Compensating Qualitative Rating Distortion of User Experience Evaluation Based on Prospect Theory

Min Chul Lee and Jaehyun Park

1 Department of Safety Engineering, Incheon National University, Academy-ro 119, Incheon 22012, Korea; LMC@inu.ac.kr
2 Department of Industrial and Management Engineering, Incheon National University, Academy-ro 119, Incheon 22012, Korea
* Correspondence: jaehpark@inu.ac.kr; Tel.: +82-32-835-8867

Received: 25 September 2019; Accepted: 26 November 2019; Published: 1 December 2019

Abstract: Psychophysical assessment may be affected by cognitive distortion. Although the theory was originally developed to revise decision making in uncertain situations, prospect theory can be applied to psychophysical measurements, which was verified in a previous preliminary study. Two case studies were used to validate the utilization of prospect theory in psychophysical measurements. Affective satisfaction dimensions were rated by participants for an experimental device using a 0–100 scale. Performance of affective satisfaction models increased with the application of prospect theory-based compensation. Hundreds of participants evaluated the user value of their own devices via an online questionnaire. Although model fit performance increased slightly with transformed data, more case studies are needed to investigate the utility of prospect theory on user value or on a range of target constructs. The application of prospect theory in various situations of psychophysical measurement can be expected to improve and compensate for measurement results.

Keywords: prospect theory; psychological measurement; subjective assessment; 0–100 scale; affective satisfaction; user value

1. Introduction

Psychophysical measurement has been widely used and will continue to be used in various fields, including the field of human-computer interaction (HCI) and user experience (UX). In particular, from the perspective of abstract constructs such as affect and user value, which are supposed to be influential factors of UX, researchers usually have no choice but to use psychophysical measurements.

Even simple, quick, and dirty assessments require psychophysical measurements, such as the Likert and Borg scales [1,2], although methodologies for quantitative assessment of UX have been studied increasingly in academic and practical fields [3]. For example, the system usability scale (SUS) and software usability measurement inventory (SUMI) depend on psychophysical measurements, such as 1–3 and 1–5 scales [4–7]. The psychophysical measurements that have been widely used have problems in two aspects. First, although they use the interval scale or ratio scale, both options have extreme values (e.g., 0 and 100). At the extremes, the scale is bound to be distorted. Second, when linguistic anchors exist, for example, even if 50 points are given at a neutral level on the 0–100 scale, people tend to be based on slightly higher scores [8]. Although it may not be a problem to allow some level distortion and analyze the data in this way, it goes without saying that it may be better to analyze the data after correction.

A previous study, designed to be preliminary to this study, intended to compensate for limitations in psychophysical measurement by using prospect theory-based transformation [8]. The theory was based on human behavior in uncertain situations including diminishing sensitivity and loss aversion.
It is assumed to be applied to psychophysical measurements in this study. An experiment with typical target selection tasks, which have been widely used to prove the usability of touch interfaces, revealed a certain possibility to confirm prospect theory in usability inspection. Consequently, if the reference point is 75 on a 0–100 scale, the performance of the transformation seemed to be the highest. That means that when a participant is asked to rate something with a 0–100 scale, there is a tendency for the participant to mark a reference point as 75 in his or her mind despite being given 50 as a medium point. At least in this specific context, it has been confirmed, but in other contexts, a verification seems to be needed.

In order to ensure the application of prospect theory on psychophysical measurement, two case studies were carried out and analyzed in the current study. The two studies target two constructs (i.e., affect and user value) that have been regarded as influential factors of UX [9]. The results of this study can be used to enlarge the utility of psychophysical measurement, which has been widely used in various fields.

2. Prospect Theory Review

2.1. Decision Making Situations with Prospect Theory

Many researchers have focused on utility, which can be defined as a measure of preferences for a set of products and services in the field of economics [10,11]. Among them, prospect theory postulates decision making in uncertain situations [12]. Prospect theory reveals individuals’ decision-making behavior: they tend to (1) set a precise reference point expected for a target, (2) have different strategies based on expected gains and losses (i.e., having more value and less value than the reference point, respectively). The value function can also be called an S-shaped function because of the concave characteristic for gains and the convex characteristic for losses.

