Review of Product Usability Index 2004 for Rapidly Ageing Society

H. Ogi, *Member, ISG*

***Abstract*—At the 5th International Conference of the International Society for Gerontechnology, we proposed the use of a product usability index as a new descriptor for senior citizen groups. As an example, we illustrated the use of PUI-2004.jp. Recently, we decided to review this methodology to clarify its mechanism. We observed the data collected for the surveys from various aspects and hence attempted to understand the meaning and the mechanism of PUI-2004.jp more clearly. This paper presents some sections of the review study.**

I. INTRODUCTION

UNIVERSAL design strategy should include some kind of evaluation process concerning the universality, whether by pre-occupancy or by post-occupancy. In our previous papers [1], [2], we reported the use of a product usability index (PUI) as a quantitative measure of the usability of products and as a descriptor for recording evaluation panel for of universal design. And as an example, we illustrated the use of PUI-2004.jp.

Concerning to PUI-2004.jp, we also discussed the composition method, the effect of age on the ND (not difficult) ratio (which indicates the ease of use of the product), and the results of the statistical tests on the dependency on the age factor, etc. At the 5th International Conference of the International Society for Gerontechnology, we pointed out the statistical correlation between the results of the psychological tests on cognitive functions, [3], [4], and the PUI (see Appendix D). After the conference, we investigated this correlation further. We also attempted to clarify the mechanism of PUI-2004.jp.

The PUI and its corresponding approach were originally developed to provide descriptors which are in the communicable and conscious level with senior citizens. At the time, "ageing" concerning to Universal Design is also intended in the communicable and conscious level with senior citizens, which is, so to speak, living function level. Recently, the psychological tests of cognitive functions advances [5], which describe the panel in rather unconscious and restrictedly communicable level only with psychological tools, which is, so to speak, physical and mental function level.

Hence, we attempted to modify the PUI in order to develop new descriptors, which acted not only as quantitative measures but also as qualitative ones. To begin, we reviewed the experimental data that we had obtained in our previous studies. In this paper, we report the results of the review that we conducted in order to obtain greater clarity regarding the mechanism of PUI-2004.jp.

In this study, we introduce four new types of diagrams, which are all closely related to each other. Through the explanation of these diagrams, we demonstrate the close relation between the patterns of the responses and the results of PUI-2004.jp, which is based on dual scaling techniques [6]

In order to resolve whether linear orderings of the data of the product usage and the senior citizens, who are the respondents in our study, are obtained when PIU-2004.jp is used, we compare the analytical solution obtained by dual scaling with the data model known as Guttman's complete scalable dichotomous items data model [4].

Finally, we present profile diagrams of PUI-2004.jp and explain and discuss its basic features.

II. GRAPHS SHOWING DETAILS OF PUI-2004.JP

A. Method-1: Analysis using Graphs

Fig. 1, which is shown below, comprises 6 parts. Figs. 1 (c) and (d) show the score charts that were subjected to the dual scaling techniques, which have already been explained in [2]. The other four figures are derived from these two figures.

We first explain the figures to the right and then proceed to the left.

B. Figure 1(Right Hand Side)

Fig. 1 (d) shows the scores of the respondents after dual scaling. We collected data for 1,092 respondents but only 157 responses were plotted because of the overlap of the responses.

Fig. 1 (b) shows the cumulative distribution of the responses, while also keeping in mind the overlapping responses. In this figure, we use the same information for Axis I as that used in Fig. 1 (d); but we call this figure simply as PUI-2004.jp, because the results indicated by Axis I are more stable and reliable than those of Axis II (see Appendix C).

Manuscript received April 30, 2008. This work was supported in part by the Science and Technology Promotion Expenditure of Ministry of Education, Culture, Sports, Science and Technology, a subsidy from the National Institute of Advanced Industrial Science and Technology, and the consignment expenses of the East Japan Railway Company.

Hajime Ogi is with the Institute for Human Science and Biomedical Engineering, National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan. (e-mail:h-ogi@aist.go.jp)

On the basis of Fig. 1(b), we observe the following features of PUI-2004.jp.

(1b-1) Both ends of the graph show highly overlapping data.

(1b-2) The area between the two ends of the graph is divided into nine areas of high density.

(1b-3) The cumulative response while considering the overlap is greater to the right of the graph than to the left.

