

Application of the Labeled Magnitude Scale in Kansei Research

Chun Yueh Chen¹ and Kuohsiang Chen²

¹ University Rd., Tainan City 701, Taiwan (R.O.C.)
karmaforza@gmail.com, kchen@mail.ncku.edu.tw

Abstract. This study intended to construct a labeled magnitude scale based on Kansei researches so that data with ratio-level can be retrieved easily and further extensive analysis can be conducted. In this study, scale derivation was generated based on the research of Green et al.[8] 32 subjects, include 23 male and 9 female, average age of 24.6 yrs with design education background, participated in the experiment. 19 car samples and 5 Kansei phrases were used in the experiment. Results showed that the intensity indicators were significantly different in the experiment ($F_{1,137}=122, p=.727$). But subjects in the experiment gave rating in different ways to each Kansei phrases (Kansei phrases * intensity indicators, $F_{20,2952}=3.55, p=.00$). This may due to the status quo bias of subjects. Comparisons with OPUS (Oral Pleasantness /Unpleasant scale) [9] and CALM (Comfort Affective Labeled Magnitude) [5] showed similar orders of intensity indicators but different maximum magnitude in each scales.

Keywords: Kansei research, labeled magnitude scale (LMS), scale method, sensory evaluation.

1 Introduction

Kansei engineering has been proved a powerful tool in product development. With the increasing successful cases in Kansei engineering, the engineering methods in Kansei engineering have evolved for specific usages. But the basic principle of Kansei engineering still the same: to know what the customers' feelings are and to measure how strong the feeling (Kansei) is.

Conventional Kansei questionnaires use Likert scale or semantic differential scale as the scaling method. Both of these two methods belong to category scale. And category scale has its own nature limitation. Category scales yield ordinal-level data, sometimes interval level data. And the deficit limits the further statistical analysis of retrieved data.

With the study of Green et al. [8], a line scale with uneven-placed intensity labels was derived and the result from the scale is proved comparable with the result retrieved by magnitude estimation. The advance of the scale compared to line scale with even-placed labels was proved. That is the labeled magnitude scale, or the LMS in the study.

This study will construct a Kansei labeled magnitude scale based on the methodology of Green et al. [8] Comparison to labeled magnitude scales in other fields will be conducted.

2 Literature Review

2.1 Kansei Engineering

2.1.1 The Definition of Kansei Engineering

This term, 'Kansei engineering', or Kansei ergonomics, was first introduced by Dr. Nagamachi in Hiroshima University in about 30 years ago [14]. And the most general definition of Kansei is *"A field that deals Kansei with engineering techniques, and to analyze Kansei by engineering means for the development of product. Therefore products with fulfillment and delight can be introduced."*

Nagamachi has divided Kansei engineering/ergonomics into three subsections. They are :

- Kansei engineering type I is a method of category classification. The method is to break down the Kansei categories elicited by classification and systemize the categories into hierarchical structures.
- Kansei engineering type II is an expert system that translates consumers' feelings and images into design details of related products.
- Kansei Engineering type III is also called KEM, Kansei Engineering Modeling. The technique is to establish a mathematic, rule-based model to get ergonomic results from inputted Kansei phrases. [13]

2.1.2 Applications of Kansei Engineering

Kansei engineering has been accepted broadly and applied in many fields for product developments. For example, Mitsubishi, Mazda, Toyota, Honda, Ford, Hyundai, Delphi Automotive Systems in automobile industries, Sharp, Sanyo, Matsushita, Matsushita Electric Works, LG, Samsung in consumer electronics business, Matsushita Electric Works, YKK Design, Tateyama Aluminum, Kansai Electric Power Plant in house construction corporations, [14] and even Asahi beer in food industries, have carried out Kansei engineering methods as means of new product design, development and information mining from customers.

