Application of Rasch Model in Validating the Content of Measurement Instrument for Blog Quality

Zuhaira Muhammad Zain[1], Abdul Azim Abd Ghani[1], Rusli Abdullah[1], Rodziah Atan[1], and Razali Yaakob[2]

[1] Department of Information Science,
Universiti Putra Malaysia
Serdang, Selangor
`zuhaira.muhdzain@gmail.com`
[2] Department of Computer Science,
Universiti Putra Malaysia,
Serdang, Selangor
`{azim,rusli,rodziah,razaliy}@fsktm.upm.edu.my`

Abstract. Research in blog quality is very crucial nowadays in order to have a good quality blog in the blogosphere. The blog quality criteria have been derived from a rigorous metadata analysis. Yet, these criteria have not been reviewed and their significance has not been proven systematically. In this paper, Rasch Model is applied to produce an empirical evidence of content validity of the blog quality criteria. This study confirms that the definitions of 11 families and the 49 criteria assigned have content validity by mean of online survey. These criteria will then be used as a basis of constructing the instrument to measure the acceptability of the criteria for blog quality.

Keywords: Blog quality, content validity, Rasch Model.

1 Introduction

Advances in technology are making use of Internet as an ever-growing phenomenon and we are witnessing a tremendous growth of blogs in the blogosphere. As reported by Pew Internet & American Life Project Surveys in 2006, there were 12 million of American adults who keep a blog [1]. We believe that the current figure has risen, due to the increased of broadband penetration rate that has exposed all levels of users to blogging. As stated in the World Broadband Statistics Report, by the end of 2009, there were 466.95 million broadband subscribers worldwide, up to 2.5 per cent on the previous quarter from 455.57 million [2]. In addition, the technical skills required to create a blog are readily available. The emergence of user-friendly and free blog tools such as Blogger, WordPress, and LiveJournal simplifies the process of web publishing, which formerly required some programming skills [3]. Blogs may offer vital information for blog readers to make decisions or keep abreast with the latest blog posting. Yet, the growth presents disorganized and uncontrolled, thus contributing to the limitation of having bad blogs in the blogosphere. Consequently, blog readers

have access to bad or poor-quality blogs that may turn off readers' interest to return to the same blog. Fulfilling readers' expectations and needs is necessary for developing good quality blog.

This study endeavoured into the criteria required in blog quality that will promote to readers satisfaction. The blog quality criteria have been derived from a thorough metadata analysis. The criteria have been consolidated from literatures and researches done in the area of blog, website design, information quality on the Web and portal data quality. However, these criteria have not been reviewed and their content validity has not been verified systematically.

Content validity test is a subjective assessment of how suitable an instrument is to a group of reviewers. It involves a systematic review of the survey's contents to ensure that it includes everything it should, and excludes everything it should not. It is also very important in providing a good foundation on which to base a rigorous assessment of validity [4]. However, Kitchenham and Pfleeger [4] argued that there is no content validity statistic. Later, this argument has been refuted by Abdul Aziz et al. [5] where they proved that content validity can be assessed accurately and fast despite the small sample size by using Rasch Measurement Model. It is a measurement model that is formed as a result of the consideration that takes into account the ability of the respondents, and the difficulty of items [6]. The graphical output provided is great which gives better clarity for quick and easy decision making [7].

The purpose of this study is to determine content validity of the measurement instrument for blog quality. It is to confirm whether the prescribed family definitions and criteria assigned are agreeable to the reviewers. This study proves that the definitions of 11 families and the 49 criteria assigned have content validity by mean of online survey.

The rest of this paper is organized as follows. Section 2 describes the basic of Rasch measurement method. Section 3 explains how content validity test is conducted. Section 4 discusses the results of this study. Finally, Section 5 touches on the conclusions and future work.

2   Rasch Measurement Method

Response from the experts on the content validity is considered rating scale in which the experts rated the criteria according to their agreement. In theory, at this phase, the study is only counting the number of positive answers from the experts which is then added up to give a total raw score. The raw score only provides a ranking order which is supposed an ordinal scale that is continuum in nature [8]. It does not have equal intervals which contradicts the nature of numbers for statistical analysis and it does not meet the fundamentals of adequate statistics for evaluation [9].

Rather than fitting the data to suit a measurement model with errors, Rasch focuses on developing the measurement instrument with accuracy. By emphasizing on the reproducibility of the latent trait measurement instead of forcing the expected generation of the same raw score, i.e. the common expectation on repeatability of results

being a reliable test, the concept of reliability takes its rightful place in supporting validity rather than being in contentions. Consequently, measuring quality in an appropriate way is vital to ensure valid quality information can be produced for meaningful use; by absorbing the error and representing a more accurate prediction based on a probabilistic model [10].