Among various studies of utility functions in uncertain situations, prospect theory seems to suit product or service evaluation. On one hand, choices in risky situations were analyzed and applied in some prominent studies, such as game theory [13]. These studies also dealt with risky and uncertain situations with forced consideration of others’ decisions. On the other hand, prospect theory has focused on individuals’ behavior when they are required to evaluate or choose an alternative. Note that product or service evaluation is generally designed to target the quality of a product or service per se.

2.2. Affective Satisfaction with Prospect Theory

Affect has been widely studied in recent decades. Since Kansei engineering was introduced in an effort to quantify users’ feelings and demands regarding products or services [14], various researchers have attempted to define affect or affective satisfaction. In the field of psychology, core affect was defined as “a neurophysiological state consciously accessible as the simplest non-reflective feelings evident in moods and emotions”. [15]. A simpler concept of affect was also suggested as an appraisal influenced by product appearance [16,17]. One of the previous studies identified six affective dimensions for mobile phones and services (Table 1) [9]. Note that although this study adopted the terms affect and affective satisfaction, other terms such as perceived affectiveness can also be used in similar contexts.

The appraisement process of affect corresponds to decision making. Regardless of the field, researchers postulate at least two states, such as neutral and affected states [15,18]. An individual receives information from multimodal senses, such as sight, hearing, and touch. If the signals are the same as expected, there would not be much change in the neutral state. However, if multimodal information is not as expected, individuals modify their perspective, giving rise to a new affected state. Both prospect and affect theory include a reference point for judgment or appraisal.
Table 1. Definitions of dimensions of the affect concept [9].

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color</td>
<td>Degree to which the color used in a product/service is likable, vivid, or colorful</td>
</tr>
<tr>
<td>Delicacy</td>
<td>Degree to which a product/service is elaborate, or finely and skillfully made</td>
</tr>
<tr>
<td>Texture</td>
<td>Degree to which a product’s texture or touch appeals to the users</td>
</tr>
<tr>
<td>Luxuriousness</td>
<td>Degree to which a product/service is luxurious or looks superior in quality and cost</td>
</tr>
<tr>
<td>Attractiveness</td>
<td>User’s perception that a product/service is pleasing, arousing, interesting, and attractive</td>
</tr>
<tr>
<td>Simplicity</td>
<td>The way a product/service looks and works is simple, plain, and uncomplicated</td>
</tr>
</tbody>
</table>

2.3. User Value with Prospect Theory

Value is one of the most important concepts across the ages. It has been widely discussed and studied in various academic fields such as sociology, psychology, marketing, and industrial design [19–24].

Parsons [22] described value as formed by a group sharing a common ideology or environment, while Rokeach [23] classified value into two types, including final states of existence and modes of behavior. These perspectives are still in effect in academia, but in the field of marketing, efforts have been made to set value limits for a certain product or service. Gutman [20] and Woodruff [24] have suggested the product or service-related value. From the perspective of engineering and user experience, Park and Han [21] defined user value as a “desirable states of existence or modes of behavior which are satisfied when using a certain product or service”. In addition, Park and his colleagues [9] identified five user value dimensions (Table 2).

Table 2. Definitions of dimensions of user value [9].

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-satisfaction</td>
<td>Degree to which a product/service gives the user satisfaction with himself or herself or their achievements</td>
</tr>
<tr>
<td>Pleasure</td>
<td>User’s feeling of being pleased or gratified by interacting with a product/service</td>
</tr>
<tr>
<td>Sociability</td>
<td>Degree to which a product/service satisfies the user’s desire of being sociable</td>
</tr>
<tr>
<td>Customer need</td>
<td>Degree to which functions or appearances of a product/service satisfy the user’s needs</td>
</tr>
<tr>
<td>Attachment</td>
<td>Ability for the user to attach subjective value to a product/service</td>
</tr>
</tbody>
</table>

Value can be assumed to be highly relevant for utility. The term “use-value” proposed by Marx [25] is similar to the term “user value” mentioned above. Marx defined use-value as “the intrinsic characteristics of a product that enable it to satisfy a human need or want”. Even if he had not attempted to quantify the concept of “use-value,” many researchers in the field of economics have studied the quantification of value or utility. Prospect theory may fit well with the concept of user value.