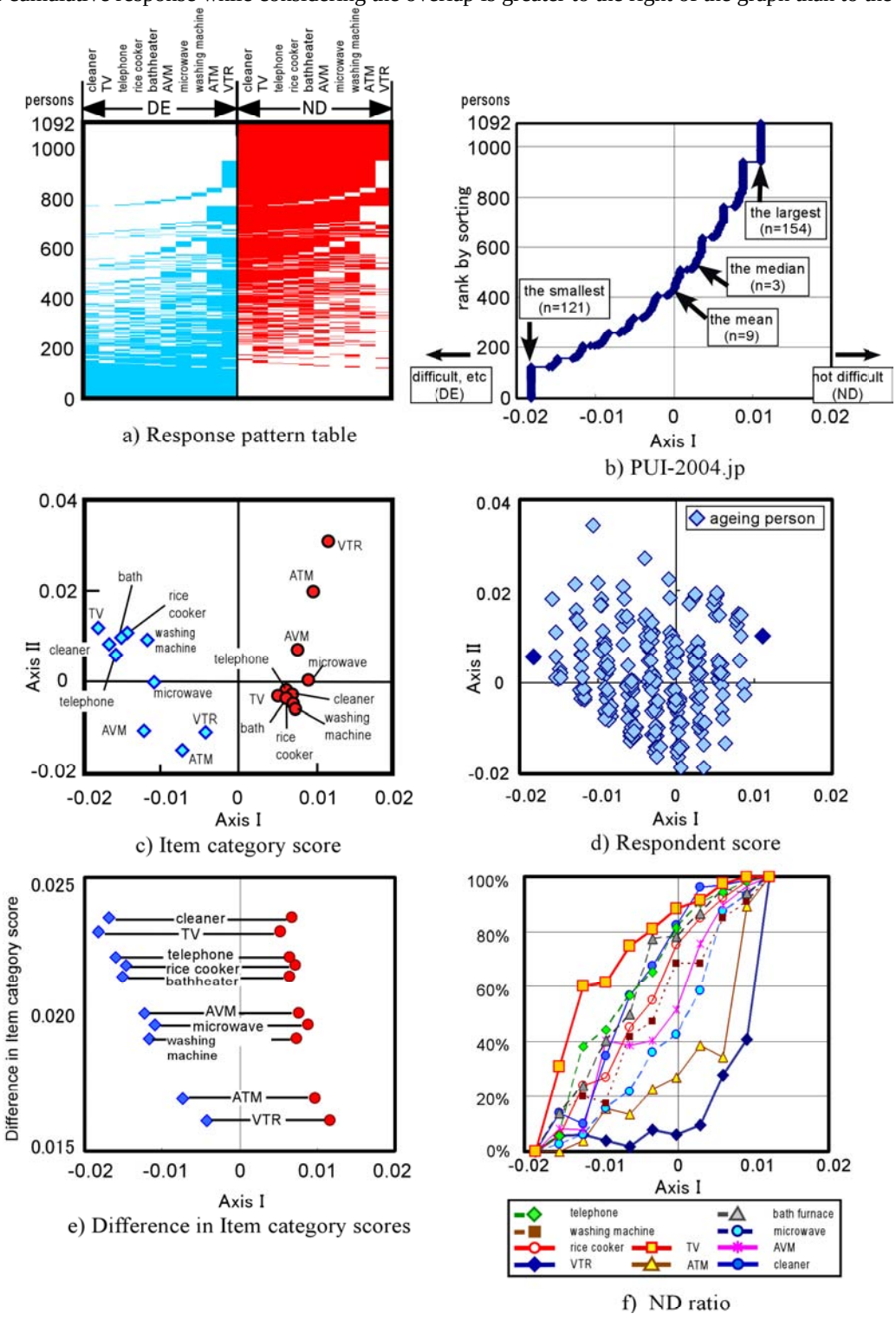

Fig.1 Relation between the response pattern table, dual scaling solutions, and PUI-2004.jp

(1b-4) The medium point is in the right side, even if the mean point is in the center of the graph.

From points (1b-1) and (1b-2), we can conclude that there are 11 high-density areas (HDAs) in the graph. In Fig. 1 (d), we can identify nine vertical rows with a slight slant. These indicate the areas between the end points of the graph in Fig 1 (b). The end points themselves are indicated by a darker color, as can be observed in Fig. 1 (d).