In Rasch measurement Model, the probability of success can be estimated for the maximum likelihood of an event as;

$$P(\theta) = \frac{e^{\beta_v - \delta_i}}{1 + e^{\beta_v - \delta_i}} \tag{1}$$

where;

e = base of natural logarithm or Euler's number; 2.7183
β_n = person's ability
δ_i = item or task difficulty

3 Methodology

Experts from the field of English and Information Technology who read blogs (comprising 50 English Language lecturers from various universities in Malaysia and 50 Information Technology executives or managers from various companies of more than 10 years working experience) form the pool of reviewers. Therefore, the total number of experts engaged at 100.

The descriptive design chosen is based on the survey objective. It is suitable for checking whether the experts agree with the proposed set of definition of family and the assigning of criteria to the respective families. The questionnaire consists of closed questions (with a Yes/ No answer being provided) and also an open question for differing views and comments. In this context, the endorsement of the experts is of interest.

Invitation by e-mail to join the online survey is being sent to the subjects. The objective of the study; its relevance; the importance of individual's participation; and the confidentiality assured, have been made known for the purpose of the survey. The feedback was then tabulated and analyzed using Rasch Measurement Model [11] with the aid of Rasch analysis software [12].

4 Results and Discussions

A total of 60 participants out of 100 subjects submitted the online survey forms (30 respondents being experts in English with another 30 from Information Technology). This represents a response rate of 60 percent. The responses are then tabulated and run in Bond&FoxSteps software in order to obtain the *logit* measures.
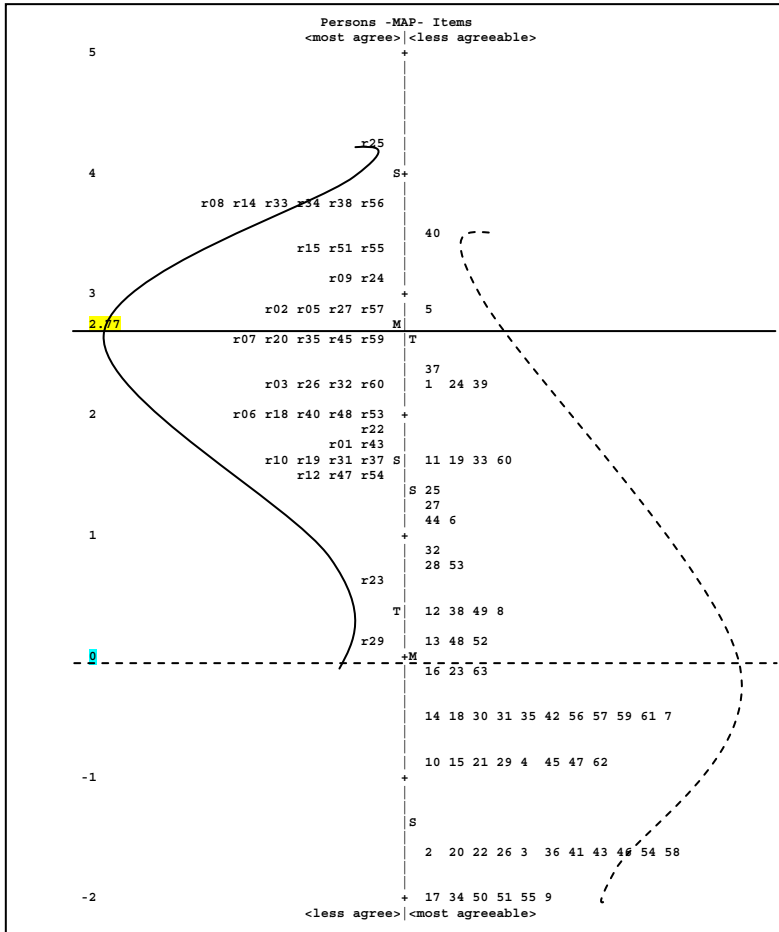
Fig. 1. PIDM

The Person-Item Distribution Map (PIDM) as shown in Fig. 1 portrays experts' agreement and the acceptability of the 11 family definitions with the 52 criteria assigned. Acceptability Test Level of Agreement given by experts for the identified family and blog quality criteria can be easily established by using formula in Equation 1:

$$P(\theta) = \frac{e^{2.77-0}}{1+e^{2.77-0}}$$

$$= 0.941$$

Thus, experts have indicated their Level of Agreement at 94.1% which is above the 70% threshold limit of Cronbach Alpha. Hence, all experts agree to the prescribed family definition and criteria assigned. This can be determined clearly from PIDM where the person mean μ_{person}=+2.77 *logit* is located higher than the item mean; μ_{item}

which is constrained to 0.00 *logit*. This indicates that all experts involved in the Content Validity Test have the tendency for agreeing to the entire family definitions and criteria assigned that have been proposed.

