Towards an Understanding of User Satisfaction Measurement in Online Discussion Site (ODS) Design using Rasch Analysis

[1]Prasanna Ramakrisnan, [2]Azizah Jaafar, [3]Noor Faezah Binti Mohd Yatim

[1] *Faculty Computer and Mathematical Sciences (FSKM), University Technology MARA (UiTM), 40450 Shah Alam, Selangor MALAYSIA*

[2]*Institute of Visual Informatics, National University of Malaysia, 43600 Bangi, Selangor MALAYSIA*

[3]*Faculty Infomation Science and Technology (FTSM), University Kebangsaan Malaysia (UKM), 43600 Bangi, Selangor MALAYSIA,*

prasanna@fskm.uitm.edu.my[1],

aj@ftsm.ukm.my[2],

nfmy@ftsm.ukm.my[3]

Abstract

In this paper, we propose Rasch analysis for studying user satisfaction towards On-line Discussion Sites (ODS) design. The proposed analysis was used to investigate user satisfaction towards the design of ODS in Universiti Teknologi MARA (UiTM) e-learning portal. The user satisfaction measurement uses factors adopted from the e-Learner Satisfaction (ELS) and Website User Satisfaction (WUS) measures. A twenty four item questionnaire was designed based on the user satisfaction measurement and administered to 86 students. The data collected were later analyzed using the Rasch measurement model for fit statistics, unidimensionality and least satisfactory items. The finding indicated the likelihood of user satisfaction with the ODS design is about 0.91 logit. The analysis also revealed the students were least satisfied with items from the learner interface and it need to be further investigated for higher probability of user satisfaction.

Keywords: *User satisfaction; On-line Discussion Site (ODS); Rasch measurement model*

1. Introduction

Over the years On-line Discussion Sites (ODS) have become platform users to share information. With the rapid growth of information technology, discussions among users were facilitated and conducted on-line anytime from anywhere. At Universiti Teknologi MARA (UiTM), students taking similar courses are connected to each other through an on-line discussion sites using an e-learning portal known as the i-Learn Portal. They are able to interact with other students in their class and lecturers for information sharing regardless of their location at any time. As the largest university in Malaysia, UiTM enrolls the highest number of students every semester. With this increase in the number of students, generic courses with discussion based sessions are now being conducted on-line for reducing classroom utilization hours. Therefore, the ODS needs to be effective in providing a positive influence on user behavior participating in on-line discussions.

User satisfaction is used as a valid measure of system effectiveness (Gatian, 1994). Hence, user satisfaction towards the ODS has become an important factor for continued student participation on-line discussions. The probability of using the ODS for discussing on-line will decrease if users are not satisfied with the design. So the success of providing a platform for on-line discussions should be guided by the analysis of user satisfaction towards the design of the ODS. There are many studies related to ODS. Many of previous studies were conducted to evaluate the ODS for promoting collaborative learning (Dixon, Dixon, & Axmann, 2008); improving students' classroom learning (Luppicini, 2007); allowing students participation in knowledge construction (Lucas & Moreira, 2010; Roschelle, Pea, Hoadley, Gordin, & Means, 2000); fostering critical thinking (Borsoto 2004; Thomas 2002); increasing student interaction (Carswell, Thomas, Petre, Price, & Richards, 2002); promoting collaboration (Murphy, 2004; Sethi & Gil, 2011; Zilouchian Moghaddam, Bailey, & Poon, 2011); investigating message quality(Kay, 2006; Weinberger, Stegmann, Fischer, & Mandl, 2007) and studying user experiences (Raitman, Augar, & Zhou, 2005; Sullivan et al., 2011).

It was also identified that user satisfaction evaluation was commonly conducted on websites and information systems. There is a lack in user satisfaction evaluation towards ODS design because most of the evaluation on user satisfaction is associated with the design of e-learning systems. As the ODS is part of the e-learning system, the design of the ODS is not widely evaluated. However, a study has shown that there is a need to further investigate ODS designs (Harman & Koohang, 2005). Therefore, a user satisfaction model is proposed to measure the design of ODS. Previous related studies were examined to identify a suitable user satisfaction model.

1.1 Purpose of Study

The purpose of this study is to examine the validity and item hierarchy of user satisfaction model by applying the Rasch measurement model. The study aims include examining item fit and establishing the item hierarchy for ordering items from greater to endorse (bottom) to lesser to endorse (top). The finding from this study provides evidence to support the validity of the instrument used to understand meaningful activities related with user satisfaction towards ODS designs. By establishing a clearer understanding of user satisfaction at this level developers would have access to information to structure effective ODS. User satisfaction outcome can be accessed with meaningful activities related with learners interface, learning community, information and personalization.

