

Rasch measurement theory in validation instruments for electronic financial technology in malaysia

Taasim, Shairil and Yusoff, Remali

Universiti Malaysia Sabah

1 January 2015

Online at https://mpra.ub.uni-muenchen.de/80736/ MPRA Paper No. 80736, posted 13 Aug 2017 09:39 UTC

Rasch Measurement Theory in Validation Instruments For Electronic Financial Technology in Malaysia.

Shairil Izwan Bin Taasim

Faculty of Business, Economics and Accountancy Universiti Malaysia Sabah Email: cheril.com@gmail.com

Remali Yusoff

Faculty of Business, Economics and Accountancy Universiti Malaysia Sabah Email: remyuf@ums.edu.my

ABSTRACT

The study develops a new instrument in measuring the validity of the questionnaire in technology banking applications using the Rasch model as an alternative method. Usually, classical method, the Cronbach alpha (a), is used to prove the validity of the instrument. In addition, the Rasch measurement model is also capable of providing guidance to proof quality items to strengthen the legitimacy of the survey instrument. Questionnaire consisting of 28 items and using a 5-level Likert scale with very unimportant to very important as the form of semantic differential was distributed to 223 respondents. Bond and Fox software analysis showed different response patterns to construct items that were measured in the same logit. Findings show the more widespread application of Rasch models would lead to a stronger justification of measurement particularly in cross-cultural studies and whenever measures of individual respondents are of interest.

Keywords: Rasch, Financial Technology, Measurement, Validity.

1.0 INTRODUCTION

K-Economy is a method for growth, and it revolutionizes the delivery of banking services with products such as internet banking and debit cards. To date, numerous studies have been conducted to investigate the factors influencing the acceptance of banking technology by using different models and theories. The existence of e-banking (electronic banking) is expected to create new markets in the banking world and provide significant benefits to both parties, application providers and users, by reducing the use of cash. Technology in banking has become a platform for banks to introduce their products that provide more efficient services. Studies by (Sharma, 2011) show that e-banking is used as a strategic tool by the world banking sector to

attract and retain customers. (Ivo and Saskia, 2011); (White, 2003) stated that the existence of information technology has enabled financial institutions to create, process and disseminate information quickly and cheaply. Study by (Murillo, Gerard and Roberto, 2010) on the adoption of internet banking among U.S. banks found that the role of internet banking is part of the bank's strategy and alternative abatement of opening new branches. Therefore, there is a need to evaluate the performance of the development of e-banking among Malaysian consumers. Feedback from respondent by questionnaires is often used to identify performance and consumer acceptance of electronic banking. Therefore, validation and strengthening of the questionnaire should be good and solid to support the objectives of the study. Confirmatory factor analysis and the Explanatory factor analysis are the methods often used by researchers.

2.0 LITERATURE REVIEW

To analyze test items, there are two types of commonly used statistical items by (Zhu, 1998), a classic item statistics (CTT) which takes into account item's difficulty and discrimination index which refers to the aggregate statistical variance, covariance and means (Thomas and Rudolf, 2005). This method has the disadvantage in which its value depends on the population analysis where the results will change when applied to different research groups due to the knowledge and skill levels of different samples. The second type of item statistics is derived from Item Response Theory (IRT), which contains the statistical difficulty of the item, calibration error and a correspondence item subset of the statistics that is able to estimate the extent to which an item complies with the expectations model of knowledgeable respondents to have a higher probability to give the correct answer. Item response theory methods are applied in the Rasch model to correct deficiencies in the Likert scale because the results are raw ordinal data, and it still needs to be processed because it does not have a regular interval. To improve the analysis, the method of Rasch measurement model (RMM) is used in this study as the primary objective is made by the best measurement. Rasch's measurement model established by (Rasch, 1980) is a measurement model that was formed as a result of the considerations on the ability of the respondents who answered the questionnaires, tests or instruments and the difficulty of each item (Rasch 1980). Previous studies by (Zamalia et al.,2013) and (Rasch, 1980) indicated the Rasch theory was able to test the item's difficulty and the ability of respondents at the same scale.

