IEICE TRANS. FUNDAMENTALS, VOL. E77-A, NO. 7 JULY 1994

LETTER

Comparison among Methods for Compounding Psychological Scale Values in the Multiple-Scale Technique

Ayumi YOSHIKAWA† *and* Takeshi NISHIMURA†, *Members*

SUMMARY In this letter, we compare the three compound methods of the Multiple-scale technique to improve the quality of the scale values estimated by the method of fuzzy categories. The results show that the maximum compound method brings higher ability to estimate the scale values than the other methods despite categories used in the scale.

key words: *psychological scaling, method of fuzzy categories, Multiple-scale technique*

1. Introduction

The method of fuzzy categories (MFC) [1], a psychological scaling method, enables to quantify a subjective extent of an attribute measured with rating categories such as "very tall", "fairly tall" and so on. It can express the vagueness both of meaning of the category and of the subjective extent by fuzzy sets, while the method of successive categories [2] has neglected the vagueness. The MFC provides a set of scale values for the categories, which is necessary to estimate the subjective extent for each object, through fuzzy sets for the linguistic truth values that have the same verbal hedges as the categories. For example, a psychological scale value (i.e., a subjective extent of the attribute, e.g. "tallness") for an object is estimated by the membership function for "very true" when an assessor rates the object at "very tall." Figure 1 shows a schema of the MFC.

However, the MFC has lower freedom in terms of the expression of the scale values for objects than the fuzzy graphic rating scale [3], [4] because of the limitation of using discrete categories. To supplement the defect, the Multiple-scale technique (MUSCAT) compounds the scale value for one object from the values estimated in the MFC using at least two different categorical rating scales [5]. Yoshikawa and Nishimura [5] concluded that the maximum compound method was suitable for improving the quality of scale values obtained with their scales. However, they did not clarify whether the same result could be obtained from different scales or not.

This letter aims to clarify the generality of the results presented in Ref. [5]. We discuss the relations between categories used to compound and the quality of the scale values obtained by three compound methods. The results from our experiment show that the maximum compound method brings higher ability to estimate the scale values than the other methods despite the similarity between the categories.

Manuscript received January 12, 1994.

† The authors are with the Faculty of Engineering and Design, Kyoto Institute of Technology, Kyoto-shi, 606 Japan.

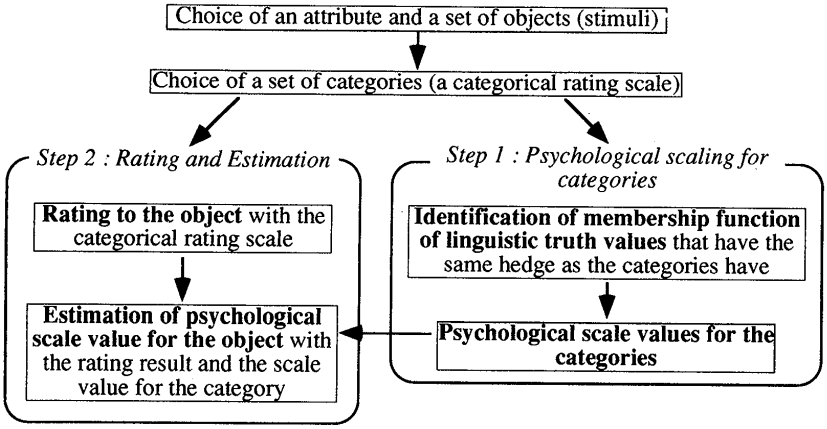
2. Experimental Conditions and Methods

This experiment was carried out with a set of questionnaires. We instructed our participants to answer their subjective extent for tallness of Japanese adult men indicated in centimeter. Both the fuzzy graphic rating scale method and the categorical rating scale method were used to measure their extent for tallness. Figure 2 shows the style of the questionnaire used in this experiment.

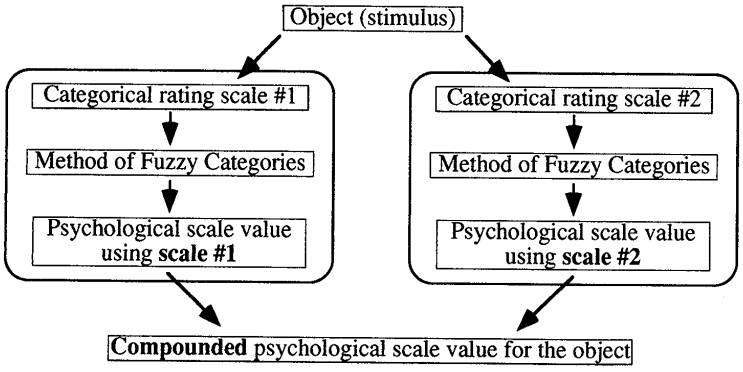
The graphic scale was at length of 10 cm and took three anchors (ticks) at both ends of the scale and its center. The right and left ends of the scale meant "perfectly tall" and "perfectly short" respectively. The participants marked two areas on the scales to answer their subjective extent for tallness of each physical height. One was the best area to show the subjective tallness of the height and the other was the maximum area to show it [1].