2.2 Sensory Evaluation

As the first human being using his or her own sensation to tell the quality of food, water, cloth or anything can be used or consumed, the history of sensory testing had begun. Human, as the only measuring instrument used in sensory evaluation, is not perfect, or prone to go wrong. Testers in a sensory evaluation are quite variable over time, variable between themselves, and apt to be biased. So scientists have developed formalized, structured, codified methodology for better assessing humans' sensory. A formal sensory study consists of at least seven stages. They are determining the project objectives, determining the test objectives, screening the samples, designing the test, conducting the test, analyzing the data and interpreting/reporting results.

2.2.1 Scaling Methods in Sensory Evaluation

There are four kinds of methods used in sensory evaluation. They are classification, grading, ranking, and scaling. [12] The method discussed further here belongs to scaling methods.

Category scale:

The earliest category scale were devised by the Greek Astronomer Hipparchus (190~120 B.C.) to assess the brightness of star. In category scale, subjects are asked to rate the intensity of preferred attributes by assigning values on a limited, often numerical, scale. The data from category scale is usually considered to be at least ordinal-level data, sometimes interval-level data.

Line scale

Line scale is a kind of scale considered to yield interval-level data. Subjects in line scale rate the intensity of a given stimulus by having a mark on a horizontal line that represents the amount of the perceived intensity. The line used in line scale has intensity marks, or anchors in both ends. In general, the left end of the line corresponds to “none” or zero and the right end of the line are a very large amount of stimulus. The line scale could be uni-polar or bipolar.

Magnitude estimation

Magnitude estimation is a scale method based on Stevens’ power law [12].

$$\text{Steven's power law: } R=KC^n$$

The first sample a subject receives is assigned a free-chosen number (the number can be assigned by the experimenters or the subject him/herself). Subjects are then asked to assign all following-up ratings for the following-up samples in proportion to the first sample rating. If a sensation of a sample is three times strong of the first sample, then the rating of the sample should be assigned three times of the first rating. Subjects are reminded to keep the rating in proportion to the ratios among sensations.

Magnitude matching

Magnitude matching is a scale method that yields ratio-level data. In this technique, subjects match the intensity of one attribute to another attribute. If the perceived magnitude and the physical intensity follow the Stevens’ power function, the attribute used to compare with would obtain the ratio property of the original attribute. Another benefit is that there is no number assigned in the technique, so people who used to use number differently will not interfere with the results.

2.2.2 Green’s Labeled Magnitude Scale (LMS)

Labeled Magnitude Scale is a line scale with uneven-positioned intensity labeled. The intensity indicators on the scale are placed according to the retrieved magnitude estimates [4]. Green and his colleagues verified the scale by holding experiments comparing the LMS with magnitude estimation and ordinary line scale with evenly-spaced intensity labels. The comparisons shown that the LMS yielded ratio-level data as magnitude estimation did and was better than ordinary even-label-placed line scale.

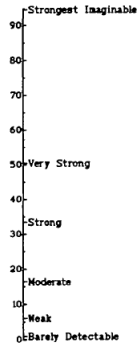


Fig. 1. The Labeled Magnitude Scale derived from Green et al.

2.2.3 Bartoshuk's General Labeled Magnitude Scale (gLMS)

In her researches, Bartoshuk and her colleagues found genetic variation may alter human oral sensations of PROP (Propylthiouracil) [3]. And prior pain experiences may also distort or change the inner pain scale for those who had suffered extremely painful events [7]. The alternation of inner sensation scale is called 'elastic ruler'.

Bartoshuk and her colleagues introduced the general Labeled Magnitude Scale based on the LMS [8] to avoid the elastic ruler conundrum, and the semantic intensity labels have been altered for better fitting cross-modality usages.

Compared with the results of magnitude matching, that of the gLMS were found to provide valid cross-group comparison as Magnitude Matching did, but invalid comparison could occur when the maximum magnitude is correlated with preference of sensation.

