COMPARISON BETWEEN DISCRETE AND ANALOG SEMANTIC DIFFERENTIAL SCALES ACCURACIES IN KE - CASE STUDY: RECEPTION CHAIRS

Nasser Koleini Mamaghani[a], Ali Dasmeh[b]

[a] *Associate Professor, Industrial Design Department, School of Architecture and Environmental Design, Iran University of Science and Technology, Iran, koleini@iust.ac.ir*

[b] *Graduate student, Design and Creativity Department, Institute of Cognitive Science Studies, Iran, alidasmeh@ut.ac.ir*

ABSTRACT

Researches show that observers have a tendency to choose the extreme ends of Likert or discrete semantic differential rating scales. Conventionally, the discrete semantic differential rating scale is used in Kansei Engineering (KE) and the above-mentioned tendency may affect the result of KE. The general aim of this paper is comparing the accuracy of two types of conventional semantic differential scales, discrete and analog. The data has been collected through a Kansei Engineering process in order to collect data in a practical situation, not an isolated laboratory situation. The result of the current research shows the previous researches' achievements are not true in the field of KE and the distribution of the scores is almost normal.

Keywords: Likert, visual analog, rating scale, normal distribution, reception chair

1 LITERATURE REVIEW

Kansei Engineering was developed in Japan as a way to translate consumers' feelings about the product to specific design decisions (Nagamachi, 1995). Kansei Engineering could be used to identify which design characteristics influence users' perception of whether or not a design is modern (Tama, 2015). The word "modern" is an instance of adjectives, any adjective can be replaced. Kansei Decomposition is used to link the desired perceptions of a design with senses and engineering characteristics, so this decomposition identifies which engineering decisions may affect user perception of a design (Ayas, 2011). In the terminology of KE, item implies the design item of the sample product, and category means the detail of the design item. For instance, color, shape, size, roundness, and so forth are examples of items; and red, yellow, green, blue, and so

forth are the categories for the color item. The Kansei/affective engineer should be very careful of the sample product's items and categories (Nagamachi, 2011).

KE benefits from a modified version of semantic differential method in data collection processes (Nagamachi, 2011). [Semantic differential] five points (five bipolar pairs of adjectives) have been proven to yield reliable findings, which highly correlate with alternative Likert numerical measures of the same attitude (Osgood, 1957). Developing self-report Likert scales is an essential part of modern psychology (Jebb, 2021). Likert scaling, introduced by Rensis Likert, is the most widely used method of measuring personality, social, and psychological attitudes. The popularity of Likert scales can be traced to a number of factors, including ease of construction, intuitive appeal, adaptability, and usually good reliability (Hodge, 2003).

The accuracy and validity of a rating scale have been one of the most important concerns of all researchers, so they proposed different methods to cover this concern. Shutte and Eklund (2010) introduced an online platform for the KE data collection step. They used VAS instead of Likert scale in SD method. Betella and Verschure (2016) proposed a modified version of Self-assessment Manikin (SAM) rating scale which is called Affective Slider (AS). AS is a digital self-reporting tool composed of two slider controls for the quick assessment of pleasure and arousal. Fernandes and Yamanaka (2020) presented Affective Impressions Scale (AIS), a dedicated tool to measure the Affective Accuracy and Affective Synchrony levels of peers in a group.

One of the principle tenets in constructing instruments is that items be as clear and concise as possible. The more items are characterized as cognitively complex, the more likely respondents are to misunderstand the question and answer incorrectly. Even small differences in wording can increase the level of cognitive noise and dramatically alter response patterns (Hodge, 2003).

In practice most participants in Likert surveys cluster to the extremes. The distribution of responses displays a characteristic "W" shape, with most participants clustering toward the extremes, some expressing indifference, and few in between. Most researchers agree that the W shape does not display the true distribution of preferences, which in reality is most likely a bell-shaped or normal curve (Posner, 2018). Jacoby and Matel (1971) mentioned: "Based upon the evidence adduced thus far, reliability should not be a factor in determining a Likert-type scale rating format, because it is independent of the number of scale steps employed". They compared the result of a survey in three-steps and five-steps Likert scales. There was no significant difference in the results. The result of that experiment supports the Posner's claim.