3. Case Study I: Affective Satisfaction

3.1. Design of Experiment

3.1.1. Participants

In advance, participants were verified with a rating ability test based on the cross-modality comparison method [26]. Each participant was asked to produce corresponding lines for randomly generated numbers and generate corresponding numbers for randomly produced lines. Using linear regression to investigate the ratio of the exponents of the power function between two modalities, the continued participation of each participant was confirmed. Note that the experiment was conducted according to the principles expressed in the Declaration of Helsinki.

An experiment was conducted with a total of 26 participants based on opportunity sampling. Gender groups were balanced. There were 13 males and females; their average age was 22.4 years.
old, with a range from 18 to 28 years old. Participants who did not have any previous experience with the experimental apparatus were recruited. Note that, when this case study was conducted, the experimental apparatus, an iPad, was not a popular device in our country and most people had no previous exposure to one.

3.1.2. Apparatus

As an experimental apparatus, an iPad was used to investigate affective satisfaction for a consumer electronic product. The experimental device was the initial iPad model by Apple. Moreover, it should be noted that the case study was not intended to confirm the degree of affective satisfaction for a certain product. Thus, control of the experimental condition (i.e., letting novice participants make ratings using the psychophysical scale) was regarded as more important than validating a certain product.

3.1.3. Measurements

The seven affective satisfaction dimensions were measured using modified magnitude estimation: delicacy, simplicity, texture, luxuriousness, color, attractiveness, and overall affect. The psychophysical rating method was implemented using a 0–100 scale. This unipolar scale had two endpoints, 0 and 100, with 50 as the anchor point on the scale described as “normal”.

The structure of affective satisfaction for mobile devices has been verified in several studies [16,27]. In particular, Cho and his colleagues used five affective dimensions, including delicacy, simplicity, luxuriousness, color, and attractiveness, which were the same as the dimensions used in this study, and revealed that those dimensions were valid measures since the average of the adjusted $R^2$ of the affective models was 0.91. With regard to the affective structure, the affective dimensions and overall affect were assumed to be highly correlated and capable of building a relationship model.

3.1.4. Tasks and Procedure

After the rating ability test mentioned above, participants were required to use the experimental devices with built-in applications (i.e., calendar, Safari, mail, settings, maps, videos, YouTube, photos, contacts, notes, iTunes, and the App store). Even if affect was defined as being related to product appearance, there was a slight possibility that product usage might influence the users. Detailed task guides were provided to participants since most users were a novice to tablet PCs and unfamiliar with iPads. For example, each participant was asked to add a new schedule in the calendar or to access a certain website with Safari. In addition, a session allowing participants to freely use the experimental device was carried out in order to provide them with diverse experiences. Consequently, participants were required to evaluate the affective dimensions and overall affect using a psychophysical rating method.

3.2. Results

3.2.1. Affective Satisfaction Model with Raw Data

The stepwise regression method was applied to build affective satisfaction models. Simple linear models that did not include any polynomials were considered. It should be noted that the performance of a simple linear model was similar to, or better than other variations, such as polynomial, conjunctive, and disjunctive models [3]. When applying a stepwise regression method, a variable whose variation inflation factor (VIF) was higher than 10 was eliminated from the models to minimize multi-collinearity [28].

As a result, an affective satisfaction model including color, delicacy, texture, luxuriousness, and simplicity was developed. The adjusted $R^2$ and the maximum VIF of the model were 0.865 and 2.441, respectively (Table 3). The performance of the model revealed that it would be reliable.
3.2.2. Affective Satisfaction Model with Transformed Data