Although with found limitation of usage, the gLMS showed 'the paradigmatic advance' in across-group comparisons. Further researches and implement in various fields were suggested.

2.3 The Application of Labeled Magnitude Scale in Kansei Research

Kansei researches suffer from the issues of invalid cross-group comparison with conventional categories scales. For example, gender difference may cause the 'elastic ruler' conundrum and direct comparisons become impossible. With the application of the LMS, direct comparisons between different groups with varied sensations are possible. And the data yielded from the LMS is ratio-level that creates more freedom for statistical analysis.

With the advantages in cross-group comparison and statistical analysis, the application of LMS to Kansei researches is beneficial.

3 Experiments

3.1 Scale Generation

The experiment employed the method of LMS with adjustments based on gLMS [2] to fit the usage of Kansei evaluation. In the experiment of scale generation, the subjects

were asked to make modulus-free magnitude estimates for a series of intensity indicators (semantic labels) of five Kansei phrases adopted from previous study. The six descriptors used were 'barely detectable', 'weak', 'moderate', 'strong', 'very strong', and 'strongest imaginable of any kind' for all 5 Kansei phrases. A set of car photos were sampled to represent the whole car category.

According to the study of Ma [11], five pairs of Kansei phrases were selected by experts of car styling for Kansei evaluations. The five chosen pairs of Kansei phrases were 'streamlined-geometric', 'staid-light', 'rounded-sharp', 'complex-simple', and 'speedy-dilatory'. In this study, one out of each pair of Kansei phrase was selected and used for Likert scaling in the experiment. The chosen Kansei phrases were 'streamlined', 'light', 'sharp', 'simple' and 'speedy'.

Consumer Report magazine classifies passenger cars into twenty four categories. Each category of passenger cars had one model chosen as experimental samples. Model photos were collected via internet from manufacturers' official websites or press resources. In order to equalize the general images delivered by samples photos, all car models were in similar posing angles in photos.

All sample photos were processed to remove the background of car models. Brand marks and model labels of each sample were also removed. All sample photos were in monochrome for avoiding color biases. [15]

Twenty-three males and nine females with average age 24.9 yrs (range from 20~46) participated in this experiment. All subjects had at least 3 years training or education in industrial design and styling.

The subjects were told to have an experiment about Kansei engineering and all they had to do was to give ratings to the samples. No prior information about the process of experiment and samples were given during the recruitment.

3.2 Procedure

Before the experiment sessions getting started, subjects were instructed about the process of the experiment and what they were expected to do in the experiment sessions. A set of sample image cards was provided for reference. The experiment had three sessions. In session 1, subjects were asked to assign an intensity indicator displayed one by one in random order on the screen to each Kansei phrase for 19 samples. Subjects were asked to press the confirm button when an appropriate intensity descriptor shown. In session 2, subjects were asked to rate the intensity magnitude for each Kansei phrase. Session 3 was the matching stage. The intensity indicator assigned in the first stage and the values given in the second stage were exhibited side by side for subjects to review and verify their ratings.



Fig. 2. The experimental interface in sessions 1, 2, and 3

In all three stages, the presenting orders of samples and Kansei phrases were randomized. And all values or intensity indicators subjects gave could be reviewed and corrected by the subjects anytime during the experiment.

4 Results

All experiments were held within one week. The data collected from the experiment were first treated with modulus equalization. After that, results of the linear mixed-model analysis indicated that gender difference in the subjects didn't alter the rating results ($F_{1,137}=.122, p=.727$), and intensity indicators did differ from each other ($F_{5,207}=43.82, p=.00$). So the meaning of each descriptor should be well accepted and understood by the subjects. But the results in linear mixed-model analysis with Kansei phrases * Intensity shown that subjects in the experiment gave rating in different ways to each Kansei phrases (Kansei phrases * intensity indicators, $F_{20,2952}=3.55, p=.00$).