The Summary Statistics in Fig. 2 shows three important indicators; Cronbach Alpha value, Person Reliability value, and Item Reliability value. The summary shows that there is a good value of Cronbach Alpha (0.87). In addition, there is a fair value of Person Reliability (0.74) and Item Reliability (0.74). The values of the three indicators (> 0.6) do confirm that the instrument for measuring content validity is reliable, reproducible, and valid for measurement. The summary also depicts that there are 12 persons and 6 items with maximum and minimum extreme score respectively. This means that there are 12 experts who agreed with all the definitions and criteria assigned. Also we have 6 items which are 100% agreed by the experts.

```
SUMMARY OF 48 MEASURED (NON-EXTREME) Persons
+--------------------------------------------------------------------------------+
|            RAW                          MODEL      INFIT        OUTFIT          |
|           SCORE     COUNT   MEASURE     ERROR   MNSQ  ZSTD   MNSQ  ZSTD         |
|--------------------------------------------------------------------------------|
| MEAN      49.3      57.0     2.77        .54    1.01    .1   1.28    .2         |
| S.D.       5.4       .0      1.18        .22     .16    .6   1.64   1.0         |
| MAX.      56.0      57.0     4.98       1.04    1.32   1.6   9.90   3.2         |
| MIN.      30.0      57.0      .07        .31     .61  -1.8    .17  -1.6         |
|--------------------------------------------------------------------------------|
| REAL RMSE   .60 ADJ.SD  1.02  SEPARATION  1.70  Person RELIABILITY  .74         |
| MODEL RMSE  .58 ADJ.SD  1.03  SEPARATION  1.78  Person RELIABILITY  .76         |
| S.E. OF Person MEAN = .17                                                       |
+--------------------------------------------------------------------------------+
MAXIMUM EXTREME SCORE:      12 Persons

     SUMMARY OF 60 MEASURED (EXTREME AND NON-EXTREME) Persons
+--------------------------------------------------------------------------------+
|            RAW                          MODEL      INFIT        OUTFIT          |
|           SCORE     COUNT   MEASURE     ERROR   MNSQ  ZSTD   MNSQ  ZSTD         |
|--------------------------------------------------------------------------------|
| MEAN      50.8      57.0     3.46        .80                                    |
| S.D.       5.7       .0      1.74        .56                                    |
| MAX.      57.0      57.0     6.23       1.86                                    |
| MIN.      30.0      57.0      .07        .31                                    |
|--------------------------------------------------------------------------------|
| REAL RMSE   .99 ADJ.SD  1.44  SEPARATION  1.45  Person RELIABILITY  .68         |
| MODEL RMSE  .98 ADJ.SD  1.44  SEPARATION  1.47  Person RELIABILITY  .68         |
| S.E. OF Person MEAN = .23                                                       |
+--------------------------------------------------------------------------------+
Person RAW SCORE-TO-MEASURE CORRELATION = .91
CRONBACH ALPHA (KR-20) Person RAW SCORE RELIABILITY = .87

     SUMMARY OF 57 MEASURED (NON-EXTREME) Items
+--------------------------------------------------------------------------------+
|            RAW                          MODEL      INFIT        OUTFIT          |
|           SCORE     COUNT   MEASURE     ERROR   MNSQ  ZSTD   MNSQ  ZSTD         |
|--------------------------------------------------------------------------------|
| MEAN      41.5      48.0      .00        .62    1.00    .1   1.25    .3         |
| S.D.       6.8       .0      1.33        .24     .11    .4   1.60    .9         |
| MAX.      47.0      48.0     3.51       1.03    1.27   2.0   9.90   3.8         |
| MIN.      17.0      48.0    -1.67        .33     .75   -.9    .14  -1.0         |
|--------------------------------------------------------------------------------|
| REAL RMSE   .68 ADJ.SD  1.14  SEPARATION  1.67  Item  RELIABILITY  .74         |
| MODEL RMSE  .66 ADJ.SD  1.15  SEPARATION  1.73  Item  RELIABILITY  .75         |
| S.E. OF Item MEAN = .18                                                        |
+--------------------------------------------------------------------------------+
MINIMUM EXTREME SCORE:      6 Items
UMEAN=.000 USCALE=1.000

     SUMMARY OF 63 MEASURED (EXTREME AND NON-EXTREME) Items
+--------------------------------------------------------------------------------+
|            RAW                          MODEL      INFIT        OUTFIT          |
|           SCORE     COUNT   MEASURE     ERROR   MNSQ  ZSTD   MNSQ  ZSTD         |
|--------------------------------------------------------------------------------|
| MEAN      42.1      48.0     -.28        .74                                    |
| S.D.       6.7       .0      1.52        .43                                    |
| MAX.      48.0      48.0     3.51       1.85                                    |
| MIN.      17.0      48.0    -2.91        .33                                    |
|--------------------------------------------------------------------------------|
| REAL RMSE   .86 ADJ.SD  1.26  SEPARATION  1.45  Item  RELIABILITY  .68         |
| MODEL RMSE  .85 ADJ.SD  1.26  SEPARATION  1.49  Item  RELIABILITY  .69         |
| S.E. OF Item MEAN = .19                                                        |
+--------------------------------------------------------------------------------+
```