There were two categorical rating scales used in the experiment. Both scales took seven primary categories with verbal hedges and six intermediate categories located between primary categories without them respectively. The primary categories included in the scale 1 were "very short, " "quite short, " "slightly short, " "neither tall nor short, " "slightly tall, " "quite tall" and "very tall" and the categories included in the scale 2 were "tremendously short, " "fairly short, " "sort of short, " "neither tall nor short, " "sort of tall, " "fairly tall" and "tremendously tall" in order from left to right. The participants checked one preferable category on each scale to present their extent.

After rating the heights, the participants decided the membership functions for the thirteen linguistic truth values using the graphic rating scale in the same way as the tallness. The twelve of them were composed from both the six verbal hedges used in the categories and "true / false" such as "very true" and "sort of false." The rest was "neither true nor false."

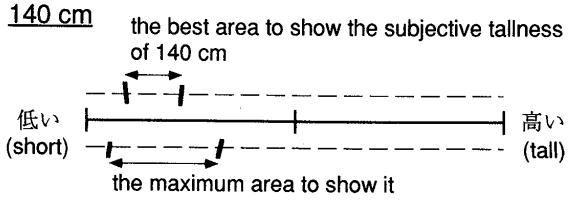


(a) schema of Method of Fuzzy Categories (MFC)

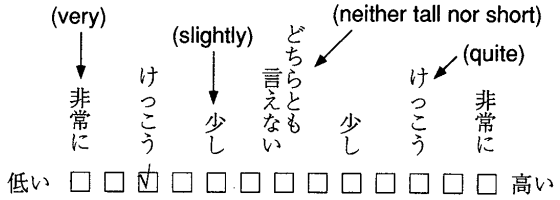


(b) schema of Multiple-scale Technique (MUSCAT)

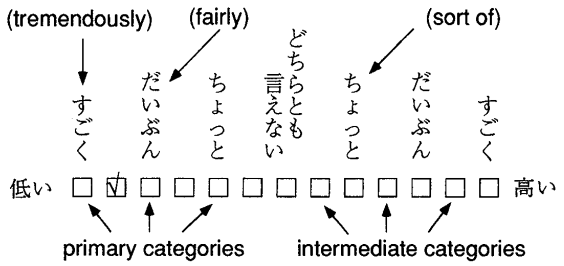
Fig. 1 Illustration of MFC and MUSCAT.



(a) continuous graphical rating scale



(b) scale 1 (categorical rating scale)



(c) scale 2 (categorical rating scale)

Fig. 2 Three rating scales used in the experiment.

All questionnaires were printed in Japanese. The nine of physical height ranged from 120 cm to 200 cm at the interval of 10 cm. The group of height were arranged in random order in the questionnaires. Fifty-four native Japanese volunteers participated in the experiment.

3. Results and Discussions

We obtained the psychological scale values for the tallness with the MFC. First, we gave the trapezoidal fuzzy sets for the thirteen linguistic truth values, matching the best area to 1 level set and the maximum area to the support set respectively and interpolating between their ends linearly. The linguistic truth values were the estimates of scale values for the primary categories with the same hedges. Second, the scale values for the twelve intermediate categories were calculated as the between-set between the neighboring primary categories [5], [6]. Last, the scale value of each height was estimated by the scale value for the category at which a participant rated it on the categorical rating scale. In the same way as the linguistic truth values, the nine trapezoidal fuzzy sets for the subjective tallness of the height were provided from the two areas measured with the fuzzy graphic rating scale.

Yoshikawa and Nishimura have adopted the maximum compound method, the minimum compound method and single scale as the MUSCAT in Ref. [5]. The maximum compound method makes a convex fuzzy set out of a union set of more than two scale values. On the contrary, the minimum compound method gives an intersection set of more than two scale values. However, the minimum compound method takes inherent disadvantage of which the maximum membership value for an intersection set is not necessarily unity. We thus use the between-set instead of intersection set and call our method intermediate compound method.

In the same manner as Ref. [5], the quality of the scale values estimated by the MUSCAT was shown in two indices, i.e. the matching measure and the similarity measure. We have adopted; a) a maximum membership value for intersection set between two fuzzy sets as the matching measure, and b) a ratio of cardinality of intersection set to union set as the similarity measure [7]. Both of the two measures take values between zero and unity. The two indices were calculated between the scale values of each height obtained from the fuzzy graphic rating scale and four scale values obtained from scale 1, scale 2, the maximum compound method, and the intermediate compound