The focus of this paper is the second aforementioned problem and the aim is comparing the result of Visual Analog rating scale (VAS) and Likert. If there was a significant difference in the distributions between two groups, we propose to researchers use SD method with VAS rating scale.

2 MATERIAL AND METHODS

The subject of the current study is reception chairs. We assumed every reception chair can be imagined by three items: backrest, seat, and arm (Figure 1). The items are selected based on the

features table for this type of chair, in an online shopping platform. For every item some categories are assumed which are presented in Table 1, these categories are the most repeated categories for each item in the online shopping platform.

Figure 1. Three items for the current research: Backrest, Seat, Arm

As presented in Table 1: two categories for the backrest, three categories for the seat, and two categories for the arm are considered for the current research. It means 12 samples (2*3*2 = 12) must be presented to observers. As it is noted in the first section, samples must be selected from the already existed products in the market. Three combinations (three samples) of mentioned categories could not be found in the Iran's online market (not found samples: b2a1s1, b2a1s3, b2a2 s3). The presented samples to observers are presented in Table 2.

Table 1. The list of items and categories

#	item	categories
1	Backrest	Whole (b1) , Half (b2)
2	Arm	With (a1) , without (a2)
3	Seat	Plastic (s1) , with pillow (s2), with thick pillow (s3)

Table 2. Sample and markers

marker	images	marker	images
b1a1s1		b1a2s3	

b1a1s2		b2a1s2	
b1a1s3		b2a2s1	
b1a2s1		b2a2s2	
b1a2s2			

Through a web engineering technique – it is called scraping - all customers' reviews for this type of chair were fetched from the online shopping platform. Common used adjectives were filtered and through the most famous Persian online dictionary (Moin Encyclopedic Dictionary) for every filtered adjective, one contradictory adjective was found. Eventually, eleven paired contradictory adjectives were presented to observers. Table 3 includes the adjectives in Persian and their English translation.

All the samples and adjectives are presented to the observers through an online platform (www.kansei.ir) which was developed by the researchers for this purpose. The platform has two types of rating scales that are discussed in the paper. Figure 2 is presenting an image of the platform in both situations. Every user filled the form with randomly ordered samples, while the order of adjectives was the same for all users.

49 observers in two groups filled the forms. The first group, filled the form with Likert rating scale (n=26, average of ages = 24, SD=3.3), The second group, filled another type of form with VAS rating scale (n=23, average of ages=25, SD =3.8). All observers were design students in Iran universities, which found the forms (the link) in their classroom online forums. The methodology of sampling in the current paper is considered as simple random sampling.

Table 3. Kansei words (adjectives)

Persian	English		Persian	English
خودمانی - رسمی	Intimate – Formal		مدرن – کلاسیک	Modern – Classic
زشت – زیبا	Ugly – Beautiful		فیک – اصیل	Fake – Original
فکستنی – شیک	Punk - Stylish		زمخت – ظریف	Clumsy - Elegant
شکننده – مقاوم	Fragile - Resistant		تزئینی – کاربردی	Decorative - Applicable
آرامش بخش – تنش زا	Soothing - Stressful		بد شکل – چشم نواز	Deformed – Eye-catching
طراحی اصولی – طراحی غیر حرفه‌ای	Ergonomic – Not ergonomic			

Figure 2. Left: a sample of Likert rating scale in Kasnei.ir platform.
Right: a sample of VAS rating scale in Kansei.ir platform.

3 RESULTS

Table 4 presents the results of the above mentioned survey. Each chart represents the average users' score in two different methods, the dashed line is showing the result of the Likert method, while the straight line is related to the VAS method. The last column of every chart is the average of all scores for the product.

A t-test is done on the average score on products between two categories. The result is showing a significant difference between the two groups (p-value < .05).

Figure 3 is presenting the distributions of scores in both methods. The VAS scores were categorized into five groups (every group includes 20 consecutive scores).