On the basis of points (1b-3) and (1b-4), it is clear that the responses of the population are biased towards the right.

Fig. 1 (f) shows the ND ratio of each industrial product for the eleven HDAs. Here, the term “ND ratio” implies the ratio of the respondents who responded that the product

was “not difficult” to use by oneself.

The following are the main characteristics of Fig. 1(f): (1f-1) From the extreme right of this figure, it is apparent that VTR declines the fastest, followed by ATM; however, the graphs of the other categories are not clear as they are too close to each other.

(1f-2) The results in the middle of the figure show a general decline, although they cannot be deciphered accurately due to the overlap of the graphs. The lowest ratios are those of VTR and ATM and the highest ratio is that of TV.

(1f-3) To the extreme left side, we find extreme confusion that rolls VTR, and a rapid decline in the response for TV.

C. Figure 1 (Left Hand Side)

Fig. 1 (c) shows the “item category score,” i.e., the score of the items based on category, obtained after application of the dual scaling techniques.

The item categories that we have used in this study are dichotomous. The products can be divided into two categories—ND (not difficult) and DE (difficult, etc.)The reason of adoption of the dichotomous is balance of amount of both categories, i.e. 6604 of ND and 4316 of DE responses are obtained from 1092 respondents..

As Fig. 1 (c) includes various types of information, we extract only the most important information required for PUI-2004.jp, i.e., the differences between the item category scores of the ND products and the DE products. The differences are shown in Fig. 1 (e), where the horizontal interval is used to indicate the correspondence of the results in this figure to those shown in Fig. 1 (c) and the vertical axis is used to indicate the difference between the two categories.

This importance comes from dual scaling. It can be observed that the respondent score is proportional to the item category score and depends on the selection of the product as ND or DE. In other words, the selection causes some difference to the respondent score, and the difference is proportional to the difference between the item category score of ND and the score of DE. We have briefly explained this concept in Appendix A.

The difference in the item category scores of PUI-2004.jp can be clarified by considering the following features of Fig. 1 (e).

(1e-1) The differences between ND and DE lie in the range of 0.015 to 0.025. It is apparent that the sum of any two differences is larger than any single difference. Therefore, the following relation holds for PUI-2004.jp.

$$NDN(i) \le NDN(j) \Leftrightarrow RS(i) \le RS(j) \qquad \cdots\cdots (1)$$

where *NDN(i)* implies the number of respondents *i* who selected ND as their response and *RS(i)* implies the respondent score of respondent *i*.

(1e-2) The responses can be divided into the following five groups—{VTR}, {ATM}, {washing machine, microwave, AVM}, {bath heater, rice cooker, telephone}, and {TV, cleaner}.

From (1e-1), it can be inferred that the HDAs in Fig. 1 (b) correspond to the number of ND responses by each respondent.

Fig. 1 (a) shows the response pattern, where the item categories are plotted along the horizontal axis on the basis of the order of the difference in Fig. 1 (e) and the respondents are plotted along the vertical line according to the order of the respondent score shown in Fig. 1 (b).

Using Fig. 1 (a), we can highlight several features of the response patterns of PUI-2004.jp.

(1a-1) The top of this figure shows the respondents who had ND responses for all the products included in the study.

(1a-2) The next set consists of the respondents with one DE response for VTR but ND responses for all the other products.

(1a-3) Closely following the above set of respondents are the respondents with one DE response but nine ND responses for all the other products.

(1a-4) The respondents with two DE responses for VTR and ATM but eight ND responses for all the other products are arranged below (1a-3). The several tails following the main body of the responses indicate the other possible combinations of responses.

(1a-5) The respondents with three DE responses are shown in a similar manner, although the set is much smaller.

(1a-6) The set of the respondents with more than three DE responses is extremely small and appears as an extremely thin pattern in the graph. Moreover, since the variety of combinations is large, it is difficult to find regular patterns.

(1a-7) The bottom of the figure indicates the respondents with DE responses for all the products.