1.2 Significance

This study can provide important data which influence both the theory and practice of user satisfaction in ODS design. Besides that, it will be also possible to identify which construct that are more difficult to be endorsed by students in activities related to ODS design.

2. Related Work

User satisfaction towards educational based ODS can be defined as a summary affective responses of varying intensity that follows asynchronous e-learning activities, and is stimulated by several focal aspects, such as content, user interface, learning community, customization, and learning performance (Wang, 2003).

There are many possible ways to evaluate user satisfaction. In the past several questionnaires have been developed to assess user satisfaction. Among them is the well-known Questionnaire for User Interaction Satisfaction (QUIS). It is a measurement tool designed to assess user satisfaction with human-computer interface (Chin, Diehl, & Norman, 1988). This tool was developed at the Human-Computer Interaction Laboratory (HCIL), University of Maryland at College Park. The QUIS Version 7.0 contains a demographic questionnaire; six scales to measure overall system satisfaction; four measures of specific interface factors. The interface factors are screen factors, terminology and system feedback, learning factors and system capabilities; and lastly an optional section to evaluate specific components of the system. The specific components are the technical manuals and on-line help, on-line tutorials, multimedia, internet access, and software installation. This instrument uses a different rating scale for each of the factors and requires a license to use it.

Authors have presented a possible user satisfaction measure using an integrated model with six dimensions: learners, instructors, courses, technology, design, and environment (Sun, Tsai, Finger, Chen, & Yeh, 2008). This measurement is used to investigate the critical factors affecting learners' satisfaction in an e-learning environment. This study was conducted to evaluate overall user satisfaction towards the e-learning environment and not a specific system design. Design was only used as one of the factors to measure user satisfaction.

Another user satisfaction measurement is the Website User Satisfaction (WUS). The WUS proposes four dimensions; layout, information, connection and language customization (Muylle, Moenaert, & Despontin, 2004). The development of this measurement is based on the IS success theory, hypermedia design theory, a qualitative exploratory pilot study, and a quantitative on-line critical incident technique. This is a general model that can be used to evaluate user satisfaction in any web-based application.

Besides that, studies have proposed a measurement of e-learner satisfaction (ELS) with asynchronous e-learning systems (Shee & Wang, 2008; Wang, 2003). The ELS identified learner interface, learning community, content and personalization dimension as the measure. The instrument for this measure was developed using samples only from Taiwan. Thus a confirmatory analysis and cross cultural validation using samples collected from outside Taiwan is required for generalization of this instrument.

3. Measurement Development

An ODS is asynchronous communication based discussion conducted in a web based e-learning system. The ODS is a part of asynchronous e-learning system where participants send or post messages at different times. Thus, the integration of ELS and WUS model was deemed the most suitable measure for identifying user satisfaction in the ODS. The items in this study were adapted from existing measures in the literature. User satisfaction of the ODS was measured with items drawn from the WUS model (Muylle et al., 2004) and the ELS model (Wang, 2003).

The construct includes items about the learner interface, learning community, information and personalization. The measures for the learner interface were derived from the ELS model, including ease of use, user friendliness, operational stability and ease of finding. Layout, guidance, structure as well as hyperlink connotation are included from the WUS model. The measures for learning community were taken directly from the ELS model. Information was measured using the up to date, exact fit, sufficient as well as usefulness from the ELS model and item reliability, ease of understanding plus clear presentation from the WUS model. Lastly, personalization used capability of learning content needed, choosing what needed to be learned, controlling learning progress, recording learning performance from the ELS model together with language customization from the WUS model.

Closed card sorting techniques were used to test the reliability and validity of the measurement items in this study since the measures came from different models. The participants were provided with four predetermined categories; learners interface, learning community, information and personalization. They were then assigned the questions in the index cards to the given four categories. This helped in identifying the degree to which the participants agreed to the items belonging to the given category. The participants consisted of one IT professional, an academic scholar and a research student. A 93 percent correct hit ratio was achieved in this round, which indicates a sufficient item-construct reliability (Moore & Benbasat, 1991) and as such a second-round of card sorting was not conducted.

3.1 Applying Rasch Analysis to Measure User Satisfaction

The Rasch measurement model was used for analyzing the data set. This measurement model is applied to measure latent traits (e.g., ability or attitude) in various disciplines. Latent traits are usually assessed trough the responses of a sample of users to a set measurement scale. Location of items and users of the measurement scale is estimated by the model from the proportion of responses of each user to each item. The probability of success depends on the differences between the ability of the person and the difficulty of the item. According to the Rasch measurement model, a user who is more developed has a greater likelihood of endorsing all the items; and easier tasks are more likely to be endorsed by all users (Bond & Fox, 2007).