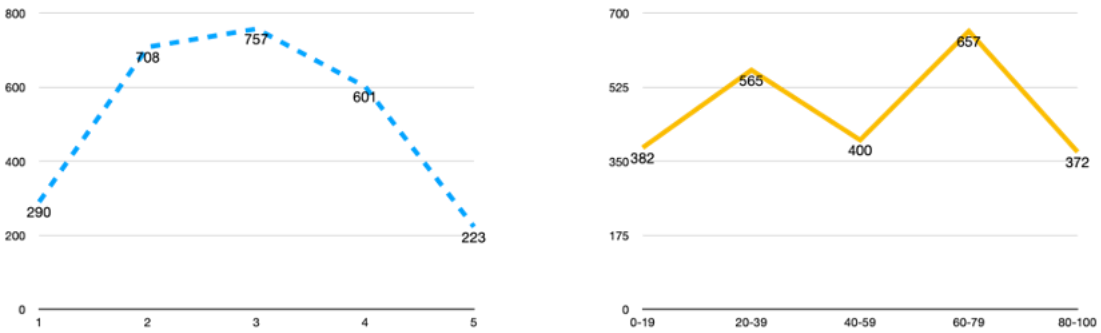
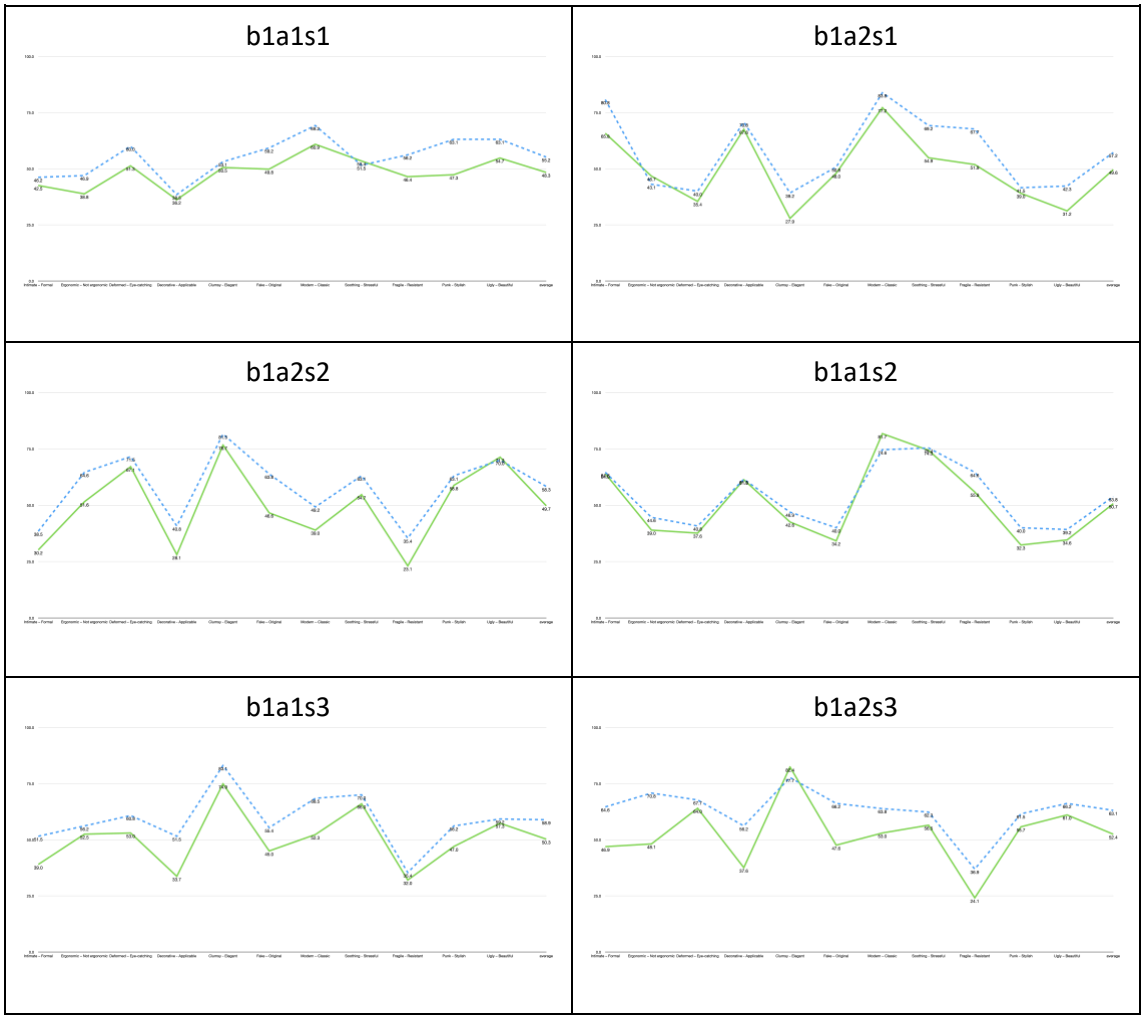