The transformation was based on the cumulative prospect theory suggested by Tversky and Kahneman [29]. The function was reversed as expressed in Equation (1). Affective satisfaction scores for all dimensions, including overall affect, were transformed with the reversed function. Transformed scores were assumed to be close to the sensation magnitudes toward the experimental devices:

$$v^{-1}(\phi_i) = \begin{cases} \alpha + (\phi_i - \alpha)^{1/r} & \text{if } \phi_i \geq \alpha \\ \alpha - ((\alpha - \phi_i)/\lambda)^{1/r} & \text{if } \phi_i < \alpha \end{cases}$$ (1)

where $\phi_i$ is the psychophysical measurement corresponding to stimulus $i$, which was a property of the experimental device supposed to manipulate each affective satisfaction dimension in this case study, $v^{-1}$ is the reversed value function that assigns a psychophysical measurement to a sensation magnitude of the stimulus, $\alpha$ is the reference score on an estimation scale, $r$ is the exponent for sensitivity, and $\lambda$ is the coefficient of loss aversion. Note that cumulative prospect theory suggested by Tversky and Kahneman [29] estimated $r$ as 0.88 and $\lambda$ as 2.25.

In total, five affective satisfaction models were constructed for five reference scores, including 0, 25, 50, 75, and 100 on the 0–100 scale (Table 3). The average adjusted $R^2$ was 0.871, which was higher than the performance with untransformed data. The model performance seemed to be best in terms of goodness of fit when the reference score of 75 on the 0–100 scale was used. The degree of multi-collinearity was regarded as low, considering that the maximum VIFs for all models were lower than 10.

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Ref. Score</th>
<th>Adj. $R^2$</th>
<th>Max. VIF</th>
<th>Model Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>-</td>
<td>0.865</td>
<td>2.441</td>
<td>$-0.135 + 0.378 \times \text{Simplicity} + 0.204 \times \text{Color} + 0.134 \times \text{Luxuriousness} + 0.129 \times \text{Delicacy} + 0.158 \times \text{Luxuriousness} + 0.136 \times \text{Color} + 0.158 \times \text{Delicacy}$</td>
</tr>
<tr>
<td>P</td>
<td>0</td>
<td>0.868</td>
<td>2.494</td>
<td>$-0.310 + 0.369 \times \text{Simplicity} + 0.207 \times \text{Color} + 0.136 \times \text{Luxuriousness} + 0.134 \times \text{Delicacy}$</td>
</tr>
<tr>
<td>P</td>
<td>25</td>
<td>0.870</td>
<td>2.532</td>
<td>$-0.275 + 0.362 \times \text{Simplicity} + 0.208 \times \text{Color} + 0.139 \times \text{Luxuriousness} + 0.137 \times \text{Delicacy}$</td>
</tr>
<tr>
<td>P</td>
<td>50</td>
<td>0.878</td>
<td>2.830</td>
<td>$-1.267 + 0.332 \times \text{Simplicity} + 0.161 \times \text{Delicacy} + 0.153 \times \text{Luxuriousness} + 0.213 \times \text{Color} + 0.159 \times \text{Texture}$</td>
</tr>
<tr>
<td>P</td>
<td>75</td>
<td>0.881</td>
<td>3.043</td>
<td>$0.598 + 0.226 \times \text{Delicacy} + 0.264 \times \text{Luxuriousness} + 0.188 \times \text{Delicacy} + 0.206 \times \text{Color} + 0.117 \times \text{Texture}$</td>
</tr>
<tr>
<td>P</td>
<td>100</td>
<td>0.888</td>
<td>2.325</td>
<td>$-0.368 + 0.404 \times \text{Simplicity} + 0.196 \times \text{Color} + 0.134 \times \text{Luxuriousness} + 0.154 \times \text{Delicacy}$</td>
</tr>
</tbody>
</table>

$a$ R and P indicate raw and transformed data under prospect theory, respectively. $b$ Ref., Adj., and Max. indicate reference, adjusted, and maximum, respectively. $c$ Overall affect is the dependent variable of the model equation.

3.3. Discussion

3.3.1. Reliability of Applying Prospect Theory on Affect Evaluation

Prospect theory was revealed as consistent with the evaluation of the products’ affect. The results show that data transformed in accordance with the theory perform better than the untransformed data. In particular, when the reference points ranged from 0 to 75, the goodness of fit of the affective models increased. Of course, the increment from 0.01 to 0.02 in adjusted $R^2$ was not noteworthy; there would be an improvement.

