, 1027–32.

Xu, R., and D. C. Wunsch. 2009. *Clustering*. Hoboken: John Wiley & Sons.

Yeomans, J. S. 2011. “Efficient Generation of Alternative Perspectives in Public Environmental Policy Formulation: Applying Co-Evolutionary Simulation-Optimization to Municipal Solid Waste Management.” *Central European Journal of Operations Research* 19 (4): 391–413. doi:10.1007/s10100-011-0190-y.

Yeomans, J. S., and Y. Gunalay. 2011. “An Efficient Modelling to Generate Alternatives Approach for Addressing Unmodelled Issues and Objectives in Public Environmental Planning.” *Asian Journal of Information Technology* 10 (3): 122–28. doi:10.3923/ajit.2011.122.128.

Zadeh, L. A. 1975. “The Concept of a Linguistic Variable and Its Application to Approximate Reasoning-I.” *Information Sciences* 8 (3): 199–249. doi:10.1016/0020-0255(75)90036-5.

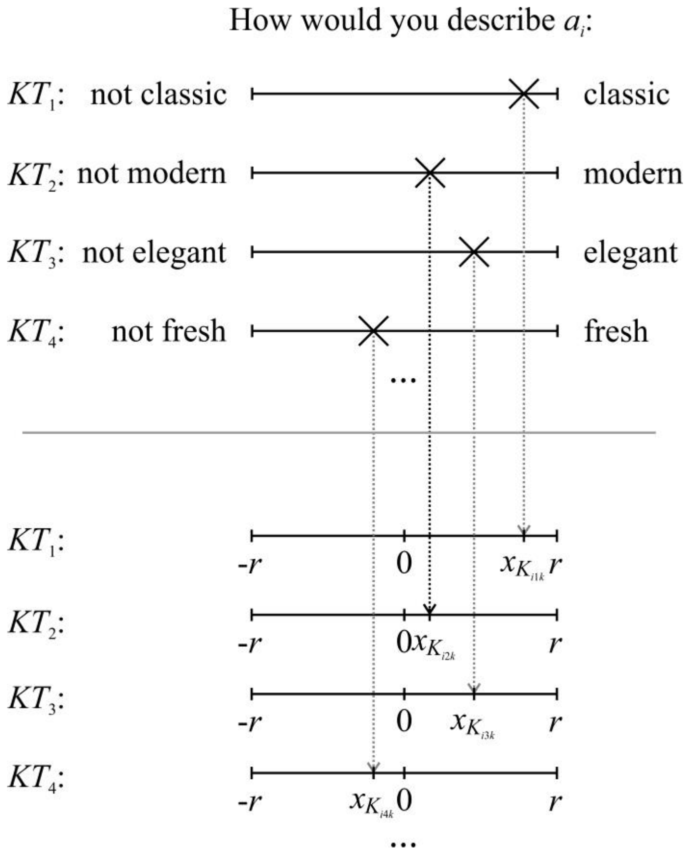

Figure 1. Evaluation form for the alternative a_i by the evaluator k with respect to the given Kansei tags, $i = 1, \dots, n$ and $k = 1, \dots, p$. The upper part represents the tool as seen and used by the evaluator, the lower part represents the conversion of the inputs into model variables' values $x_{K_{ijk}} \in [-r, r]$, where $KT_j, j = 1, \dots, m$, represents the j-th Kansei tag.