Normally, the Cronbach alpha (α) is only used to prove the validity of the instrument but Rasch measurement model is also capable of providing guidance to prove the quality of items to further strengthen the validity of the survey instrument (Azrilah et al., 2013). Rasch Measurement Theory by Georg Rasch comprises a model of item response (IRT), which later made famous by Ben Wright. Ordinal data does not have the same interval, so the data must be converted to the form of requirement ratio for statistical analysis. Rasch model was developed to determine the relationship between person's ability and item's difficulty where findings enable a high level of ability to be able to answer questions with a lower difficulty level (Bond and Fox, 2007).

(Choppin, 1983) provided explanations for the Rasch model in a mathematical equation. (Choppin,1983) essentially described the probability in Rasch model is the result when the respondent can answer an item to disable that feature singles and Rasch item. Based on the assumption that certain individuals respond properly to a particular item and the item does not depend on the answer to the previous item.

Probability
$$[X_{vi} = 1] = \frac{A_V}{A_V + D_I}$$
 (1)

Where, X_{vi} Value 1 if individual V responds to item i, and 0 otherwise

 A_V Parameters reflecting an individual's ability v.

 D_I Parameters describing the item difficulty i.

In this formula, A and D may vary from 0 to ___. Changes to these parameters are often introduced to simple mathematical analysis. New parameters are defined for individual ability () and item difficulty () to satisfy the equation:

 $A_v = W$ and $D_i = W$ for W constant. Rasch introduced and used in previous studies, for constant W are fixed proportion to natural Logarithmic Base, e. Therefore, the model can be written:

Probability
$$[X_{vi} = 1] = \frac{e^t}{1 + e^t}$$
 where $t = (\alpha_v - \delta_i)$ (2)

In this formula, α and δ can be taken into consideration for the ability and the difficulty of measuring respectively in the same logic scale. If $\alpha > \delta$, the result of the probability obtained is the correct response, and if $\alpha < \delta$, the actual results are incorrect response. Rasch also defined the ratio of the probability to obtain a probability than the one on display in the following simple equation;

Odds
$$[X_{vi} = 1] = \frac{e^t}{1 - \frac{e^t}{1 + e^t}} = e^t$$
 or $t = log_e(odds)$ (3)

For these conditions, the Rasch model is sometimes called and referred to as 'log-odds' model (Choppin, 1983). Considered good ordinal score categorize the data X_{nijk} for linear parameter equation (4) and (5) controlled by differential residual mean square between data X_{nijk} in forecast model E_{nijk} for fit equation respondent to pattern measurement respond B_n .

Infit
$$\sum_{ijk} (X_{nijk} - E_{nijk})^2 / \sum_{ijk} V_{nijk}$$
 (4)

Outfit
$$\left[\sum_{ijk} (X_{nijk} - E_{nijk})^2 / \sum_{ijk} V_{nijk}\right] / \sum_{ijk} 1$$
 (5)

Good subset of the statistics for each parameter measured expectations. Infit aim is to focus on the evolution of the reaction suitability as a conventional item biserial correlation and IRT item discrimination. Advantage of infit is the variance ratio formula. The outfit is the variance ratio sensitive to outliers possible off-target and detect anomalies such as guessing the tough questions, the negligence of a simple question. (Velo and Rosna, 2009) identify that the types of responses on the Likert scale using Rasch model is good for studying the validity and reliability of the instrument to maintain the accuracy of the questionnaire from exposure to disability. It means the more accurate the data, the higher the value for the validity and reliability of the questionnaire. (Rosenni et al., 2009) refers to reliability coefficient Cronbach alpha to measure the reliability of the items in a questionnaire. It is referred to the model that is commonly used on True Score Theory Test (TSTT) otherwise known as the classical model. Rasch model uses a mathematical formula that is roughly similar

as the measurement of the parameters in the Item Response Theory (IRT) or also known as Latent Trait Theory.