The item difficulty and person ability are expressed in logits through transformation of the raw score (ordinal scale) percentage into success-to-failure ratio or odds. This odds value is then converted to its natural logs (interval scale). The scale resulting from the Rasch analysis of the ordinal response has the properties of an interval scale. This scale is linear, and the numbers tell how much more of the attribute of interest is present. The basic assumption of the Rasch measurement model are each user is categorized by his or her ability; each item by a difficulty; user and item can be presented by numbers along one line and lastly the probability of observing any particular scored responses can be computed from the differences between the numbers (Bond & Fox, 2007). Thus the model can be used to link the person to the items that have relative ordering of latent variables.

The validity testing of the user satisfaction instrument was done by applying Rasch measurement model for data analysis. The item hierarchy obtained from the person-item distribution map provides an indication of construct validity (Smith Jr, 2001). The use of the Rasch measurement model is for validating measures of constructs such as learners interface, learning community, information and personalization in user satisfaction model.

4. Method

4.1 Sample

The data was collected from full time undergraduate students at Universiti Teknologi MARA (UiTM) who use the i-learn portal, the e-learning portal for their academic discussions. 86 responses were collected from the survey distributed manually in four different classrooms and all the 86 responses were used as data set for analysis. Participants must be a user of the online discussion environment in i-learn portal. Participants varied in terms of gender, year and area of study. The sample details are presented in Table 1. A large proportion of the samples were female (74.4%). Most participants were either in their second year (50%) or third year (34.9%) of study. While the overall samples were selected from computing (55.8%) and non-computing (44.8%) areas.

Table 1. Sample Details

	Gender		Year of Study				Area of Study	
	Male	Female	≤ 1	2	3	≥ 4	Computing	Non-Computing
Total (n=86)	22	64	7	43	30	6	48	38
%	25.6	74.4	8.1	50.0	34.9	7.0	55.8	44.2

4.2 Instrument

An instrument was developed using measures adopted from the ELS and WUS model. The instrument is a 24-item that reflects the construct of meaningful activity of user satisfaction towards the ODS design. Four-point scales were used to obtain the responses from the participants. The scale used for this instrument was (1) strongly disagree, (2) disagree, (3) agree, (4) strongly agree. Table 2 summarizes the layout of the instrument.

4.3 Data Analysis

All analysis was conducted using the SPSS and WINSTEPS software. The SPSS software version 19 was used to analyze the participant's personal data (demographic data) while WINSTEPS software version 3.68.2 was used to analyze the data obtained from item-level responses. The WINSTEPS software is used to construct linear measures from ordered qualitative observations. Data collected were participant's item-level responses, gender (male/female), year of study (1/ 2/ 3/ 4) and area of study (computing/ non-computing). WINSTEPS was designed to construct Rasch measurement from the responses of a set of persons to a set of items. The collected responses were based on a rating scale of four categories. The categories are (1) strongly disagree, (2) disagree, (3) agree and (4) strongly agree. Output from the data analysis includes tables, graphs and plots which are included in analysis and finding writing.

Table 2. The Variables and Measures of the Instrument

Variables	Number of Item	Measures	Model Reference
Personnel Data	3	Gender	
		Year of Study	
		Area of Study	
Learner Interface	8	Ease of Use	ELS
		User Friendliness	ELS
		Operational Stability	ELS
		Ease of Finding	ELS
		Guidance	WUS
		Layout	WUS
		Structured	WUS
		Hyperlink Connotation	WUS
Learning Community	4	Discuss with Student	ELS

		Discuss with Lecturer Access Content Share Learning	ELS ELS ELS
Personalization	5	Learn Required Content Choice of Learning Control Learning Progress Records Learning Performance Choice of Language	ELS ELS ELS ELS WUS
Information	7	Up-to-date Information Exact Required Information Sufficient Information Useful Information Rely on Information Easy to Understand Information Clearly Presented	ELS ELS ELS ELS WUS WUS WUS

5.0 Analysis and Finding

5.1 Summary Statistics

The analysis was conducted to identify how well each item fit within the underlying construct in the user satisfaction model. The item difficulty (satisfactory) and person abilities to endorse were quantified in logits. The results yielded a Chi-Square value of 2387.97 with a 1950 degree of freedom. The reliability of the overall test can be seen in the Cronbach-α value of 0.92 based on Rasch analysis.

The summary statistic for the 24 measured items is shown in Table 3. It shows that all the item estimates from the mean item 0 logit. This is because in the Rasch analysis, the mean of an item is set to 0 logit as a starting point for calibration. The item reliability was 0.77 and considered fair on a scale of 0 to 1 (Fisher, 2007).