The results of the transformation correspond to the results from the previous experiment with target selection tasks. In the experiment, the correlation coefficient between performance and subjective measures was highest if the reference point was set to 75. Similarly, in this case study, when the reference score was 75, the adjusted $R^2$, as a measure of goodness of fit, demonstrated the highest value.

To address another aspect, the same limitation that occurred in the previous study could occur in this case study as well. A measure that targets goodness of fit does not truly assess the degree of fit with prospect theory. Even if the adjusted $R^2$ increased through applying the value reverse function to
affective satisfaction scores, it cannot ensure that prospect theory is suitable for affect evaluation. In order to compensate for this limitation, more case studies need to be pragmatically analyzed.

3.3.2. Prospect Theory and Affective Satisfaction

Affect can be divided into several types. Types vary across researchers’ perspectives, but consensus exists that the dimensions of the subjectivities are distinct from one another [17,18]. However, it is clear that the evaluation of affective satisfaction is subjective. A situation where an evaluator is supposed to indicate his or her subjective feelings or image for a product can cause uncertainty. Similarly, the uncertainty provides the precondition of prospect theory and the subjectivity would provide the basis of prospect theory toward affect evaluation.

4. Case Study II: User Value

4.1. Design of Experiment

4.1.1. Type of Experiment with User Value Assessment

This case study was designed to assess user experiences, especially user value. Unlike other constructs, user value for products or services involves a high level of judgment and is formed over a long period [21]. Thus, in the case of user value, a laboratory experiment may not be an adequate assessment method because of temporal limitations. This was the reason for the case study conducted with an online survey.

4.1.2. Participants and Their Own Smartphones as Target Products

A total of 295 qualified participants consequently carried out the online experiment. Because participants in an online experiment tend to act insincerely, conservative criteria suggested by Cho et al. [27] to screen out insincere participants were modified slightly and applied. Participants who were not authenticated with duplicate IP addresses or who were not motivated were eliminated. For example, unmotivated participants were denominated if their average response time per questionnaire item was less than 5 s. The average age of qualified participants was 28.7; 247 of them were male and 48 were female. Note that the experiment was conducted according to the principles expressed in the Declaration of Helsinki.

As mentioned above, this case study intended to rate user value, which takes a reasonable amount of time to be evaluated because a user may not perceive or pursue value for a product or service in immediate interactions. In order to appropriately investigate user value, participants were allowed to evaluate their own smartphones. In turn, devices from Apple, Samsung Electronics, HTC, LG, Motorola, RIM, and Nokia were evaluated; most devices were from Apple and Samsung Electronics.

4.1.3. Measurements

A hierarchical structure of user experience suggested by Park and his colleagues [9] was used in this study. The structure was intended to situate the concept of user value into the user experience assessment model. All of the user value dimensions mentioned above were included in this structure: self-satisfaction, pleasure, sociability, customer need, and attachment.

A questionnaire was used whose items were verified in terms of validity and reliability in a previous study [9] (Table 4). The user value survey (UVS) includes 16 questionnaire items as follows. Each participant was asked online to rate questionnaire items using a 0–100 scale. The same verbal anchor used in the first case study was applied. Participants were informed that 0 and 100 denoted the endpoints and 50 represented “neutral”. Note that since the survey targeted smartphone users, the survey consisted of smartphone-related items such as manufacturer brands and applications.
Table 4. Questionnaire items of the user value survey (UVS).