III. PUI-2004.JP AND ITS LINEAR AND STRUCTURAL ORDERING

A. Method-2: Analysis Using Data Model 1: Guttman’s Complete Scalable Dichotomous Items Data Model

From the features of the response pattern, i.e., from (1a-1), (1a-2), and (1a-4), we can easily develop a model based on Guttman’s complete scalable dichotomous item data model [4], which is illustrated in Fig. 2. The ordering of the products is based on the item category scores of PUI-2004.jp.

B. Analysis based on Fig. 1

This type of data model implies a linear and structural ordering of the item categories and the respondents. Then, using Fig.1 and Fig. 2, we investigate the mechanism of PUI-2004.jp from the viewpoint of linear and structural ordering.

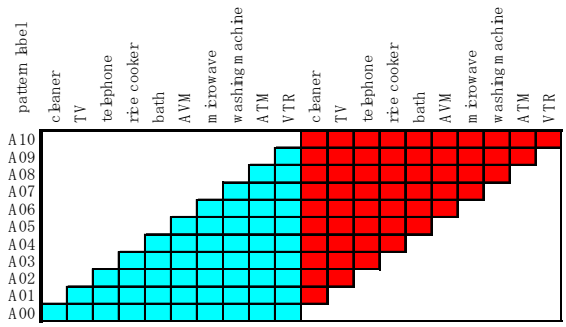

Fig. 2 Guttman's complete scalable dichotomous items data model

In general, Figs. 1 (e) and (f) display a linear ordering of the item categories but they include a little bit different shape of ordering in the detail. In Fig. 1 (e), we observe a linear ordering of the differences between ND and DE but we also observe groups, as described in (1e-2). Fig. 1 (f) shows the linear ordering of the ND ratios of VTR, ATM, TV, etc.; however, we need to bear in mind the confusion in the results due to the overlapping of the graphs, as mentioned in (1f-1) to (1f-3).

With regard to the ordering of the respondents, (1a-1), (1a-2), and the sets of respondents mentioned in (1a-4) to (1a-6) show linear ordering; however, some contradiction is present in the form of the tails mentioned in (1a-3) and (1a-4), and the thinness of the body and the different combinations mentioned in (1a-5) and (1a-6).

Therefore, it is not entirely clear whether PUI-2004.jp shows linear and structural ordering.

C. *From the View Point of the Eigenvalue*

The following equation holds true for Guttman's complete scalable dichotomous item data model [7]:

$$\lambda = \frac{(n+1)}{2n} \qquad \cdots\cdots\cdots (2)$$

where

n = half number of item categories

$(n = 10, \quad \text{in PUI - 2004.jp})$

From equation 2, it is apparent that the eigenvalue λ should be larger than 0.5; however, the eigenvalue in case of PUI-2004.jp is approximately 0.49.

Consequently, from the view point of the eigenvalue, it is also not easy to judge whether PUI-2004.jp shows linear and structural ordering .

IV. Using the Ordering Data Model as a Feature

A. *Method-3: Analysis Using Data Model 2: Linear Ordering with Five Groups*

In the above section, we have clarified the mechanism of PUI-2004.jp and discussed the possibility of linear ordering. In this section, we examine the response patterns and the linear ordering in more detail.

Considering the five groups mentioned in (1e-2), we modified the data model as shown in Fig. 3. The response pattern labeled "Axx" is based on Fig. 2, while the response pattern labeled "Bxx" depends on the combination of the groups ("xx" denotes the number of ND responses).

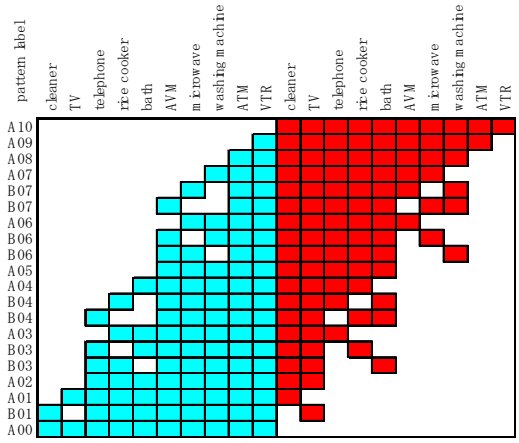

Fig. 3 Data model 2: linear ordering with five groups

B. *Profile of Data Model 2*

The percentage of the respondents with the response patterns of "Axx" and "Bxx" according to Data Model 2 was 52% (including 8% of type Bxx) in PUI-2004.jp. However, with the exception of respondents with A10 and A00, the percentage decreased to 36% (including 11% of Bxx). We labeled the response patterns that do not belong to Data Model 2 as "Cxx." We observed that there were 260 types of response patterns in PUI-2004.jp, of which 93% were of the type "Cxx."