Table 1: Criteria for the validity of the Questionnaire Items

Criteria	Statistic	Result			
Validity Item	Polarity item	PTMEA CORR>0.3			
	Item Fit	Mean square infit and outfit 0.6-1.4			
	PCA	Varians 29.6%			
	Respondent reliability	0.83			
	Item reliability	0.96			
Distribution of	The estimated distance	4 logit (-1.0 hingga+3.0)			
respondents	of understanding				
The validity of	Percentage of	Infit			
the response of	respondents mean	10.2% < 0.4			
respondents	square between 0.4 -	18.3% > 1.6			
	1.6	Outfit			
		11.5% < 0.4			
		15.6% > 1.6			

Source: Bond and Fox (2007)

Table 1 shows Rasch's measurement, the validity of an instrument by reference to analysis such as polarity items, the item-person map, mismatch-individual items, item-individual isolation, unidimensional, compatibility and individual-item rating scale of (Rasch, 1980; Bond & Fox, 2007). Therefore, this study was undertaken to produce empirical evidence to strengthen the validity and reliability of the questionnaire for e-banking performance by using the Rasch's measurement model to test the questionnaire. According to (Thomas and Rudolf, 2005) theoretical distinction between CFA and Rasch is that a CFA assumes metric scale even though we know it is doubtful while Rasch relies on the number of respondents and do not have normal or form set.

3.0 METHODOLOGY

The study obtained data from a sampling of 470 respondents from Malaysia. Testing instruments used the Rasch model via software Bond & Fox. The instrument included 26 questions based on seven (7) constructs to examine the performance of e-banking in Malaysia. Data analysis has been conducted in several stages to prove the normal distribution of the data and it is also a requirement to meet the conditions of the test statistic. All items in the questionnaire were measured using a Likert scale from 1 (strongly disagree) - 5 (strongly agree) based on (Davis, 1989); (Hung-Pin Shih, 2004); (Yong, 2013); (Pasharibu et al., 2012); (Thompson, 2005); (Chen et al., 2007); (Widjayan, 2011); (Hung-Pin Shih, 2004); (Koi and Sze, 2002).

The selection of the sample size of the study represents a population Krejcie and Morgan (1970) method was applied in this study based on the population of people aged 15-74 years the number of 18,931,200 people in 2013 (Malaysia, 2013).

$$SIZE = \frac{X^2NP(1-P)}{\delta^2(N-1) + X^2P(1-P)}$$

 X^2 = Chi-square value of 1 degree of freedom at the desired confidence level. (0.05)

N = Population size

P = The proportion of the population (assumed 0.50) maximum sample size

 δ = The level of accuracy is expressed as a proportion (0.05)

$$s = \{(0.05)^2(18,931,200)(0.50)[1-0.5]\} / 0.05(18,931,200-1) + (0.05)^20.5(1-0.5)$$

$$s = \frac{3.8416(9465600)(1 - 0.5)}{47328 + 0.9604}$$

$$s = \frac{18181524.48}{47328.96}$$

 $s = 384.15 \approx 400$ respondent

4.0 RESULT AND DISCUSSION

Table 2 presents summary of the statistics from the Rasch's model analysis of 470 respondents who answered for 26 items on the instrument. Table 2 presents a high person reliability index (0.95) and a high item reliability index (0.85). These are considered good index for both item and person.