Table 3. Summary Statistics of Measured 24 Items

	Reliability	Raw Score	Count	Measure	Model Error	Infit Statistics		Outfit Statistics	
						MNSQ	ZSTD	MNSQ	ZSTD
Mean		256.4	85.9	0.00	0.25	1.00	0.0	0.95	-0.20
SD		8.1	0.3	0.53	0.01	0.17	1.0	0.19	1.0
Max		277.0	86.0	1.01	0.26	1.36	2.1	1.35	1.7
Min		240.0	85.0	-1.32	0.24	0.75	-1.7	0.64	-2.0
Real	0.77								

The summary statistic for 86 measured persons is illustrated in Table 4. The mean for a person was displayed at 2.01 logits and the maximum item difficulty to endorse is 1.01 logits (SE=0.26). The distance of the maximum item difficulty to endorse from the standpoint of the average group is -1.0 logits. Thus the average group of students comparatively endorsed (agree) the items in the instruments. The overall test was satisfactory. Therefore, the goodness of this instrument (user satisfaction model) to measure user satisfaction towards the ODS design was further analyzed.

Table 4. Summary Statistics of Measured 86 Person

	Reliability	Raw Score	Count	Measure	Model Error	Infit Statistics		Outfit Statistics	
						MNSQ	ZSTD	MNSQ	ZSTD
Mean		71.5	24.0	2.01	0.48	0.95	-0.3	0.95	-0.30
SD		7.6	0.2	1.80	0.06	0.55	1.8	0.58	1.8
Max		91.0	24.0	6.25	0.58	2.85	5.4	3.01	5.2
Min		55.0	23.0	-1.32	0.39	0.04	-3.7	0.03	-3.7
Real	0.91								

5.2 Model Unidimensionality

Unidimensionality is concerned if the obtained data forms a single factor (Bond & Fox, 2007). It is used to identify if the single latent trait explains all the variance in the data. The unidimensionality of the model is tested using a point-measure correlation, fit statistics and the principle component analysis (PCA). The point-measure correlation in the Rasch analysis provides information on the relationship between the observations on an item and the corresponding person measure. The items with a negative point-measure correlation will provide initial evidence of multidimensionality (Linacre, 1995). The recommended value for point-measure correlation ranged from 0.32 to 0.80. The point-measure correlations for all the items in the analysis were within the stated range. Thus, there was no initial evidence of multidimensionality.

Next, the fit statistics test was conducted to further evaluate the unidimensionality of the model. There are two types of fit statistics; infit and outfit. Fit statistics are observed for items aimed at a good fit between data. Fit statistics is used to evaluate the correspondence of data with the model. The infit statistics (weighted) report patterns of responses to items targeted on the person while outfit statistics (un-weighted) give the response pattern to items with difficulty far from a person. The value of infit and outfit statistics are reported as MnSq (mean of the squared residuals) and standardized Z values (Zstd) to show the size of randomness in the measurement. Generally, the recommended values for infit/outfit MnSq range from 0.6 to 1.4 and Zstd values ranging from -2 to 2 (Bond & Fox 2007). Items with a MnSq of more than 1.4 or Zstd of more than 2 or both indicate too much variance in the response pattern. Those items with erratic response patterns are considered as misfits to the model. This item could have been wrongly scored or belong to a different construct. Whereas items with a MnSq of less than 0.6 or Zstd of less than -2 or both indicate less variance in the response pattern. Those items are identified as over fitting to the model and might be a redundant measurement for the test. Usually items with a MnSq of more than 1.4 or Zstd of more than 2 or both are investigated first because it causes more distortion to the measurement compared to a MnSq of less than 0.6 or a Zstd of less than -2 or both. Therefore, if any item with poor fit statistics value were identified, it should be considered for removal from the instrument.

The analysis indicated a good fit as most of the items did not show values outside the stated range (see Table 5). Thus all the items measured the same underlying unidimensional latent trait. This indicates that the instrument fits the model, and all the items belonged to a different construct and there was no redundant measurement.

The final unidimensionality test was done using the PCA of the standardized residual from the Rasch analysis. The PCA transforms correlated items into principle components. It is expected that after removal of Rasch measures from the data, the residuals for pairs of item or person should be normally distributed and uncorrelated (Chien & Bond, 2009). Explicitly, there should be no presence of principle components. The unidimensionality was examined using three criteria from the PCA. First the variance explained by the measure need to be identified to judge the strength of the measurement dimension. The following guidelines for variance explained by measure were used (Conrad, Conrad, Dennis, Riley, & Funk, 2009): ≥ 40% is considered a strong measurement dimension, ≥ 30% is considered a moderate measurement dimension, and ≥ 20% is considered a minimal measurement dimension. Next is the variance explained by the first construct in the residuals needs to be less than 10% and last the eigenvalue of the first construct has to be below 3.0 (Eakman, 2012).