Fig. 2. Summary statistic

The Item Measure in Fig. 3 lists the family and criteria in an ascending order of their acceptability level. Close study reveals that the 6 items with extreme score or have minimum estimated *logit* measure are: *Availability of blog*; *Easy to read info*; *Clear layout of info; Family of Readability; Family of Info Representation; Family of Currency;* and *Appropriate explanatory text.*

```
+-----------------------------------------------------------------------------------+
|ENTRY   RAW                    MODEL|  INFIT  |  OUTFIT |PTMEA|EXACT MATCH|          |
|NUMBER  SCORE  COUNT  MEASURE   S.E.|MNSQ ZSTD|MNSQ ZSTD|CORR.| OBS%  EXP%| Item     |
|-----------------------------------------------------------------------------------|
|   40     17     48    3.51     .34| .95  -.2|1.40  1.6| .66| 75.0  74.4| 40-Must-have sound        |
|    5     23     48    2.84     .33| .93  -.5| .94  -.3| .62| 68.8  70.4| 5-Relevant info           |
|   37     27     48    2.41     .33| .99   .0|1.10   .5| .54| 66.7  69.9| 37-Info in diff format    |
|   24     28     48    2.30     .33| .90  -.7| .78 -1.0| .59| 68.8  69.9| 24-Easy to remember add   |
|   39     28     48    2.30     .33| .91  -.7| .81  -.8| .58| 72.9  69.9| 39-Must-have photos       |
|    1     29     48    2.19     .33|1.27  2.0|1.48  1.8| .39| 62.5  70.0| 1-F.Accuracy              |
|   11     34     48    1.62     .35| .89  -.7| .76  -.7| .51| 77.1  73.8| 11-Avail. Blog owner info |
|   19     34     48    1.62     .35| .86  -.9| .68 -1.0| .53| 72.9  73.8| 19-Real-occurences info   |
|   33     34     48    1.62     .35|1.03   .3|1.01   .1| .44| 68.8  73.8| 33-Technorati rank        |
|   60     34     48    1.62     .35|1.17  1.1|1.08   .3| .39| 64.6  73.8| 60-Chat box               |
|   25     36     48    1.37     .36| .97  -.1|1.54  1.4| .41| 79.2  77.1| 25-Emotional support      |
|   27     37     48    1.23     .37|1.15   .8| .91  -.1| .37| 72.9  78.9| 27-Personal feel          |
|    6     38     48    1.08     .39|1.07   .4|1.18   .5| .35| 79.2  80.8| 6-Originality             |
|   44     38     48    1.08     .39|1.01   .1| .97   .1| .38| 83.3  80.8| 44-Interactivity          |
|   32     39     48     .93     .40| .91  -.4| .69  -.6| .43| 85.4  82.6| 32-Rewarding experience   |
|   28     40     48     .76     .42|1.00   .1| .82  -.2| .37| 83.3  84.3| 28-Surprises              |
|   53     40     48     .76     .42|1.06   .3| .93   .0| .34| 83.3  84.3| 53-Readable font          |
|    8     42     48     .38     .46|1.03   .2|1.10   .4| .30| 85.4  87.9| 8-Amount of info          |
|   12     42     48     .38     .46|1.04   .2| .79  -.1| .31| 85.4  87.9| 12-Easy to understand     |
|   38     42     48     .38     .46|1.02   .2| .86   .0| .31| 89.6  87.9| 38-Multimedia             |
|   49     42     48     .38     .46|1.08   .4|1.37   .7| .25| 89.6  87.9| 49-Intuitive interface    |
|   13     43     48     .15     .50| .97   .0| .59  -.4| .33| 87.5  89.7| 13-Informative            |
|   48     43     48     .15     .50| .98   .1|1.21   .5| .28| 91.7  89.7| 48-Good use of colours    |
|   52     43     48     .15     .50| .93  -.1| .66  -.3| .33| 91.7  89.7| 52-Legibility             |
|   16     44     48    -.13     .55|1.13   .4| .78   .0| .23| 91.7  91.7| 16-Link to info           |
|   23     44     48    -.13     .55|1.10   .4| .77   .0| .25| 91.7  91.7| 23- Cognitive advancement |
|   63     44     48    -.13     .55| .75  -.5| .37  -.7| .38| 91.7  91.7| 63-Trackback              |
|    7     45     48    -.46     .62|1.18   .5|1.30   .6| .14| 93.8  93.8| 7- F.Completeness         |
|   14     45     48    -.46     .62| .81  -.3|1.10   .4| .28| 93.8  93.8| 14-Objective info         |