Table 2: Summary Statistic Instrument: Respondent and Item

Persons		470 Input		INFIT		OUTFIT	
	Score	Count	Measure	MNSQ	ZSTD	MNSQ	ZSTD
Mean	91.7	26.0	1.03	1.01	-0.9	1.00	-0.9
S.D	15.5	0.0	1.52	0.90	3.4	0.90	3.4
Person Reliability: 0.94							
Items		26 Input	INFIT		OUTFIT		
Mean	1579.3	448.0	0.00	0.99	-0.2	1.00	0.0
S.D	37.5	0.2	0.22	0.15	2.1	0.16	2.2
Item Reliability: 0.87							

Person Raw Score-to-Measure Correlation = 0.98

Cronbach Alpha (KR-20) Person Raw Score Reliability = 0.98

The mean infit and outfit are 1.01 for person and 0.99 for items mean squares. This indicates that the item fulfills requirement of (Bond and Fox, 2007) where 0.4 - 1.6 is accepted. The table also shows that the z-scores for infit and outfit are -0.9 (person) and -0.2 (items). This indicates that the data fit the model somewhat better than would be expected which could be due to some redundant items. The data also shows an overall acceptable fit as the value for standard deviation for person (1.52) and item (0.22).

According to the Rasch's measurement model, the validity of a questionnaire can be identified by analyzing the program output. The main output is a polarity item and should be referred to a correlation coefficient-point measurement known as the point measure correlation coefficient (PTMEA CORR). In addition, values are also referred to individual items such as maps, the mismatch-individual items, item-individual isolation, unidimensional, compatibility and individual-item rating scale of (Linacre, 2003). If the PTMEA CORR is high, an item will be able to distinguish between respondents capability. According to (Linacre, 2003), negative of zero value indicates joint response to the item or the respondent is contrary to the variables or constructs. The item sags if the value is less than 0.30 PTMEA CORR (Nunnally and Bernstein,

1994). Based on analysis, after removal of PC4 items because misfit is greater than the MNSQ Outfit > 1.6, as recommended by Bond and Fox (2007), showed high validity and reliability for item in the questionnaire. For the PTMEA CORR is more than 0.30, i.e., from 0.65 to 0.83. It can be concluded that the items contributing to the performance assessment for e-banking questionnaire could discriminate or differentiate between the use of e-banking applications respondent.

GUTTMAN SCALOGRAM OF RESPONSES:

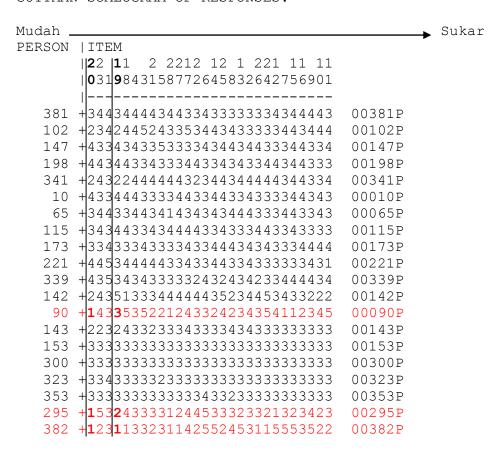


Table 3: Guttman Response Pattern Scalogram

To reinforce that there are two items PC3 (entry 20 no) and PC4 (19) should be eliminated in Table 3 with the MNSQ not meet the minimum criteria proposed strengthening the Rasch model is necessary for removal based on the analysis of data obtained support through Scalogram Guttman. Rasch measurement model states that items such as PC3 and PC4 show patterns of response that does not meet its tough item (Bond and Fox, 2007).

Infit MNSQ large items also display the probability that this item of outstanding individuals who are negligent in reply to the questionnaire. This assumption is reinforced by the PMC low as in appendix 1. Rasch measurement model suggests referring to Guttman Scalogram respondents as a means of detecting such a condition occurs, such as in Table 3, which shows the number of respondents who answered the questionnaire with the older respondents are respondents competent and item to the right is a difficult item. Excellent response of 90, 295 and 382 on the item PC3 (20) and PC4 (19) is unreasonable and more likely as a negligence (Azrilah et al, 2013), because the answer is more difficult items to errors in answering questions questionnaire. This also proves that the Rasch Measurement Model and removal of the items have a reason and the reason that allows the item to be removed by taking into account the difficulty of the items and the ability of the respondents said item.