<table>
<thead>
<tr>
<th>User Value Dimensions</th>
<th>Questionnaire Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attachment</td>
<td>I am loyal to my phone’s manufacturer brand</td>
</tr>
<tr>
<td></td>
<td>I want to recommend my smartphone to others</td>
</tr>
<tr>
<td></td>
<td>I feel I can trust the manufacturer and telecommunication company of my phone due to their brand awareness</td>
</tr>
<tr>
<td>Sociability</td>
<td>I feel better if I get in touch with my friends</td>
</tr>
<tr>
<td></td>
<td>I maintain a close relationship with my friends through calling and texting</td>
</tr>
<tr>
<td></td>
<td>I use SNS or messenger applications to contact friends when I have not heard from them in a long time</td>
</tr>
<tr>
<td></td>
<td>I feel a sense of intimacy when I exchange information about smartphones with others</td>
</tr>
<tr>
<td>Self-satisfaction</td>
<td>I feel superior to others in terms of practical use</td>
</tr>
<tr>
<td></td>
<td>I am satisfied with the fact that I have more applications than others</td>
</tr>
<tr>
<td></td>
<td>My phone is superior to others</td>
</tr>
<tr>
<td>Customer need</td>
<td>My phone seems to be useful because there are many things I can do with it</td>
</tr>
<tr>
<td></td>
<td>I am satisfied with that I can download applications that I want to have</td>
</tr>
<tr>
<td></td>
<td>I can obtain the information that I want through the mobile phone</td>
</tr>
<tr>
<td>Pleasure</td>
<td>I feel better if I listen to music or watch videos on my mobile phone</td>
</tr>
<tr>
<td></td>
<td>It is pleasant to discover a new function</td>
</tr>
<tr>
<td></td>
<td>I feel better if I buy a new phone or application</td>
</tr>
</tbody>
</table>

4.2. Results

4.2.1. Measurement Model with Raw Data

A confirmatory factor analysis was conducted with the raw data. The questionnaire items developed in a previous study [9] were used, as mentioned above. A measurement model consisting of 16 items was developed (Figure 1) and verified in terms of several fit indexes, including the absolute, incremental, and parsimonious fit indexes: goodness of fit index (GFI) and root mean square error of approximation (RMSEA) for the absolute fit indexes, as well as the normed fit index (NFI), comparative fit index (CFI), adjusted goodness-of-fit index (AGFI), and the ratio of chi square to its degrees of freedom (CMIN/DF) for the incremental and parsimonious fit indexes.

For this case study, the GFI, RMSEA, NFI, CFI, AGFI, and CMIN/DF were 0.91, 0.075, 0.852, 0.9, 0.87, and 2.664, respectively, indicating that the model was reliable. Note that acceptable model fit is generally indicated by values over 0.9 for the GFI, AGFI, and CFI, and an RMSEA value less than 0.06.
4.2.2. Measurement Model with Transformed Data

As mentioned in the first case study, scores for the questionnaire items were transformed using a prospect theory value function, which is based on the method described in Section 3.2.2 and the previous study [8]. Five score sets based on five reference scores (i.e., 0, 25, 50, 75, and 100) were investigated using confirmatory factor analysis. The results indicate that the model had the best fit when the reference score was 50 and the worst fit was shown at a reference score of 75. In general, the model fit data varied slightly (Table 5).
4.3. Discussion

4.3.1. Reliability of Applying Prospect Theory with User Value

First, the model fit summary was not improved as much as the performance of affective satisfaction when the data were transformed, even if the performance of the models increased to some degree after transformation. One of the absolute fit indexes, GFI, was 0.91 with the raw data, but 0.911 with the transformed data. Things were not much different for the other indexes. NFI, one of the incremental and parsimonious fit indexes, was 0.852 with the raw data and the highest NFI with the transformed data was 0.86.

In addition, when the reference score was 50, the performance of the model recorded the highest values. Unlike the case study for affective satisfaction and the previous experiment with target selection tasks, a reference score of 75 did not demonstrate the highest performance. Even if the target and methodology of the evaluation were different from those of the previous studies, the reference scores need to be investigated further. There is a high possibility that the optimal or practical range fell between 50 and 75. A future study could be carried out to investigate the optimal value or distribution of the reference scores.

4.3.2. Prospect Theory and User Value

As mentioned above, user value was originally based on the value that a product or service can provide [21,30]. However, the result of analyzing user value with prospect theory-based transformation did not turn out to have as high a performance as with affect evaluation. These results can be explained by two points.

First, the model fit summary does not reflect how measurement data were precisely assessed. This characteristic was also described in the affect evaluation section. On one hand, the reliability of the user value questionnaire items was slightly better than or almost the same as with the raw data, while the reliability of the affective satisfaction model with transformed data was certainly superior to the one using raw data. On the other hand, even if the cases assessed different aspects, the fact remains that the reliability measure restrictedly represents the correspondence of prospect theory with affect or user value.