|   18     45     48    -.46     .62|1.11   .4| .85   .2| .20| 93.8  93.8| 18-Real time info         |
|   30     45     48    -.46     .62|1.10   .4| .74   .1| .21| 93.8  93.8| 30-Reputation of blog     |
|   31     45     48    -.46     .62| .90   .0| .79   .1| .26| 93.8  93.8| 31- Reputation of blogger |
|   35     45     48    -.46     .62| .97   .1| .54  -.2| .27| 93.8  93.8| 35-Exciting content       |
|   42     45     48    -.46     .62|1.00   .2| .84   .2| .24| 93.8  93.8| 42-Ease of ordering       |
|   56     45     48    -.46     .62| .87  -.1| .48  -.3| .30| 93.8  93.8| 56-Blog responsiveness    |
|   57     45     48    -.46     .62| .87  -.1| .48  -.3| .30| 93.8  93.8| 57-Ease of info access    |
|   59     45     48    -.46     .62| .91   .0| .82   .2| .26| 93.8  93.8| 59-Blogroll               |
|   61     45     48    -.46     .62| .93   .0| .43  -.4| .30| 93.8  93.8| 61-Comment field          |
|    4     46     48    -.92     .75| .91   .1|7.69  2.8| .10| 95.8  95.9| 4-Reliable source         |
|   10     46     48    -.92     .75|1.13   .4|2.48  1.3| .08| 95.8  95.9| 10-Appropriate level      |
|   15     46     48    -.92     .75|1.14   .4|1.34   .7| .11| 95.8  95.9| 15-Provide info sources   |
|   21     46     48    -.92     .75|1.08   .3| .72   .1| .17| 95.8  95.9| 21-F.Engaging             |
|   29     46     48    -.92     .75|1.12   .4|1.44   .7| .12| 95.8  95.9| 29-F.Reputation           |
|   45     46     48    -.92     .75|1.08   .3| .76   .2| .17| 95.8  95.9| 45-F.Visual design        |
|   47     46     48    -.92     .75|1.05   .3| .57   .0| .20| 95.8  95.9| 47-Clear layout           |
|   62     46     48    -.92     .75| .83  -.1| .36  -.3| .28| 97.9  95.9| 62-Search tool            |
|    2     47     48   -1.67    1.03| .80   .1| .14  -.6| .25| 97.9  97.9| 2-Correct info            |
|    3     47     48   -1.67    1.03|1.11   .4|4.61  1.9|-.02| 97.9  97.9| 3-Reliable info           |
|   20     47     48   -1.67    1.03| .80   .1| .14  -.6| .25| 97.9  97.9| 20-Up-to-date             |
|   22     47     48   -1.67    1.03|1.09   .4|1.52   .8| .06| 97.9  97.9| 22- Appreciate comments   |
|   26     47     48   -1.67    1.03|1.12   .4|9.90  3.8|-.11| 97.9  97.9| 26-Fun                    |
|   36     47     48   -1.67    1.03| .91   .2| .21  -.4| .22| 97.9  97.9| 36-Fresh perspective      |
|   41     47     48   -1.67    1.03|1.08   .4|1.04   .5| .09| 97.9  97.9| 41-F.Navigation           |
|   43     47     48   -1.67    1.03|1.10   .4|2.41  1.2| .03| 97.9  97.9| 43-Easy to navigate       |
|   46     47     48   -1.67    1.03| .91   .2| .21  -.4| .22| 97.9  97.9| 46-Attractive layout      |
|   54     47     48   -1.67    1.03|1.09   .4|1.52   .8| .06| 97.9  97.9| 54-F.Blog accessibility   |
|   58     47     48   -1.67    1.03| .80   .1| .14  -.6| .25| 97.9  97.9| 58-F.Blog Tech Features   |
|    9     48     48   -2.91    1.85| MINIMUM ESTIMATED MEASURE |     9-Appropriate exp. text |
|   17     48     48   -2.91    1.85| MINIMUM ESTIMATED MEASURE |    17-F.Currency            |
|   34     48     48   -2.91    1.85| MINIMUM ESTIMATED MEASURE |    34-F.Info representation  |
|   50     48     48   -2.91    1.85| MINIMUM ESTIMATED MEASURE |    50-F.Readability          |
|   51     48     48   -2.91    1.85| MINIMUM ESTIMATED MEASURE |    51-Easy to read info      |
|   55     48     48   -2.91    1.85| MINIMUM ESTIMATED MEASURE |    55-Availability of blog   |
|-----------------------------------------------------------------------------------|
| MEAN   42.1   48.0    -.28     .74|1.00   .1|1.25   .3|     88.3  88.8|          |
| S.D.    6.7     .0    1.52     .43| .11   .4|1.60   .9|     10.2   9.3|          |
+-----------------------------------------------------------------------------------+
```