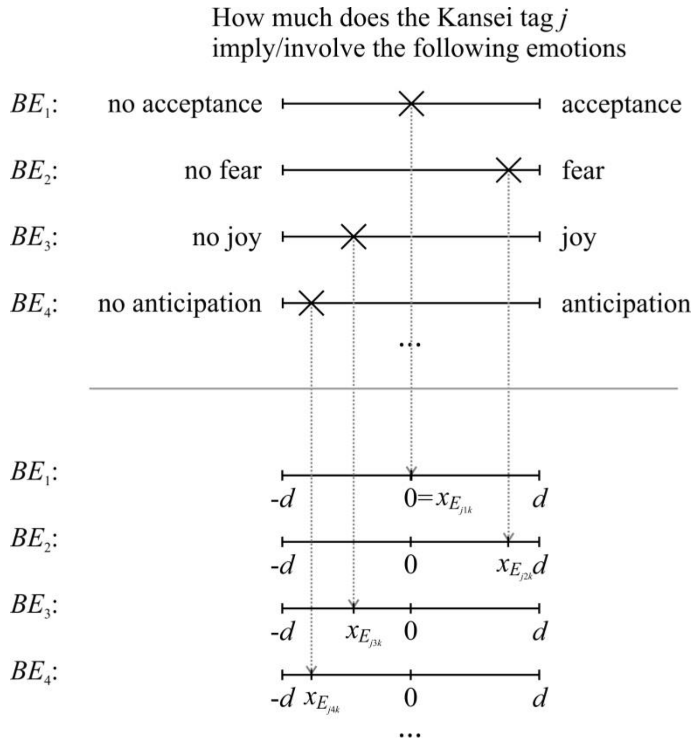

Figure 2. Assessment form for the Kansei tag j by the evaluator k with respect to the pre-specified basic-emotions, $j = 1, \dots, m$ and $k = 1, \dots, p$. The upper part represents the tool as seen and used by the evaluator, the lower part represents the conversion of the inputs into model variables' values $x_{E_{jlk}} \in [-d, d]$, where $BE_l, l = 1, \dots, q$, represents the l-th basic-emotion