Table 5. Item Order and Fit Statistics in the Rasch Analysis of the 24 items (n=86)

Item Number	Item Measures	Item Difficulty (logits)	Point Measure Correlation	Infit Statistics		Outfit Statistics	
				MnSq	Zstd	MnSq	Zstd
23	Operational Stability	1.01	0.49	0.95	-0.2	0.84	-0.8
12	Layout	0.95	0.62	0.88	-0.8	0.81	-1.0
5	Ease of Use	0.48	0.68	1.00	0.0	0.96	-0.1
19	Records Learning Performance	0.37	0.57	0.98	-0.1	0.83	-0.9
20	Sufficient Information	0.37	0.65	0.75	-1.7	0.70	-1.7
16	Hyperlink Connotation	0.18	0.51	0.84	-1.0	0.78	-1.1
21	Guidance	0.18	0.59	1.01	0.1	0.90	-0.4
4	Control Learning Progress	0.13	0.59	1.08	0.5	0.98	0.0
13	Learn Required Content	0.12	0.60	0.76	-1.6	0.64	-2.0
14	Exact Required Information	0.12	0.48	1.01	0.1	0.97	-0.1
17	Discuss with Student	0.12	0.67	1.35	2.0	1.35	1.7
24	Rely on Information	0.12	0.50	0.82	-1.1	0.75	-1.3
2	Structured	0.06	0.69	0.85	-0.9	0.78	-1.1
8	Access Content	0.06	0.57	1.36	2.1	1.35	1.7

10	User Friendliness	0.06	0.67	1.21	1.3	1.26	1.3
22	Up-to-date Information	0.06	0.58	0.98	-0.1	1.0	0.1
9	Easy to Understand	-0.06	0.80	1.06	0.4	0.93	-0.3
15	Choice of Learning	-0.06	0.48	0.99	0.0	1.01	0.1
6	Information Clearly Presented	-0.13	0.66	1.23	1.4	1.15	0.8
18	Ease of Finding	-0.43	0.62	0.84	-0.9	0.75	-1.3
7	Choice of Language	-0.45	0.38	1.17	1.0	1.17	0.9
11	Discuss with Lecturer	-0.71	0.56	0.85	-0.9	0.82	-0.8
3	Useful Information	-1.25	0.47	1.01	0.1	1.09	0.5
1	Share Learning	-1.32	0.52	0.95	-0.2	1.06	0.4

Note: MnSq = Mean Square, Zstd = Z-Standardized

The observation of the PCA value was done based on Table 6 obtained from the Rasch analysis. The Rasch measurement model explained 37.1% of the raw variance in the instrument and was nearly identical to the variance expected by the model (37.0%). It reveals a moderate measurement dimension (≥ 30%) and thus a low likelihood of additional components being present. While the PCA resulted in the first component eigenvalue from the Rasch analysis displayed value of 2.8 representing 7.3% of the residual variance. An eigenvalue of less than 3.0 is considered good and less than 1.5 is measured to be excellent (Linacre, 2006). From the analysis, the eigenvalue of unexplained variances in the first contrast were less than 3.0 indicating unidimensionality within each construct. The interpretation of fit statistics and PCA of the standardized residuals from the Rasch measurement model indicated that the user satisfaction instrument actually assessed a unidimensional construct.

Table 6. Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)

	Empirical (Eigenvalue units)	**Empirical (Percentage of total variance)**	**Empirical (Percentage of unexplained variance)**	**Modeled**
Total raw variance in observation	38.2	100%		100%
Raw variance explained by measures	14.2	37.1%		37.0%
Raw variance explained by persons	11.2	29.3%		29.2%
Raw variance explained by items	3.0	7.8%		7.8%
Raw unexplained variance (total)	24.0	62.9%	100.0%	63.0%
Unexplained variance in 1st contrast	2.8	7.3%	11.6%	
Unexplained variance in 2nd contrast	2.2	5.8%	9.2%	
Unexplained variance in 3rd contrast	2.0	5.1%	8.2%	
Unexplained variance in 4th contrast	1.8	4.7%	7.5%	
Unexplained variance in 5th contrast	1.7	4.4%	7.0%	

5.3 The Person-Item Distribution Map (PIDM)

The Person-Item Distribution Map (PIDM) is used to display graphically the person and item hierarchy in a vertical dotted line ruler. The vertical scale in the ruler is an interval level iterative scale (Callingham & Bond, 2006). The spaces between the items; between the persons; and; between the items and persons in the scale have essential meanings. The unit of logit is used to state the measures. The ability of students to endorse an item is referred to as the person measure and the difficulty of an item to be endorsed by students is referred to as the item measure.