Fig. 3. Item Measure

By looking at the Point-Measure Correlation (see column titled PTMEA CORR.), it can be found that 15 items are in the acceptable range; $0.32<x<0.80$. On the other hand, 37 items fall outside the range. A further verification for these items is done by looking at the OUTFIT column for MNSQ, y value; $0.5<y<1.5$. Sixteen items are found beyond this parameter. Further check on the Z-Std value, $-2<z<2$; shows *Reliable source* and *Fun* are beyond the upper limit, +2. Counter check against the Guttman scalogram (see Fig. 4) indicates that the two items, *Reliable source* (item 4) and *Fun* (item 26) have been under rated by respondent 41 and respondent 58 respectively.
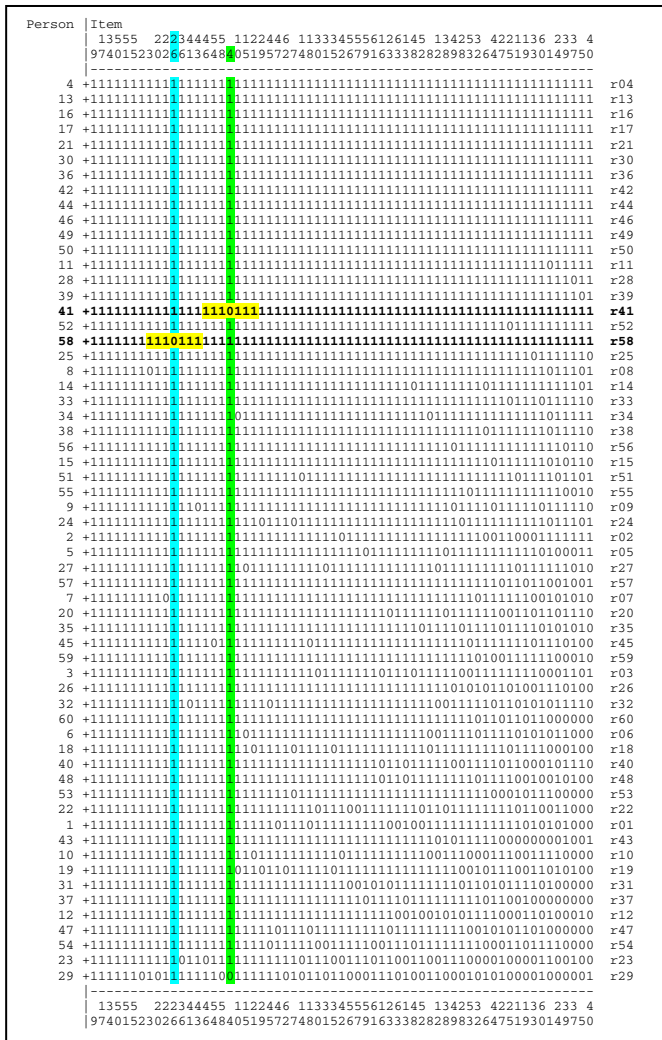
Fig. 4. Guttman scalogram

One possible reason is that they could have been careless in attempting their decisions which lead to such a grossly under rated work. After verifying that the Infit value (see INFIT column in Fig. 3) is within range (MNSQ: $0.5<y<1.5$; Z-Std: $-2<z<2$), the two misfits are acceptable.

As stated in the objective, there are two different aspects to this analysis, firstly the definition of family and secondly the assigning of criteria to the respective families. In order to analyze experts' views and comments from the open question provided in the content validity test, the percentage of the probability of the two aforementioned aspects to be agreed is calculated based on the *logit* measure. This is to decide whether to review them or not. A threshold value to 70% is set in line with the standard threshold limit of Cronbach Alpha. It can then be construed as follows:

- Definition of family and the assigning of criteria with percentage of probability to be agreed of more than 70% will be accepted without being reviewed.
- Definition of family and the assigning of criteria with percentage of probability to be agreed of less than 70% will be reviewed if comments are provided by the experts. The family will then be redefined whereas the criteria will be discarded or amended, if required.