Secondly, the user value questionnaire items suggested might not encompass the prospect theory. For example, the questionnaire item, “I want to recommend my smartphone to others” for the attachment dimension, and “I feel better if I get in touch with my friends” for the sociability dimension can lead to quick answers by many of us. The questionnaire items generally assess subjective ideas or feelings, but the degree of uncertainty may be low, which indicates that prospect theory might not be suitable.

4.3.3. Parameters of Value Function

Parameters of the function originally suggested by Kahneman and Tversky [12] consisted of \( r \) and \( \lambda \) which indicate the exponent for sensitivity and the coefficient of loss aversion, respectively. In addition, a reference score, \( \alpha \), was postulated in this study to convert the interval scale data into ratio
scale data. In the case of the reference score, as mentioned above, further study will be expected to investigate the optimal or practical point, which is estimated to be between 50 and 75 on the 0–100 scale.

Above all else, the exponent for sensitivity, $r$, and the coefficient of loss aversion, $\lambda$, needs to be investigated in the future study. Tversky and Kahneman [29] have already estimated $r$ as 0.88 and $\lambda$ as 2.25. However, because of differences in the situation that a decision-maker is faced with, the value of the parameter may vary. Note that previous study has suggested the concept of disappointment aversion instead of loss aversion. Even if parameters suggested by Tversky and Kahneman [29] performed well in this study, improved performance would occur if better parameter values are discovered.

5. Conclusions

This study tried to confirm a methodology to make compensations in psychophysical measurement using prospect theory in two case studies. First, a prospect theory-based approach was originally verified in a previous study through a usability assessment experiment consisting of serial target selection tasks [8]. In order to enlarge the base and ensure the methodology per se, two case studies were used in this study, targeting affect and user value concepts.

The first case study included a psychophysical assessment of affect constructs. Six sub-dimensions of affect (i.e., delicacy, simplicity, texture, luxuriousness, color, and attractiveness) as well as the overall affect were assessed by psychophysical measurement with a 0–100 scale. A tablet PC was used as an experimental device in the laboratory experiment with a total of 26 participants. As demonstrated by the results, the goodness of fit of the suggested models increased after the prospect theory-based transformation.

The second case study probed user value questionnaire items with a psychophysical assessment using a 0–100 scale. The overall user value concept was supposed to have five dimensions: attachment, sociability, pleasure, self-satisfaction, and customer need. Three to four questionnaire items were designed for each dimension and were verified [31]. A total of 295 participants rated their own mobile phone in terms of user value. As the results indicate, the goodness of fit of the suggested models increased slightly with data transformation based on prospect theory.

One of the limitations of this study is that it is not easy to directly prove something related to human cognition. In particular, the problem of judging whether a psychophysically rated one is right or not is really difficult. Various data analysis can confirm this indirectly. This study accepts the results of previous studies [8] and assumes that humans are based on specific scores when evaluating something. As a result, 75 points on the 0–100 scale, which were important criteria in previous studies, appear to have worked in this study. In both case studies, the baseline score of 75 improved the performance of the various models. This is only what researchers can determine from this study, and further research is needed to reveal further facts.

Although an increase in model performance was observed, further study needs to be conducted to compensate for the methodology. Considering gender and age groups were not fully balanced for a part of experiments, further study needs careful recruitment of participants for increasing the generalizability of our findings. Additional case studies including other psychophysical scales, such as a 0–10 scale, would confirm the reliability and validity of the methodology suggested in this study. In addition, optimal or practical parameter values (i.e., the exponent for sensitivity, the coefficient of disappointment aversion, and the reference score), which would be revealed in further studies, will enlarge the theory utilization. Most importantly, the methodology would be expected to help designers to precisely confirm their design alternatives, and researchers to understand human behavior during the assessment of abstract constructs.

Author Contributions: Conceptualization, M.C.L.; supervision, J.P.

Funding: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. 2018R1C1B6008848).

Conflicts of Interest: The authors declare no conflict of interest.
References

26. Stevens, S.S. Cross-modality validations of subjective scales for loudness, vibrations, and electric shock. *J. Exp. Psychol.* 1959, 57, 201–209. [CrossRef]