Pairwise tests also shown that all intensity indicators were well differentiated ($p<.05$) but exception still existed. In the case of Kansei phrase 'streamlined', the descriptor pair 'very strong - strongest imaginable of any kind' was not well differentiated. In the phrase 'simple', three descriptor pairs 'barely detectable - weak', 'strong - very strong', and 'very strong - strongest imaginable of any kind' were not significantly different. In the phrase 'speedy', the descriptor pairs 'barely detectable - weak' and 'very strong - strongest imaginable of any kind' were not significantly different. For the phrase 'light', 'barely detectable - weak', 'strong - very strong', and 'very strong - strongest imaginable of any kind' were not significantly differentiated. In the pairwise test for 'sharp', between the intensity indicators 'barely detectable - weak' and 'very strong - strongest imaginable of any kind' were not significantly different.

The geometric means and standard deviation of each intensity descriptor were calculated based on the research of Alf and Grossberg [1]. Values of geometric means and standard deviation are shown in table 1.

Table 1. Geometric means and the standard deviations of magnitude estimates (n = 32) given to six different intensity indicators in each category of Kansei phrase

Intensity indicators	Stream-lined	simple	light	Sharp	Speedy	Over all
<i>Barely detect-able</i>	2.70 (0.18)	3.37 (0.34)	3.74 (0.21)	3.75 (0.11)	3.39 (0.26)	3.44 (0.12)
<i>Weak</i>	4.96 (0.27)	5.57 (0.22)	6.29 (0.33)	6.13 (0.54)	4.85 (0.31)	5.58 (0.16)
<i>Moderate</i>	8.36 (0.22)	6.05 (0.24)	10.45 (0.41)	8.41 (0.37)	8.02 (0.27)	8.44 (0.14)
<i>Strong</i>	12.57 (0.47)	10.27 (0.28)	15.87 (0.88)	13.61 (1.03)	10.94 (0.35)	12.09 (0.27)
<i>Very strong</i>	16.80 (1.39)	14.33 (0.97)	19.53 (2.06)	17.66 (2.97)	16.53 (1.25)	16.53 (0.74)
<i>Strongest imaginable of any kind</i>	19.44 (1.33)	16.26 (2.04)	22.85 (2.62)	26.68 (6.05)	17.84 (1.88)	19.63 (1.04)

Standard deviation in parentheses (Alf & Grossberg, 1979)

5 Discussions

Results from thirty-two subjects showed that there were significant differences between the ratings of intensity indicators of five Kansei phrases. But subjects also rated each Kansei phrase in significantly different way. So do the results suggest that a universal labeled magnitude scale of Kansei doesn't exist, or the experiments were influenced by some unexpected factors?

5.1 Generality of the Scales

Results indicated that subjects did give ratings in different ways for different Kansei phrases. And significant differences between Kansei phrases were only found in the middle segment of intensity indicators, such as 'moderate', 'strong', 'very strong'. There are several questions laid among the findings. It may be caused by magnitude equalization which was used to equalize the range of data and keep the structure inside. Since the range of the data was about the same, the anti log magnitude estimation value of the first two intensity indicators and the last one would be about the same. The difference of data would appear in the middle section of intensity indicators.

5.2 Status Quo Bias in Kansei

As previous mentioned, Kansei is the general images or attributes aroused by sensory stimuli of certain objects or events. Since Kansei phrases represent the attributes of general images, Kansei phrases may be assigned a default value by the subjects. The default value in Kansei phrases, may cause status quo bias [10] because there is always a status quo of specific sample used as the standard or reference no matter the referential sample is assigned or not. If a reference is not assigned in a Kansei evaluation, the panelists in the evaluation still compare all the samples with their own prior experiences.

5.3 Compare with Existing Scales

This research is the first try to adopt labeled magnitude scales into Kansei researches. And the comparison to the labeled magnitude scales from other fields reveals agreements and differences.