The results for the 11 families are presented in Table 1. The summary shows that the definitions of 9 families are agreeable by the experts with the percentage of probability to be agreed is between 70% to 95%. On the other hand, the definition of *Family of Accuracy* and *Family of Completeness* need to be reviewed, with the percentage of probability to be agreed below 70%. However, the definition of the *Family of Completeness* is accepted without being reviewed because there is no comment available for it. On the other hand, *Family of Accuracy* has been redefined as suggested by the experts. See Table 3 for the accepted definitions of the 11 families.

Table 1. Percentage of possibility to be agreed for the definitions of 11 families

Family		$P(\Theta)$ (%)	Family		$P(\Theta)$ (%)
1	Accuracy	10.07	7	Blog Accessibility	84.16
2	Completeness	61.30	8	Blog Technical Features	84.16
3	Engaging	71.50	9	Currency	94.83
4	Reputation	71.50	10	Info Representation	94.83
5	Visual Design	71.50	11	Readability	94.83
6	Navigation	84.16			

The findings for the assigning of criteria to the respective families are shown in Table 2. It can be seen that 16 criteria (percentage of possibility to be agreed > 70%) remain in their respective families. Based on the findings, there are 36 criteria that require to be reviewed. However, there is no comment provided for the 31 criteria, means that they remain in their respective families. Yet, there are 5 criteria have been revisited; (1) *Relevant info* from *Family of Accuracy*, (2) *Easy to remember address* from *Family of Engaging*, (3) *Must-have sounds*, (4) *Info displayed in different format*, and (5) *Must have photos*. The later 3 criteria are from *Family of Info Representation*. Consequently, as suggested by the experts, the following actions have been taken for each of them:

- *Relevant info* is deleted from its family and has been transferred to the *Family of Completeness*.
- *Easy to remember address* is replaced by *Memorable content*.
- *Info displayed in different format* is eliminated from its family for sharing the same meaning as *Multimedia*.
- *Must-have photos* is discarded from its family for it is an integral part of *Multimedia*.
- *Must-have sounds* is removed from its family for it is also an integral part of *Multimedia*.

Table 2. Percentage of probability to be agreed for the assigning of 52 criteria

Family		$P(\Theta)$ (%)	Family		$P(\Theta)$ (%)
1	Must-have sound	2.90	27	Objective info	61.30
2	Relevant info	5.52	28	Real time info	61.30
3	Info in different format	8.24	29	Reputation of blog	61.30
4	Easy to remember address	9.11	30	Reputation of blogger	61.30
5	Must-have photos	9.11	31	Exciting content	61.30
6	Availability of blog owner info	16.52	32	Ease of ordering	61.30
7	Real-occurrence info	16.52	33	Blog responsiveness	61.30
8	Technorati rank	16.52	34	Ease of information access	61.30
9	Chat box	16.52	35	Blogroll	61.30
10	Emotional support	20.26	36	Comment field	61.30
11	Personal feel	22.62	37	Reliable source	71.50
12	Originality	25.35	38	Appropriate level of content	71.50
13	Interactivity	25.35	39	Provide information source	71.50
14	Rewarding experience	28.29	40	Clear layout of info	71.50
15	Surprises	31.86	41	Search tool	71.50
16	Readable font	31.86	42	Correct info	84.16
17	Amount of info	40.61	43	Reliable info	84.16
18	Easy to understand	40.61	44	Up-to-date	84.16
19	Multimedia	40.61	45	Appreciate comments	84.16
20	Intuitive interface	40.61	46	Fun	84.16
21	Informative	46.26	47	Fresh perspective	84.16
22	Good use of colours	46.26	48	Easy to navigate	84.16
23	Legibility	46.26	49	Attractive layout	84.16
24	Link to info	53.25	50	Appropriate explanatory text	94.83
25	Cognitive advancement	53.25	51	Easy to read info	94.83
26	Trackback	53.25	52	Availability of blog	94.83

See Table 3 for the final assigning of the 49 criteria to the 11 families concerned. They will then be used in the construction of questionnaire for measuring the acceptability of criteria for blog quality.

Table 3. Final result of content validity test

Family		Definition	Quality criteria	
1	Accuracy	The extent to which information is exact and correct, certified as being free-of-error.	1 2 3 4	Correct information Reliable info Reliable source Originality
2	Completeness/ Comprehensiveness of Info	The extent to which the information provided is sufficient.	5 6 7 8 9 10 11 12 13 14	Amount of information Appropriate explanatory text Appropriate level of content Availability of blog owner information Easy to understand information Informative Links to information Objective information Providing information sources Relevant info
3	Currency, Timeliness, Update	The extent to which the blog provides non-obsolete information.	15 16 17	Real time info Real-occurrence info Up-to-date info
4	Engaging	The extent to which the blog can attract and retain readers.	18 19 20 21 22 23 24	Appreciation for readers' comments Cognitive advancement Emotional support Fun Surprises Personal feel Memorable content
5	Reputation	The extent to which the information is trusted or highly regarded in terms of their source or content.	25 26 27 28	Reputation of blog Reputation of bloggers Rewarding experiences Technorati rank
6	Info Representation	The way information is presented, maybe in different formats/ media with customized displays.	29 30 31	Exciting content Fresh perspective Multimedia
7	Navigation	The extent to which readers can move around the blog and retrieve information easily.	32 33 34	Ease of ordering Easy to navigate Interactivity