If the any space is detected between the vertical dotted line ruler, it is a point where the items or persons located in that area has the same measure. The distribution of the person is represented on the left of the ruler and ordered from less able (at the bottom of the scale) to more able (at the top of the scale) to endorse. This mean of persons with a lower satisfaction level will be located at the bottom of the ruler and persons with a higher satisfaction level will be located at top of the ruler. While the item spread is on the right of the ruler and based on difficulty to endorse the items. The items are arranged from less difficult to endorse (at the bottom of the scale) to more difficult to endorse (at the top of the scale). Thus, students with higher satisfaction (at the top of the scale) have less difficulty in endorsing the items in the instrument while students with lower satisfaction (at the bottom of the scale) have more difficulty to endorse even the easiest items.

The characters ‘M’, ‘S’ and ‘T’ can be seen along both sides of the ruler. The representation of letters on the left of the ruler is a person’s obverted value while on the right of the ruler is item the related value. The character ‘M’ indicates the average mean for the item and persons. The character “S” denotes one standard deviation away from the mean while “T” is an indication of two standard deviations away from the mean.

The PIDM linking the item difficulties to the person endorsement ability of the sample across the four response scale (strongly agree, agree, disagree and strongly disagree) is presented in Figure 1. The mean for item is always zero because the Rasch measurement model sets the mean of the item as a starting point (0 logits) for the calibration. It is a location where the item on a ruler corresponds with the person location at which there is a 0.5 (50:50) probability of endorsing the question. The item locations are scaled first based on the endorsement probability. The probability of a person endorsing a question with a difficulty lower than that of person's location is greater than 0.5 while the difficulty greater than the person's location is less than 0.5.

The item difficulty to endorse was plotted in relation to the mean satisfaction level (0 logits) for the group of students towards the ODS design. The higher the item estimate the more difficult that item was for the group to endorse and less satisfied the group was on that item. It was found that from the standpoint of the average group of students (person mean = 2.01 logit), all the items were plotted “much more easy to endorse”. These groups of students were judged to be more satisfied on most of the items.

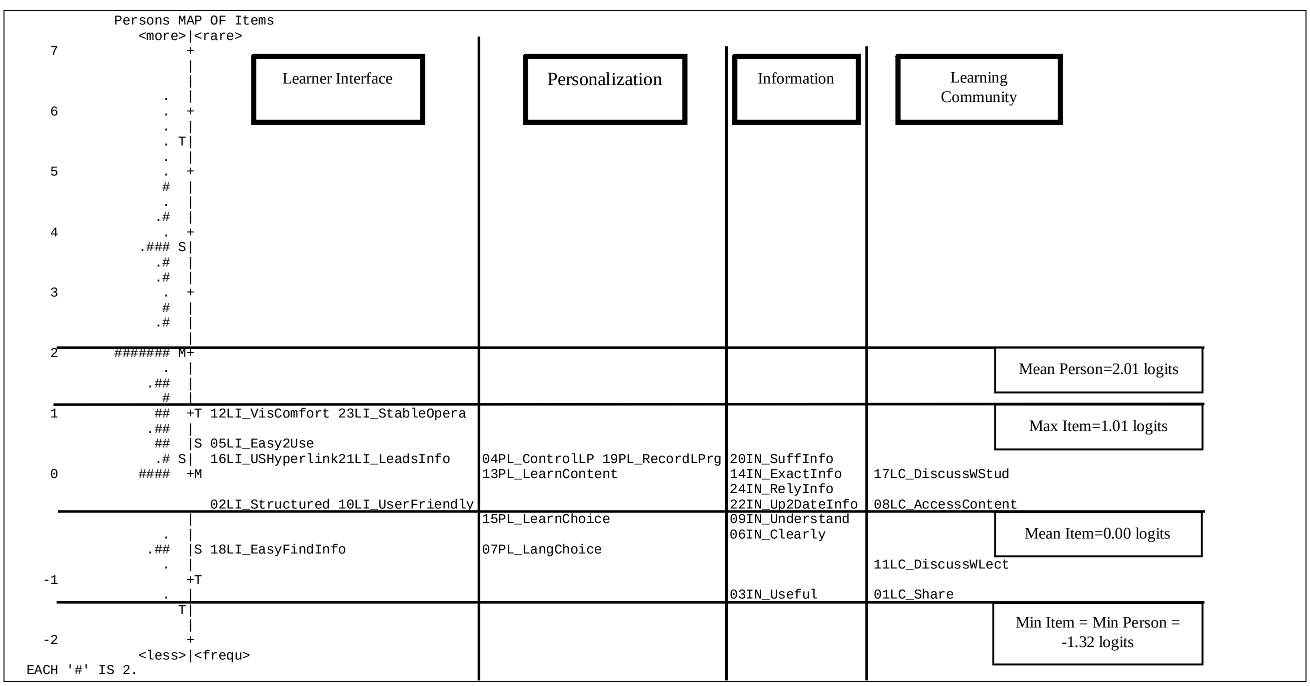

Figure 1. The Person-Item Distribution Map (PIDM)

The items distributions on the PIDM were in descending order of the item endorsement difficulty. Although most of the students were more satisfied on most of the items, there are a few items at the top of the ruler where the students had more difficulty to endorse. The most difficult items to endorse (see Table 5) were “Operational Stability” (1.01 logits), “Layout” (0.95 logits) and “Ease of Use” (0.48 logits). While the three least difficult items in ascending order of difficulty included “Share Learning” (-1.32 logits), “Useful Information (-1.25 logits) and “Discuss with Lecturer” (-0.71 logits). There were also item measures presented with an equivalent logit value (see Table 5).