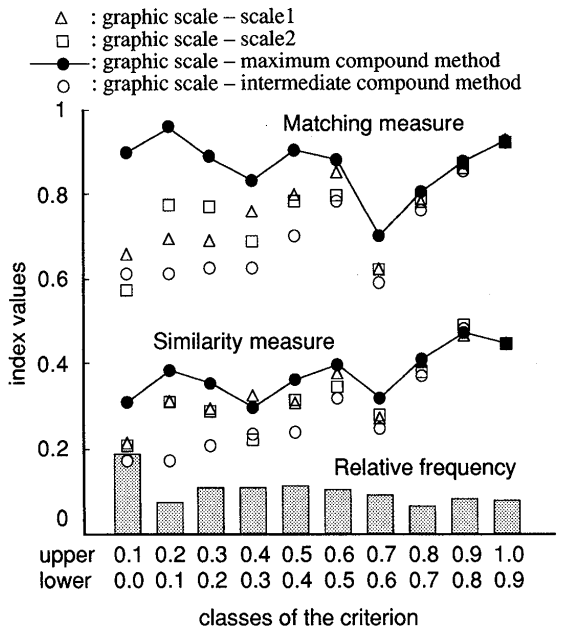


Fig. 3 The relations between the criterion and index values.

method. That is, we considered the scale values obtained from the fuzzy graphic rating scale as the estimates of the true subjective tallness [1], [5].

The averages and the standard deviations (shown in parentheses) of the matching measure over 486 data (nine of height by fifty-four participants) were 0.750 (SD: 0.363), 0.737 (0.358), 0.871 (0.274) and 0.693 (0.377), corresponding to scale 1, scale 2, maximum and intermediate compound method in order. The averages of the similarity measure were 0.322 (0.268), 0.310 (0.261), 0.363 (0.240) and 0.270 (0.250) in the same order. Those averages of two indices show that the maximum compound method brings higher quality of the scale values than others, but the differences between averages are not significant. The results reveal a similar tendency to the results obtained in the previous study [5].

As mentioned in the introduction, the results are insufficient to show that the maximum compound method is superior to the other methods. In fact the maximum compound method does not always give higher index values to all data than the others. If we could select one from the four scale values on a certain criteria, we would make the quality of the scale values, namely two index values, higher. However, the criteria for choosing the compound method must not relate to the true scale values for tallness because they are usually unknown. Therefore, the candidates for the criteria are the matching measure and the similarity measure between the scale values of two categories at which a participant rates the object on the two scales. The matching measure is not suitable for the criterion since the values take more than 0.9 at almost 70% of all data. In contrast, the values of the similarity measure range from zero to unity with almost same frequencies for each class (see Fig. 3). We thus adopt the similarity measure as the criterion.

Figure 3 shows the relations between the criterion and the two index values. The set of points shown in the figure indicates the averages of the two indices for each class of the criterion. As is evident from Fig. 3, both of the index values for maximum compound method, especially the matching measure, are higher than the others in the range that the class is less than 0.7. Meanwhile, the index values for intermediate compound method are lower than the others. As the class gets closer to unity, the differences between indices for the four scale values become smaller. It is reasonable because the criterion shows a similarity between the two scale values for categories participants elected. Therefore, the obtained results suggest that the maximum compound method is preferable to estimate the scale values despite the similarity between the scale values for the categories.

In the MUSCAT, we implicitly assume that two scale values obtained from each scale get same weights without measuring suitability of two categories, so that the important and interesting information has not been used at all. If the suitability of two categories would be measured and they would be different, for example the one is rated at "most suitable" and the other is rated at "nearly suitable, " then using only scale value for "the most suitable category" will be more suitable for estimating the true scale value than using the maximum compound method. Since the answer to above problem needs more studies, we will report them at another opportunity.

4. Conclusion

We discussed the relationship between the categories and the scale values obtained by three compound methods in the MUSCAT to clarify the general advantage of the maximum compound method. The experimental results show that the maximum compound method makes the ability to estimate the true scale values for objects higher on the assumption that the suitability of two categories rated are equal. In addition, choosing compound methods based on the suitability of the categories possibly makes the quality of the scale values higher than using the maximum compound method only.

References

[1] Yoshikawa, A. and Nishimura, T., "Psychological scaling with method of fuzzy categories and its experimental verifications, " *Journal of Japan Society for Fuzzy Theory and Systems*, vol. 5, no. 4, pp. 719–731, Aug. 1993.

[2] Guilford, J. P., *Psychometric methods*, McGrow-Hill Book, New York, 1954.

[3] Hesketh, B., Pryor, R., Gleitzman, M. and Hesketh, T., "Practical applications and psychometric evaluation of a

computerized fuzzy graphic rating scale," in *Fuzzy set in psychology*, Zetenyi, T. (ed.), North-Holland, pp. 425-454, 1988.

[4] Hesketh, T., Pryor, R. and Hesketh, B., "An application of a computerized fuzzy graphic rating scale to the psychological measurement of individual differences, " *International Journal of Man-Machine Studies*, vol. 29, pp. 21-35., 1988

[5] Yoshikawa, A. and Nishimura, T., "Psychological scaling by The Multiple-scale Technique and The Between-sets, " *IEICE Trans.* , vol. J76-D-II, no. 1, pp. 154-161, Jan. 1994.

[6] Yoshikawa, A., "Mathematical Properties of Between-sets, " *Journal of Japan Society for Fuzzy Theory and Systems*, vol. 4, no. 1, pp. 150-159, Feb. 1992.

[7] Dubois, D. and Prade, H., *Fuzzy Sets and Systems: Theory and Applications*, Academic Press, New York, 1980.