Fig. 4 shows the constituent ratio of the response patterns with linear ordering. The vertical axis is the same as that used in Figs. 1 (a) and (b), but the unit for this axis is the percentile in respondents.

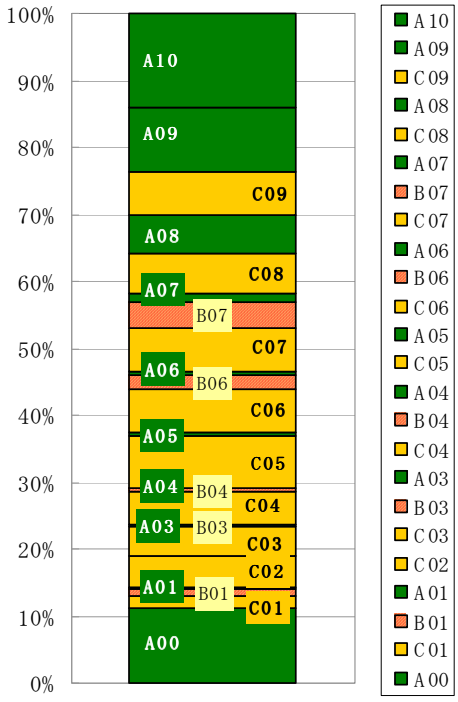

Fig. 4 Constituent ratio of types of response pattern

From the figure, we can observe that the percentage of A09 is more than C09 and that the percentage of A08 is the same as that of C08. However, the sum of A07 and B07 is less than the value of C07. We also observe that as the

number of ND responses decreases, the linear ordering of the responses also reduces.

C. Linear Ordering Feature of the Data Model

We examine the type of respondents whose responses resulted in the response patterns shown in Data Model 2. Fig. 5 roughly illustrates the remarkable difference between the four surveys (see Appendix B).

The first survey and the third survey both depended on the "age" factor, whereas the second survey and the fourth survey were independent of this factor [2]. Therefore, we can distinguish the survey panels by considering the effect of the age factor on PUI.

In this study, we have used two new tools, i.e., profiles of number of ND responses and profiles of the data model.

Based on the profiles of the ND responses, it can be observed that the second survey and fourth survey were similar to each other. However, there was a considerable difference between the results of the first survey and the third survey. It should be noted that the third survey differed from the first only at the for the ND number 0.

The profiles of the data model further clarify the difference between the first and the third surveys. In the first survey, the percentage of the response patterns according to Data Model 2 was 64 % (including 15 % of type B) but in the third, this percentage was 30% (including 7% of B). The percentages of respondents with the response patterns of the second and fourth surveys were roughly the same and were approximately 65% (including 1-3% of B).

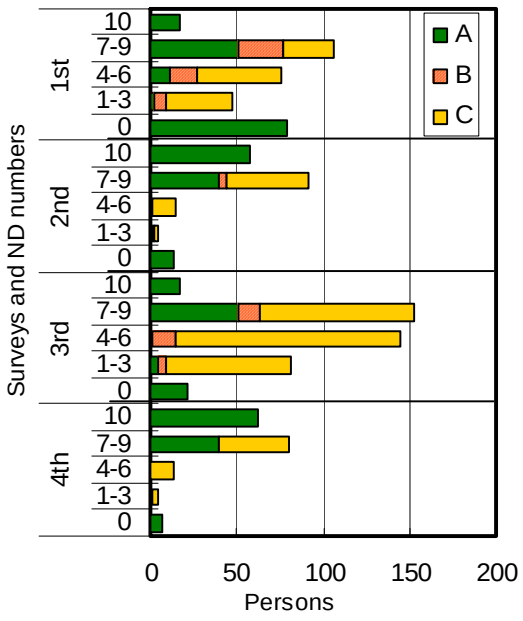

Fig. 5 Differences in surveys shown by the response patterns

V. Future Outlook

A. Summary

We can summarize the results of this paper on the basis of its positive and skeptical aspects as follows:

Positive aspects:

(i) We clarified the mechanism of PUI-2004.jp.