Person map of item is the last determination for validity of our data and item. Figure 4 shows the Rasch analysis to produce a mapping of the distribution of the items to the distribution of the ability or tendency of respondents. According to (Bond and Fox, 2007), the purpose of this mapping is to show the relationship between the ability of respondents and the level of difficulty of the items. Respondents with high abilities and items with the highest difficulty level are at the top of the scale, while respondents with low abilities and items at the lowest difficulty level are located at the bottom. This is because the measurement using the logit scale shown above is based on the simplest to the most difficult level. Since most of the respondents are in the vicinity of the mean logit of 0 to 0.5 of the ability level of the respondents, the mean logit values 0 was set for the item. Mapping depicts most individuals to have much higher ability levels to answer the most difficult item in the questionnaire. Figure 5, it is seen that the most difficult item (PB1) is at the top of the scale and the easiest item (PC3) is located at the very bottom of the scale. The estimated distance for respondents to understand e-banking is approximately 3 logit (from -1.0 to +2.0).

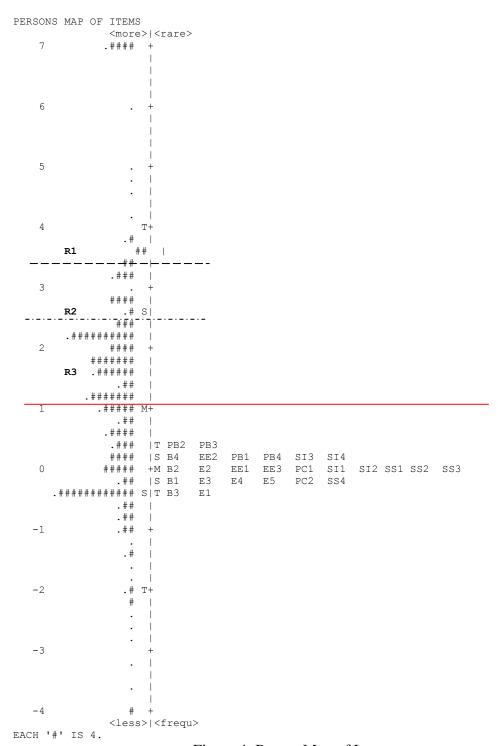


Figure 4: Person Map of Items

6.0 CONCLUSION

Validity and reliability of each item in the questionnaire is important to ensure accuracy and data entry are as meant and contributes to the validity and reliability of the results. If the reliability or validity of the questionnaire was high, then the questionnaire is reliable and valid. Although the questionnaire used by researchers has been previously tested for validity and reliability, the questionnaire should be tested again because the inference obtained is only suitable for the purpose and samples of the particular study, especially if it was analyzed using Classical Test Theory or True Score Theory Test (TSTT). In this study, by using the Rasch's measurement model, researchers have obtained the high reliability of test of the items and they also indicate that the questionnaire is valid and reliable to measure e-banking. In addition, the questionnaires were administered to the appointed time and enjoy the respondents, thus no mismatch problem items and respondents (50% fit) found during the process of data analysis. One of the advantages of modern psychometric methods is the ability to identify his formula items and respondents misfit. Respondent should be able to answer very clever questions easily. To obtain more accurate results and consistency, it is proposed for future research questionnaire, the same data to test the construct validity by using structural equation modeling method known as structural equation modeling (SEM).