Table 3. (*continued*)

Family		Definition	Quality criteria	
8	Visual Design	Visual appearances that can attract readers.	35	Attractive layout
			36	Clear layout of info
			37	Good use of colours
			38	Intuitive interface
9	Readability	Ability to comprehend the meaning of words or symbols.	39	Easy to read info
			40	Legibility
			41	Readable font/ text
10	Blog Accessibility	The extent to which the blog can be accessed faster and easier.	42	Availability of info
			43	Blog responsiveness
			44	Ease of information access
11	Blog Technical Features	Features such as search tools, chat box, blogroll, and comment field.	45	Blogroll
			46	Chat box
			47	Comment field
			48	Search tool
			49	Trackback

5 Conclusion and Future Work

This paper has described the content validity test to confirm whether the prescribed family definitions and blog quality criteria assigned are agreeable to the reviewers by means of online survey.

In conclusion, this study confirms the content validity of the 49 criteria in the 11 families for blog quality through a reliable and valid content validity test. These criteria will then be used as a basis in the construction of the instrument to measure the acceptability of blog quality criteria based on users' perception. It is also found that, despite the small sample size, Rasch Measurement Model is an effective tool in assessing content validity accurately and fast.

The content validity test is crucial to ensure that the content of our questionnaire is significant for meaningful measurement. Means that the instrument's content is essential to identify the important criteria from the blog readers' viewpoints in determining a blog quality. Consequently, it is an initial step towards the achieving a valid blog quality model. The model can be used as guidelines for blog readers to verify whether the visited blog is of quality or not. Besides, the model can help bloggers to promote readers' satisfaction.

However, this study does not establish the construct validity and criterion validity of the measurement instrument. Also, there is no evidence as to whether it is reliable or not. Therefore, as for future work, we plan to continue identifying the three aforementioned aspects in an effort to develop an accurate measurement instrument for getting a precise and correct blog quality model.

References

1. Lenhart, A., Fox, S.: A portrait of the internet's new storytellers (2006)
2. Vanier, F.: World Broadband Statistics: Q4 2009 (2010)
3. Tan, J.-E., Ibrahim, Z.: Blogging and Democratization in Malaysia. A New Civil Society in the Making. SIRD, Petaling Jaya (2008)
4. Kitchenham, B.A., Pfleeger, S.L.: Personal Opinion Surveys. In: Shull, F., Singer, J., Sjøberg, D.I.K. (eds.) Guide to Advanced Empirical Software Engineering, pp. 71–92. Springer, London (2008)
5. Abdul Aziz, A., Mohamed, A., Arshad, N., Zakaria, S., Zaharim, A., Ahmad Ghulman, H., Masodi, M.S.: Application of Rasch Model in validating the construct of measurement instrument. International Journal of Education and Information Technologies 2(2) (2008)
6. Rasch, G.: Weblogs models for some intelligence and Student test. The University of Chicago Press, Chicago (1980)
7. Masodi, M.S., Abdul Aziz, A., Mohamed, A., Arshad, N., Zakaria, S., Ahamd Ghulman, H.: Development of Rasch-based Descriptive Scale in profiling Information Professionals' Competency. In: IEEE IT Simposium, Kuala Lumpur, pp. 329–333 (2008)
8. Sick, J.: Rasch Measurement in Language Education Part 3: The family of Rasch Models. Shiken. JALT Testing & Evaluation SIG Newsletter 13(1), 4–10 (2009)
9. Wright, B.D.: Rasch Model from Counting Right Answers: Raw Scores as Sufficient Statistics. Rasch Measurement Transactions 3(2), 62 (1989)
10. Wright, B.D., Mok, M.M.C.: An overview of the family of Rasch measurement models. In: Everett, J., Smith, V., Smith, R.M. (eds.) Introduction to Rasch Measurement: Theory, Models, and Applications, p. 979 (2004)
11. Abdul Aziz, A.: Rasch Model Fundamentals: Scale Construct and Measurement Structure. Perpustakaan Negara Malaysia, Kuala Lumpur (2010)
12. Bond, T.V., Fox, C.M.: Applying The Rasch Model: Fundamental Measurement in the Human Sciences, 2nd edn. Lawrence Erlbaum Associates, New Jersey (2007)