The currently developed LMS can be divided into the sensory scale and the affective scale. In the work of Guest and his colleagues [9], they suggested that there is a general dichotomy between affective and intensive sensations. The following is how they discriminate sensory scale and affective scale.

"For example, Gracely et al.(1978a) found that the most intense general sensory descriptor in a set of 15 was rated as 85 times more intense than the weakest descriptor, i.e., the sensory range was 85:1. Similarly, the ratio between most and least intense descriptors in Green et al.'s LMS is 69:1. In contrast, for classes of affective intensity indicators, the sensory range has consistently been found to be much smaller, varying from 9:1 for positive affect in Schutz and Cardello's LAM, through 10:1 (Gracely et al., 1978a) to 15:1 for our OPUS" [9]

According to the definition by Dr. Nagamachi, Kansei is more likely to be an affective sensation than a sensory sensation. And the result of experiments in this study agrees with Dr. Nagamachi's words. Therefore, the LMS of each Kansei phrase will be compared with the oral sensation scale for comparison here is OPUS, oral pleasantness /unpleasant scale and the tactile sensation scale for comparison is CALM, Comfort Affective Labeled Magnitude [5].

Figure 7 is the comparison of five Kansei phrases scale with OPUS and CALM. OPUS was conducted by Cardello and Shuntz [6] with an adverb intensity descriptor system, and is different from the LMS [8]. As the figure shows, the order of intensity indicators in OPUS is similar to the scales of five Kansei phrases. The most difference occurs in the maximum magnitude. The mean of the maximum magnitude of five Kansei phrases is 19.63, with a standard deviation of 1.04. However, the maximum magnitude of OPUS is about two times larger than the maximum value in Kansei scales. The maximum pleasant magnitude is 51.90 (1.67), and the maximum unpleasant magnitude is 54.44 (1.72). The same circumstance is also found in CALM. The most comfort magnitude originally retrieved in the experiments of CALM is 366.72 (0.10), and the most uncomfortable magnitude in CALM is 350.67 (0.10).

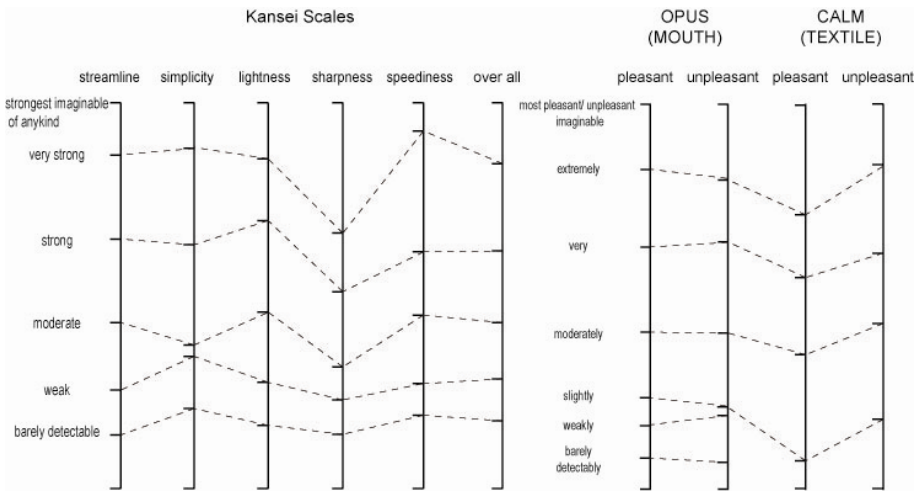


Fig. 3. The comparison of the scales of five Kansei phrases with OPUS and CALM

6 Conclusions

With this study, a Kansei labeled magnitude scales have been devised based on the method suggested by Green et al. [8]. An experiment comparing intensity indicators with magnitude estimation has been carried out. Kansei phrases used in this experiment are acquired in the study of Ma [11] and the intensity indicators in the experiment are from the study of Bartoshuk and his colleague [2]. In the experiments, corresponding geometric means of each intensity indicators have been retrieved and these values are used for deriving the labeled magnitude scale for these Kansei phrases.