(ii) Using the feature of PUI-2004.jp listed in (1e-1), we derived the following equation:

$$NDN(i) \le NDN(j) \Leftrightarrow RS(i) \le RS(j).$$

(iii) On the basis of the rearranged shape of the response pattern table shown in Fig. 1 (a), we determined the relation of the linear and structural ordering as Fig. 3 hidden in PUI-204.jp.

(iv) It is clarified that concerning the total respondents of PUI-2004.jp, it is unable to determine whether there is the linear and structural ordering, even if as a numeric descriptor, PUI-2004.jp gives apparent ordering.

(v) The state of inclusion of respondents with the response patterns appeared in data models of the linear ordering is usable as a marker for the survey panels.

Skeptical aspects:

(vi) We have not determined which is more essential to describe the evaluation panels for universal design yet, respondents with the response patterns with the linear and structural ordering or the respondents with the response patterns without linear and structural ordering.

(vii) We have not made any attempt to apply data models to clarify the statistical correlation between the psychological tests on cognitive functions and the statistics among PUI yet

B. Future Studies

Further investigations are required to resolve the doubts regarding the negative aspects of PUI-2004.jp.

Still now, we have not any good estimation concerning to the more experiments (see Appendix C). We plan to continue our investigations on PUI-2004.jp and develop new and better descriptors.

Appendix

A. Interpretation of the difference between the respondent scores and the dichotomized item category scores

The following relation exists between the "item category score" $\boldsymbol{x}$ and the "respondent score" $\boldsymbol{y}$ of dual scaling [6].

$$\boldsymbol{y} = \frac{1}{n} D_n^{-1} F \boldsymbol{x} \qquad \cdots\cdots (3)$$

where

$\boldsymbol{y}$ = respondent score (vector, length m)

D_n = diagonal matrix (matrix, $m \times m$)

F = data matrix (matrix, $m \times n$)

$\boldsymbol{x}$ = item category score (vector, length n)

η = square root of correlation ratio

m = number of respondents

n = number of item categories

Then, score y_i of respondent i with response f_i becomes

$$y_i = \frac{1}{\eta} \frac{1}{n_i} \boldsymbol{f}_i \boldsymbol{x}, \qquad \cdots\cdots (4)$$

where

y_i = score of respondent i

n_i = number of response of respondent i

$\boldsymbol{f_i}$ = response of respondent i (1/0 vector, length n_i)

It should be noted that DE includes the response "No Answer." In the above equation

$$n_i = \frac{n}{2} = \text{constant} \qquad \cdots\cdots(5)$$

and y_i is determined simply by the summation of the switching functions with scores as follows:

$$y_i = c\sum_p \begin{cases} s_{p\ ND} \cdots \text{if response of } i \text{ is ND for p} \\ s_{p\ DE} \cdots \text{if response of } i \text{ is DE for p} \end{cases} \qquad \cdots(6)$$

where

$s_{p\ ND}$ = item category score of "ND for p"

$s_{p\ DE}$ = item category score of "DE for p"

c = constant

p = each product given for selection

Thus, the selection causes some difference to the respondent score and the difference is proportional to the difference between $s_{p\,ND}$ and $s_{p\,DE}$.

B. *Outline of Four Surveys*

In this section, we provide an outline of the surveys used in this study, i.e., the objectives of the surveys, the methods of distribution and collection of the questionnaires, the population density in the areas of the survey, the years of practice, and the number of respondents.

(i) The first survey was conducted on housing in an agricultural area during the FY93–95 in an area with low population density. The survey questionnaires were distributed in advance and were collected by visiting the respondents. A total of 326 respondents were surveyed.

(ii) The second survey was conducted on outdoor behavior in meeting rooms in the FY97–99. A total of three rooms were considered with medium population density. A total of 182 respondents were surveyed.

(iii) The third survey was conducted in the FY01-02 on the use of industrial products. The surveys were mainly conducted in meeting rooms. A total of seven rooms were used. Partly the survey was conducted by layaway in institutions for senior citizens and collection by standard mail, The surveys were held in five institutions. Partly, the questionnaires were distributed and collected all over Japan by standard mail. A total of 417 respondents were surveyed.