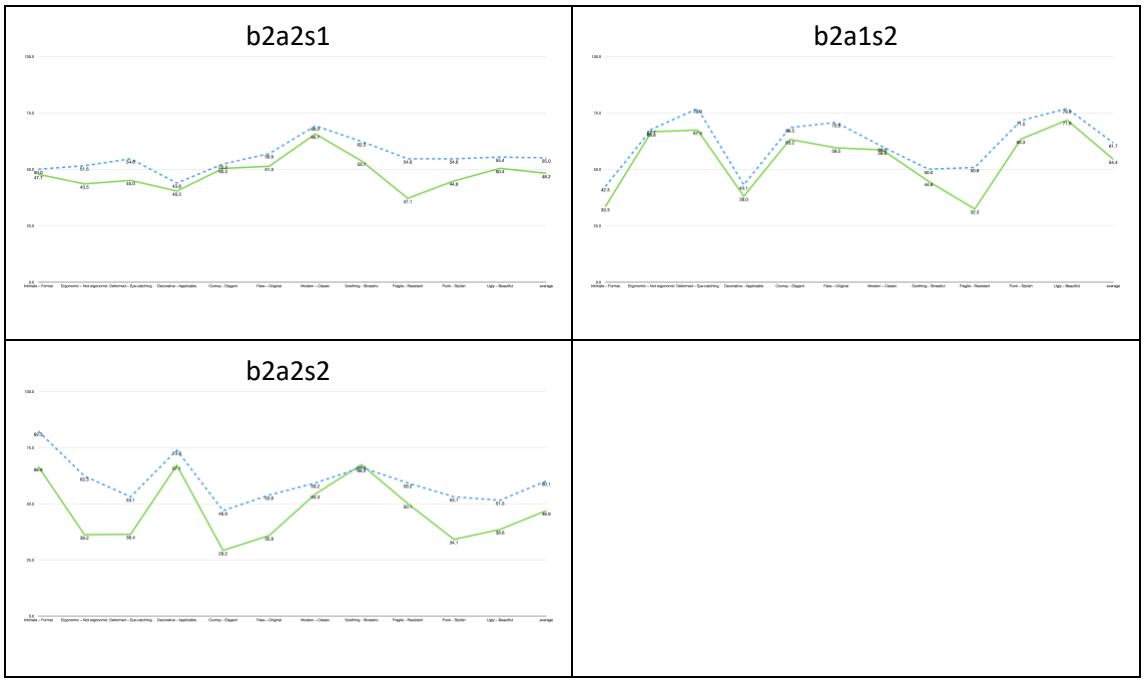