6. Conclusion

User satisfaction evaluation is widely used for websites and e-learning tools, but its measurement properties have not been explored in an ODS design setting. The objectives of this study were to examine the validity and items that are less satisfactory to students in the ODS design. This study applied the Rasch measurement model to investigate fit statistics and unidimensionality in items of the user satisfaction evaluation. Evaluation of the item fits was done using general rule outfit MnSq more than 1.4 as cutoff for items in the instrument. Based on this rule, all the items displayed an acceptable fit to the Rasch Measurement Model. Findings from this study also indicated that all the underlying items were assessing a unidimensional construct. This is proven from the value obtained from the PCA of the Rasch model residuals. The analysis shows a considerable promise in determining user satisfaction towards the ODS design.

Applying the Rasch measurement model in the analysis provided potential information on satisfaction assessment on the construct of activities participated by the students in ODS. The items reflecting least satisfaction of activities are "Operational Stability", "Layout" and "Ease of Use". These aspects of activities meaning are commonly referred to learner interface. The findings offer evidence to suggest that students are least satisfied with some of the elements in the user interface of the current ODS. In conclusion, the findings from these analysis indicate validity and unidimensional measurement in the instrument used to identify user satisfaction towards the ODS design. The developer or designer can use these instruments as a basic assessment tool to aid them in the development and deployment of the ODS. Besides that researchers can use the measure student's satisfaction with the respect to activities in learner interface, personalization, information and learning community. The finding shows that least satisfaction towards ODS design is commonly from items belonging to the learner interface. Therefore, future studies need to be conducted to fully explore the elements in the ODS user interface to provide better student interaction experiences.

7. References

Bond, T., & Fox, C. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences*. Lawrence Erlbaum.

Borsoto, C. B. (2004). Community of inquiry a precondition of higher learning in online journalism courses. *Asia Pacific Media Educator*, *1*(15), 10.

Callingham, R., & Bond, T. (2006). Research in mathematics education and rasch measurement. *Mathematics Education Research Journal*, *18*(2), 1–10 LA – English. doi:10.1007/BF03217432

Carswell, L., Thomas, P., Petre, M., Price, B., & Richards, M. (2002). Distance education via the Internet: The student experience. *British Journal of Educational Technology*, *31*(1), 29–46.

Chien, C. W., & Bond, T. G. (2009). Measurement properties of fine motor scale of Peabody developmental motor scales-: A Rasch analysis. *American Journal of Physical Medicine & Rehabilitation*, *88*(5), 376. doi:10.1097/PHM.0b013e318198a7c9

Chin, J. P., Diehl, V. A., & Norman, K. L. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 213–218). New York, NY, USA: ACM. doi:10.1145/57167.57203

Conrad, K. J., Conrad, K. M., Dennis, M. L., Riley, B. B., & Funk, R. (2009). Validation of the Substance Problem Scale to the Rasch Measurement Model GAIN Methods. *Chestnut Health Systems.*, 3. Retrieved from http://www. chestnut. org/li/gain/psychometric_reports/Conrad et al

Dixon, R., Dixon, K., & Axmann, M. (2008). Online student centred discussion: Creating a collaborative learning environment. *Hello! Where are you in the landscape of educational technology? Proceedings ascilite Melbourne 2008*. Retrieved from http://pdf.aminer.org/000/287/659/enhancing_collaborative_learning_through_online_discussion_and_peer_assessment.pdf

Eakman, A. M. (2012). Measurement characteristics of the Engagement in Meaningful Activities Survey in an age-diverse sample. *The American Journal of Occupational Therapy*, *66*(2), e20–e29.

Fisher, W. P. J. (2007). Rating Scale Instrument Quality Criteria. *Rasch Measurement Transactions*, *21*(1), 1095.