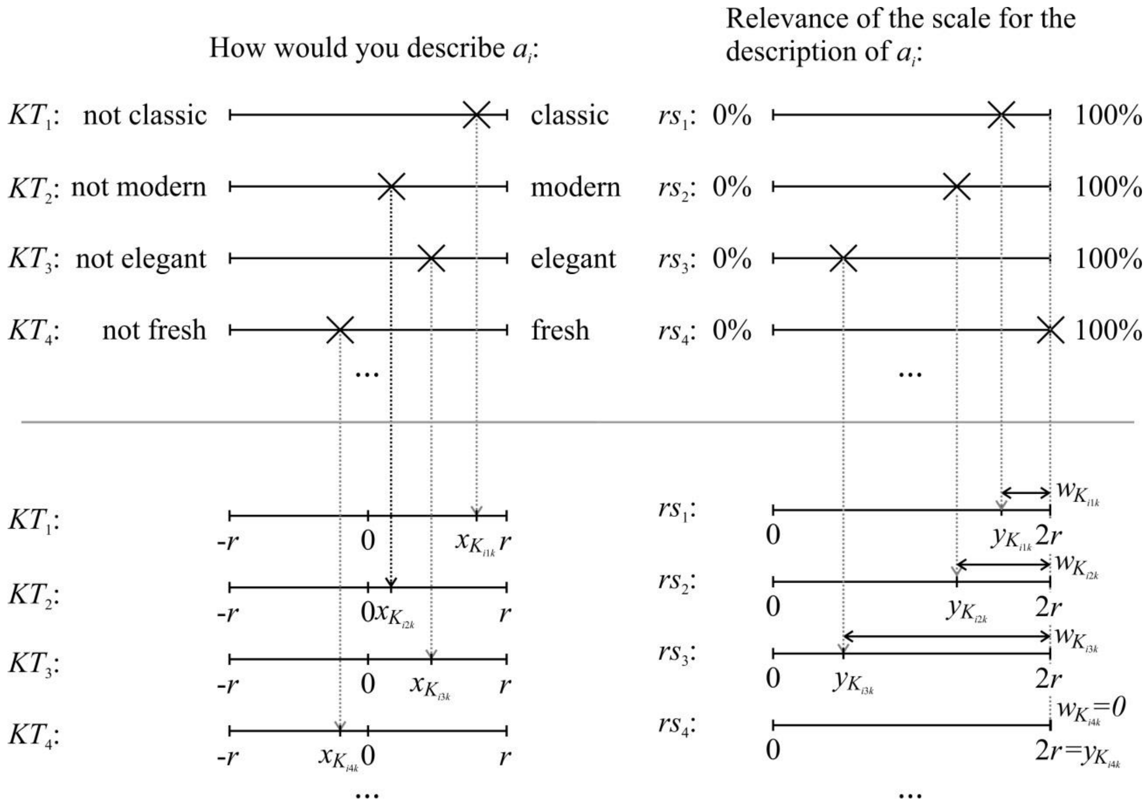

Figure 3. Evaluation form for the alternative a_i by the evaluator k with respect to the given Kansei tags, $i = 1, \ldots, n$ and $k = 1, \ldots, p$ - extended version inspired by Stoklasa, Talášek and Stoklasová (2016) and Stoklasa, Talášek and Stoklasová (2017). The upper part of the figure represents the tool as seen and used by the evaluator, the lower part represents the conversion of the inputs on the Kansei tag scales into model variables' values $x_{K_{ijk}} \in [-r, r]$, where KT_j, $j = 1, \ldots, m$, represents the j-th Kansei tag, and of the perceived scale relevance into uncertainty regions of the width $w_{K_{ijk}}$. The right part of the figure (titled "Relevance of the scale for the description of a_i:") denotes the addition with respect to the original semantic differential method.

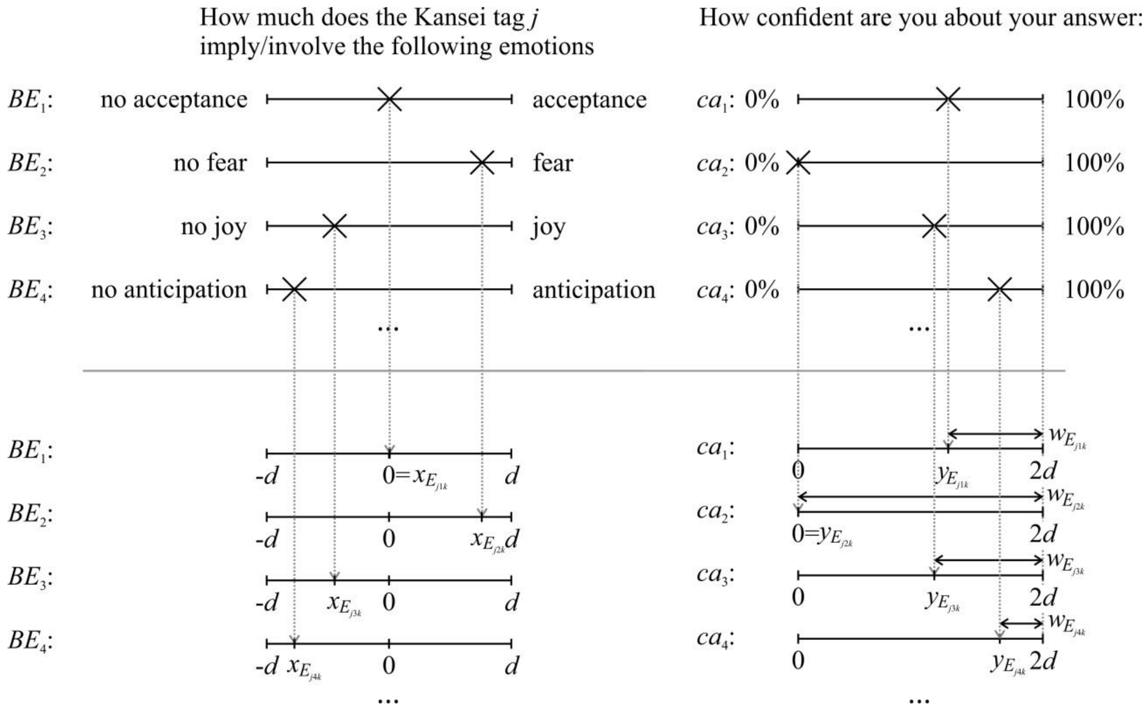

Figure 4. Assessment form for the Kansei tag j by the evaluator k with respect to the pre-specified basic-emotions, $j = 1, \dots, m$ and $k = 1, \dots, p$ - extended version inspired by Stoklasa, Talášek and Stoklasová (2016). The upper part of the figure represents the tool as seen and used by the evaluator, the lower part represents the conversion of the inputs on the Kansei tag scales into model variables' values $x_{E_{jlk}} \in [-d, d]$, where BE_l, $l = 1, \dots, q$, represents the l-th basic-emotion, and of the perceived confidence of the answer into uncertainty regions of the width $w_{E_{jlk}}$. The right part of the figure (titled "How confident are you with your answer:") denotes the addition with respect to the original semantic differential method.