REFERENCE

- Arsaythamby Veloo and Rosna Awang Hashim. (2009). Kesahan dan kebolehpercayaan alat ukur orientasi pembelajaran matematik (OPM). *International Journal of management Studies*. Vol 16(1): 57-73.
- Azrilah, A.,B., Saidfudin, M.,M. dan Azami, Z. 2013. *Asas Model Pengukuran Rasch. Pembentukan Skala dan Struktur Pengukuran*. Universiti Kebangsaan Malaysia.
- Bruce Choppin. (1983). *The Rasch Model For Item Analysis*. CSE Report No.219. https://www.cse.ucla.edu/products/reports/R219.pdf
- Bond, T.G and Fox, C.M. (2007). Applying the Rasch Model; Fundamental Measurement in the Human Science. New Jersey: Lawrence Erlbaum Associates.
- Chun, D, C., Yi W, F., and Cheng, K., F. (2007). Predicting Electronic Toll Collection Service Adoption: An Integration of the Technology Acceptance Model and

- the Theory of Planned Behavior. *Transportation Research Part C.* Vol 15:300-311.
- Fred, D. Davis. (1989). Perceived Usefulness, Perceived Ease of Use and User acceptance of Information Technology. *MIS Quarterly*. Vol 13 (3). pp 319-340.
- Sharma, H. (2011). Bankers Perspectives on E-Banking. NJRIM, Vol.1 (1): 71-85.
- Ivo J.M. Arnold dan Saskia E. van Ewijk. (2012). The quest for growth: The impact of bank strategy on interest margins. *International Review of Financial Analysis*, Vol 25: 18-27.
- Hung, P.,S. (2004). Extended Technology Acceptance Model of Internet Utilization Behavior. *Journal Information & Management*. Vol 41: 719-729.
- Pasharibu dan John J.O.I Ihalauw. (2012). E-banking: What is in Prospects' Mind?. *Proceedings of the IEEE ICMIT*.: 325-330.
- Thompson S.H Teo dan Yuan, Y., Y. (2005). Online buying behavior: a transaction cost economics perspective. *The International Journal of Management Science*. Vol 33: 451-465.
- Widjana M., A and Basuki, R. (2011). Factors Determining Acceptance Level of Internet Banking Implementation. *Journal of Economics, Business and Accountancy Venture* Vol.14 (2): 161-174.
- Loo, L., S., and Sze, M., K. (2002). Singapore's Internet shoppers and their impact on traditional shopping patterns. Journal of Retailing and Consumer Services Vol.9. 115–124.
- Linacre, J., M. (2011). Rasch Measures and Unidimensionality.Rasch Measurement Transactions. 24:4: 1310. http://www.rasch.org/rmt/rmt244f.htm
- Nunnally and Bernstein. (1994). Psychometric theory (3rd). New York: McGraw-Hill.
- Hun, P.,S. (2004). Extended Technology Acceptance Model of Internet Utilization Behavior. *Journal Information & Management*. Vol 41: 719-729.
- Roberto Fuentes and Rubén Hernández-Murillo & Gerard Llobet, (2007). "Strategic online-banking adoption," Working Papers 2006-058, Federal Reserve Bank of St. Louis.
- Murillo, R. H., Llobet, G., and Fuentes, R. (2010). Strategic Online Banking Adoption. *Journal of Banking & Finance*. Vol.34: 1650-1663.
- Robert V.Krejcie dan Daryle W.Morgan (1970) Determining Sample Size For Research Activities. Educational and Psychological Measurement. Vol 30. 607-610.