(iv) The fourth survey was conducted in FY03 on the use of the railway station. The surveys were conducted in meeting rooms and a total of five rooms were used. The population density was medium to high. A total of 167 respondents were surveyed.

C. *Conformity through Correlation Coefficients*

This section of the appendix is related to the content of our previous papers [1], [2].

Fig. A shows the correlation coefficients of PUI-2004.jp to the dual scaling scores of 46 subgroups of PUI-2004.jp. The correlation coefficients are for the scores indicated by Axis I and Axis II and are arranged by the logarithmic scale of the group size.

In this figure, we can see the shape of the convergence as the size of group. The correlation coefficients of the scores of Axis I rapidly converge to 1. On the other hand, the correlation coefficients of the scores of Axis II converge slowly to 1. The difference in the speed of the convergence appears to be approximately within one graduation of the logarithmic scale, i.e., 10 times.

Therefore, we estimate that the 10 times of survey respondents may fix the features of Axis II.

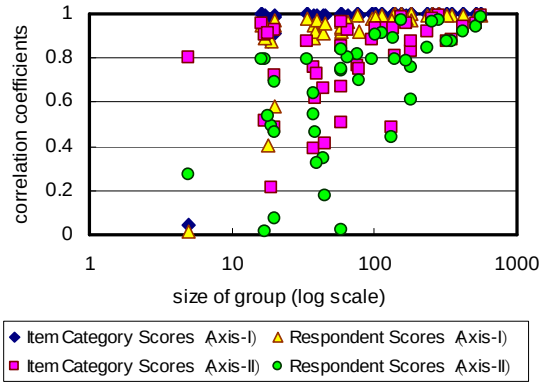

Fig. A Correlation coefficients of PUI-2004.jp to Dual Scaling scores for 46 sub-groups of the respondents of PUI-2004.jp

D. *Correlation between Tests on Cognitive Functions and PUI*

At the 5th International Conference of the International Society for Gerontechnology, we pointed out the correlation between the tests on cognitive functions [3], [4], and PUI. The correlation is shown in Table A. The notations for the tests on cognitive functions are modified according to [5]. The survey described in this table is based on PUI-2004.jp, which is the fifth survey. However, the PUI and RSUI (Railway Station Usability Index) are calculated using a combination of the results of PII-2004.jp and the fourth survey.

TABLE A. CORRELATION COEFFICIENTS AMONG PUI, RSUI, AND THE TESTS ON COGNITIVE FUNCTIONS

Usability Index	PUI		RSUI	
Tests on Cognition	cor.coef	t value	cor.coef	t value
Planning-1(Seq)	-0.2571	-3.28**	-0.2581	-2.66**
Planning-2(Switch)	-0.3195	-4.16**	-0.3969	-4.30**
Attention	-0.1868	-2.34*	-0.3287	-3.46**
Working Memory	-0.2046	-2.57*	-0.3186	-3.34**
comment:	n=154, df=152		n=101, df=99	

REFERENCES

[1] H. Ogi, "Products usability index and its dependency on the age factor", *Gerontechnology* 2005;3:232

[2] H. Ogi, *Proc. 5th Conference on Gerontechnology*, Nagoya, 2005, PS4a-6

[3] M. Kitajima, T. Kumada, M. Akamatsu, H. Ogi, and H. Yamazaki, “Effect of cognitive ability deficits on elderly passengers’ mobility at railway stations: Focusing on attention, working memory, planning”, *Gerontechnology* 2005;3:231
[4] M. Kitajima, T. Kumada, H. Ogi, M. Akamatsu, H. Yamazaki, *Proc. 5th Conference on Gerontechnology*, Nagoya, 2005, PSa-4.
[5] M. Kitajima, T. Kumada, H. Ogi, M. Akamatsu, H. Tahira, H. Yamazaki, “Usability of Guide Signs at Railway Stations for Elderly Passengers – Focusing on Planning, Attention, and Working Memory”, *Japanese Journal of Ergonomics*, to be published.
[6] S. Nishisato, *Analysis of Categorical Data: Dual Scaling and Its Applications.* Toront Buffalo London, University of Toront Press, 1980, pp. 54-59.
[7] S. Iwatsubo, *Data Analysis and Informatics* Amsterdam, New York, Oxford, North-Holland, 1983, ch. 13, pp. 31–40.

Keywords: usability, universal design, sampling, dual scaling, PUI