Gatian, A. W. (1994). Is user satisfaction a valid measure of system effectiveness? *Information & Management*, *26*(3), 119–131. Retrieved from http://www.sciencedirect.com/science/article/pii/0378720694900361

Harman, K., & Koohang, A. (2005). Discussion Board: A Learning Object. *Interdisciplinary Journal of Knowledge and Learning Objects*, *1*, 67–77. Retrieved from http://ijello.org/Volume1/v1p067-077Harman.pdf

Kay, R. H. (2006). Developing a comprehensive metric for assessing discussion board effectiveness. *British Journal of Educational Technology*, *37*(5), 761–783.

Linacre, J. M. (1995). Prioritizing misfit indicators: an Insight based on Log-Linear Rasch Modeling. *Rasch Measurement Transactions*, *9*(2), 422. Retrieved from http://www.rasch.org/rmt/rmt92b.htm

Linacre, J. M. (2006). A user's guide to FACETS Rasch-model computer programs. *2009-06-30]. http://www. winsteps, corn/a/facets, pdf.*

Lucas, M., & Moreira, A. (2010). Knowledge Construction with Social Web Tools. In M. Lytras, P. Ordonez De Pablos, D. Avison, J. Sipior, Q. Jin, W. Leal, L. Uden, et al. (Eds.), *Technology Enhanced Learning. Quality of Teaching and Educational Reform SE - 40* (Vol. 73, pp. 278–284). Springer Berlin Heidelberg. doi:10.1007/978-3-642-13166-0_40

Luppicini, R. (2007). Review of computer mediated communication research for education. *Instructional Science*, *35*(2), 141–185 LA – English. doi:10.1007/s11251-006-9001-6

Moore, G. c., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research*, *2*(3), 192–222. doi:10.1287/isre.2.3.192

Murphy, E. (2004). Recognising and promoting collaboration in an online asynchronous discussion. *British Journal of Educational Technology*, *35*(4), 421–431.

Muylle, S., Moenaert, R., & Despontin, M. (2004). The conceptualization and empirical validation of web site user satisfaction. *Information & Management*, *41*(5), 543–560. Retrieved from http://www.sciencedirect.com/science/article/pii/S0378720603000892

Raitman, R., Augar, N., & Zhou, W. (2005). Employing wikis for online collaboration in the e-learning environment: Case study. *Information Technology and Applications, 2005. ICITA 2005. Third International Conference on* (Vol. 2, pp. 142–146). IEEE.

Roschelle, J. M., Pea, R. D., Hoadley, C. M., Gordin, D. N., & Means, B. M. (2000). Changing how and what children learn in school with computer-based technologies. *The future of children*, 76–101.

Sethi, R. J., & Gil, Y. (2011). A Social Collaboration Argumentation System for Generating Multi-Faceted Answers in Question & Answer Communities. *Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence*.

Shee, D. Y., & Wang, Y.-S. (2008). Multi-criteria evaluation of the web-based e-learning system: A methodology based on learner satisfaction and its applications. *Computers & Education*, *50*(3), 894–905. Retrieved from http://www.sciencedirect.com/science/article/pii/S0360131506001424

Smith Jr, E. V. (2001). Evidence for the reliability of measures and validity of measure interpretation: a Rasch measurement perspective. *Journal of Applied Measurement*.

Sullivan, F., Hamilton, C., Allessio, D., Boit, R., Deschamps, A., Sindelar, T., Vargas Ramos, G., et al. (2011). Representational guidance and student engagement: examining designs for collaboration in online synchronous environments. *Educational Technology Research and Development*, *59*(5), 619–644 LA – English. doi:10.1007/s11423-010-9178-x

Sun, P.-C., Tsai, R. J., Finger, G., Chen, Y.-Y., & Yeh, D. (2008). What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction. *Computers & Education*, *50*(4), 1183–1202. Retrieved from http://www.sciencedirect.com/science/article/pii/S0360131506001874

Thomas, M. J. W. (2002). Learning within incoherent structures: the space of online discussion forums. *Journal of Computer Assisted Learning*, *18*(3), 351–366.

Wang, Y.-S. (2003). Assessment of learner satisfaction with asynchronous electronic learning systems. *Information & Management*, *41*(1), 75–86. Retrieved from http://www.sciencedirect.com/science/article/pii/S0378720603000284

Weinberger, A., Stegmann, K., Fischer, F., & Mandl, H. (2007). Scripting argumentative knowledge construction in computer-supported learning environments. *Scripting computer-supported collaborative learning*, 191–211.

Zilouchian Moghaddam, R., Bailey, B., & Poon, C. (2011). IdeaTracker: An Interactive Visualization Supporting Collaboration and Consensus Building in Online Interface Design Discussions. In P. Campos, N. Graham, J. Jorge, N. Nunes, P. Palanque, & M. Winckler (Eds.), *Human-Computer Interaction – INTERACT 2011 SE - 23* (Vol. 6946, pp. 259–276). Springer Berlin Heidelberg. doi:10.1007/978-3-642-23774-4_23