In the experiments, significant difference between intensity indicators was found. But analysis also showed that there was significant difference between the five Kansei phrases. This phenomenon may due to the effect of loss aversion. Comparison between the labeled magnitude scales in other sensations shown great difference in the maximum intensity magnitude.

Acknowledgement. Prof. Chin-Fang Sheu and Prof. Kuohsiang Chen helped a lot to foster the idea, and the support from my family and friends. Thanks to all of them.

References

1. Alf, E.F., Grossberg, J.M.: The geometric mean: Confidence limits and significant tests. *Perception & Psychophysics* 5(26), 419–421 (1979)
2. Bartoshuk, L.M., Duffy, V.B., Green, B.G., Hoffman, H.J., Ko, C.-W., Lucchina, L.A.: Valid across-group comparisons with labeled scales: The gLMS versus magnitude matching. *Physiology & Behavior* 82, 109–114 (2004)
3. Bartoshuka, L.M., Duffy, V.B., Fast, K., Green, B.G., Prutkin, J., Snyder, D.J.: Labeled scales (e.g., category, Likert, VAS) and invalid across-group comparisons: what we have learned from genetic variation in taste. *Food Quality and Preference* 22, 125–138 (2002)
4. Butler, G., Poste, L.M., Wolynetz, M.S., Agar, V.E., Larmond, E.: Alternative analyses of magnitude estimation data. *Journal of Sensory Studies* 2, 243–257 (1987)
5. Cardello, A.V., Winterhalter, C., Shuntz, H.G.: Predicting the handle and comfort of military clothing fabrics from sensory and instrumental data: development and application of new psychophysical methods. *Textile Research Journal*, 221–237 (2003)
6. Cardello, A.V., Schutz, H.G.: Research note: Numerical scale-point locations for constructing the LAM (labeled affective magnitude) scale. *Journal of Sensory Studies* 19(4), 341–346 (2004)
7. Dionne, R.A., Bartoshuk, L., Mogil, J., Witter, J.: Individual responder analyses for pain: does one pain scale fit all? *TRENDS in Pharmacological Sciences* 26(3), 125–130 (2005)
8. Green, B.G., Shaffer, G.S., Gilmore, M.M.: Derivation and evaluation of a semantic scale of oral. *Chemical Senses* 18(6), 683–702 (1993)
9. Guest, S., Essick, G., Patel, A., Prajapati, R., McGlone, F.: Labeled magnitude scales for oral sensations of wetness, dryness, pleasantness and unpleasantness. *Food Quality and Preference* 18, 342–352 (2007)
10. Kahneman, D., Knetsch, J.L., Thaler, R.H.: Anomalies: The Endowment Effect, Loss Aversion, and Status Quo bias. *Journal of Economic Perspective* 5(1), 193–206 (1991)
11. Ma, C.-C., Shieh, M.-D.: Using Neural Networks in Automobile Shape Feature Design. Tainan: NCKU, Master Thesis (in Chinese) (2005)
12. Meilgaard, M.C., Civille, G.V., Thomas Carr, B.: *Sensory evaluation techniques*, 4th edn. CRC Press, Taylor & Francis Group, Boca Raton (2007)
13. Nagamachi, M.: Kansei Engineering: A new ergonomic consumer-oriented technology for product development. *International Journal of Industrial Ergonomics*, 3–11 (1995)
14. Nagamachi, M.: Technical note: Kansei engineering as a powerful consumer-oriented technology for product development. *Applied Ergonomics* 33, 289–294 (2002)
15. Wong, J.-S., Chen, L.-L.: A Study on the Relationship of Automobile Shape Morphing to Affective and Aesthetic Responses. Taipei: NTUST, Master thesis (in Chinese) (2004)