- Rosseni, D., Mazalah, A., Faisal, Norhaslinda, M, S., Aidah, A, K., Nur, A, J., Kamaruzaman, J., Mohamad, S, Z., Khairul, A, M., and Sit, R, A. (2009). Validity and Reliability of the e-Learning Style Questionnaire (eLSE) Version 8.1 Using the Rasch Measurement Model. *Journal of Quality Measurement and Analysis*. Vol 5(2): 15-27.
- Salzberger, T., and Sinkovics, R., R. (2005). Reconsidering the problem of data equivalence in international marketing research Contrasting approaches based on CFA and the Rasch model for measurement. *International Marketing Review*. Vol. 23(4): 390-417.
- Venkatesh, Morris, Davis G.B and Davis F.D. (2003). User Acceptance of Information Technology: Toward A Unified View. *MIS Quarterly*. 27(3): 425-478.
- Benjamin D.Wright. (1996). Comparing Rasch Measurement and Factor Analysis. Structural Equation Modelling. Vol 3 (1): 3-24.
- Thomas Salzberger and Rudolf R. Sinkovics. (2005). Reconsidering the problem of data equivalence in international marketing research Contrasting approaches based on CFA and the Rasch model for measurement. *International Marketing Review*. Vol. 23(4): 390-417.
- Zamalia, M., Nor, A., and Rosli, A, R. (2013). Assessing Students' Learning Ability In A Postgraduate Statistical Course: A Rasch Analysis. Procedia *Social and Behavioral Sciences* Vol 89: 90 894.
- Zhu, W-M. (1998). *Test equating: What, why, how?*. Research Quarterly for Exercise and Sport. Vol. 1: 11-23.

APPENDIX 1

ITEM STATISTICS: MEASURE ORDER

ENTRY NUMBER		COUNT		MODEL IN S.E. MNSQ	ZSTD MNSQ	ZSTD	CORR.	OBS%	EXP%	+ ITEM
11	1593	470	.49	•	3.0 1.26					
10	1599	470	.46		2.8 1.24					PB2
1 9	1631	470	.29		2 .99					PB1
15	1645	470	.21	.07 1.07	1.0 1.08			60.7	58.4	SI3
16	1645	470	.21		.5 1.05			63.3	58.4	SI4
7	1652	470	.17	.07 .92	-1.2 .93	-1.0	.72	68.1	58.5	EE2
12	1653	470	.17		.2 1.01	.1	.72	64.0	58.5	PB4
24	1653	469	.15	.07 .89	-1.6 .89	-1.6	.75	65.9	58.5	B4
26	1662	470	.12	.07 .89	-1.6 .89	-1.5	.74	65.7	58.8	SS2
2	1664	470	.11	.07 1.11	1.5 1.15	2.1	.69	60.0	58.8	E2
13	1672	470	.06	.07 .78	-3.4 .79	-3.1	.76	69.2	58.9	SI1
8	1676	470	.04	.07 .84	-2.4 .83	-2.6	.73	68.6	58.9	EE3
25	1680	470	.02	.07 .97	4 .99	1	.71	62.6	58.9	SS1
14	1681	470	.01	.07 .94	9 .93	-1.0	.74	65.7	58.9	SI2
6	1684	470	01	.07 .88	-1.7 .86	-2.0	.72	68.6	59.0	EE1
22	1685	470	01	.07 1.08	1.2 1.05	.8	.71	63.7	59.0	B2
17	1686	470	02	.07 .90	-1.5 .90	-1.4	.74	68.1	58.9	PC1
27	1687	470	02	.07 .83	-2.6 .83	-2.5	.73	66.6	58.9	SS3
5	1707	470	13	.08 .90	-1.4 .88	-1.7	.73	66.2	59.0	E5
28	1707	470	13	.08 .86	-2.0 .85	-2.2	.76	68.4	59.0	SS4
21	1711	470	16	.08 .71	-4.6 .70	-4.6	.76	76.3	59.0	B1
3	1712	470	16	.08 .88	-1.8 .87	-1.9	.73	66.6	59.0	E3
4	1713	470	17	.08 .76	-3.7 .75	-3.8	.76	69.2	59.0	E4
18	1725	470	24	.08 1.00	.0 .99	1	.72	66.4	59.1	PC2
19	1727	470	25	.08 1.41	5.4 2.21	9.9	.58	63.1	59.1	PC3
1	1740	470	32	.08 1.35	4.6 1.32	4.0	.68	59.6	59.2	E1
23	1747	470	36	.08 .97	4 .96	5	.73	67.7	59.3	B3
1 20	1772	470	51						59.4	
	1619 5	455 O	.00		2 1.02					
				.00 .18						ľ