Figure 3. The left chart represents the distribution of score in the Likert method. The right chart represents the score in the VAS method. Y-axis in both graphs are representing the frequencies of every score and the X-axis in the right graph represents the scores between 0 and 100 in five clusters.

Table 4. The dashed lines represent the result of the Likert method and the line is related to VAS method. The Y-axis shows the average score from 0 to 100. The result of the Likert method is scaled from 0-5 to 0-100. The x-axis is the list of Kansei words. The Kansei words (in order of x-axis): intimate-formal, ergonomic-not ergonomic, deformed-eye-catching, decorative-applicable, Clumsy-elegant, fake-original, modern-classic, soothing-stressful, fragile-resistant, punk-stylish, Ugly-Beautiful, average





4 DISCUSSION

As can be seen in Figure 3, the distribution of scores in the Likert method is almost normal, contrary to what Posner claimed (2018) that the Likert's survey distribution should be "w-like". However, it seems that Posner's assumption is based on the gap instinct, the result of the current research shows that Posner's assumption is not true in the field of Kansei Engineering. It is possible that phenomenon occurs in the field of economics and other behavioral science studies.

In addition, in almost all averages of adjectives 'scores for each product, the score of the Likert method was greater than in the VAS method, as can be seen in chart 4. The result of the t-test for Likert and VAS scores shows a significant difference. We claim that the difference is due to the differences in accuracies in these two methods.

Strangely, the distribution chart of the VAS scores is "M-like". Scores in the range from 40 to 60 were less than expected. We claim the reason for this occurrence is a bias that could be arisen because of the default mode in the survey VAS version. In Figure 2, the right image, a VAS sample can be seen. The default version of any VAS sample has a scroll button in the middle of the range between two adjectives. It is possible that participants think they are obligated to move it. So the scores in the middle of the range are less than expected.

5 CONCLUSION

Although it was claimed that the distribution of scores in the Likert surveys is not normal, the result of the current experiment shows the Likert version of the semantic differential is working well in the Kansei Engineering data collection step. In addition, the results of this research suggest the possibility of better accuracy in the semantic differential VAS version in comparison with the Likert version, however, the distribution of VAS version was not normal.

REFERENCES

Ayas, E. (2011). Engineering Quality Feelings. Doctoral Thesis. *KTH Engineering Science.* Stockholm, Sweden.

Betella, A., & Verschure, P. F. M. J. (2016). The Affective Slider: A Digital Self-Assessment Scale for the Measurement of Human Emotions. *PLOS ONE*, *11*(2), e0148037. https://doi.org/10.1371/journal.pone.0148037

Fernandes, R., Yamanaka, T. (2020). Affective Impressions Scale (AIS): The Development of a Visual Scale to Measure Affective Synchrony Among Children. Proceedings of the 8th International Conference on Kansei Engineering and Emotion Research. KEER 2020. Advances in Intelligent Systems and Computing, vol 1256. Springer, Singapore. https://doi.org/10.1007/978-981-15-7801-4_5

Hodge, D. R., & Gillespie, D. (2003). Phrase completions: An alternative to Likert scales. *Social Work Research*, *27*(1), 45–55. https://doi.org/10.1093/swr/27.1.45

Kobayashi, M., & Kinumura, T. (2016). A method of gathering, selecting and hierarchizing kansei words for a hierarchized kansei model. *Computer-Aided Design and Applications*, *14*(4), 464–471. https://doi.org/10.1080/16864360.2016.1257188

Maver, T. (1991). Principles of optimal design: modelling and computation. *Design Studies*, *12*(1), 62. https://doi.org/10.1016/0142-694x(91)90014-n

Nagamachi, M. (1995). Kansei Engineering: A new ergonomic consumer-oriented technology for product development. *International Journal of Industrial Ergonomics*, *15*(1), 3–11. https://doi.org/10.1016/0169-8141(94)00052-5

Nagamachi, M. (2010). *Kansei/Affective Engineering (Systems Innovation Book Series)* (1st ed.). CRC Press.

Posner, E. A., & Weyl, G. E. (2018). *Radical Markets: Uprooting Capitalism and Democracy for a Just Society*. Princeton University Press.

Shutte, S., & Eklund, J. (2010). Ratingscales in Kansei Engineering : Modification for an European Context. *International Conference On Kansei Engineering and Emotion Research 2010*.

Tama, I. P., Azlia, W., & Hardiningtyas, D. (2015). Development of Customer Oriented Product Design using Kansei Engineering and Kano Model: Case Study of Ceramic Souvenir. *Procedia Manufacturing*, *4*, 328–335. https://doi.org/10.1016/j.promfg.2015.11.048