Scientific
Research
Publishing

Psychometric Evaluation on Mathematics Beliefs Instrument Using Rasch Model

Siti Mistima Maat

Department of Teaching & Learning Innovations, Faculty of Education, National University of Malaysia, Bangi, Malaysia
Email: sitimistima@ukm.edu.my

Abstract

The purpose of this study was to provide the evidence of psychometric evaluation on mathematics beliefs instrument (MBI) using Rasch model. 36 items of 5 point Likert scale with three constructs of mathematics beliefs towards nature, mathematics belief in teaching and mathematics beliefs in learning. 254 of secondary mathematics school teachers were selected to be the samples. The collected data were analysed using Winstep 3.69.1.11 in order to obtain the Rasch Model output. The findings revealed that the MBI has reliability value of 0.98 with separation index 6.27. The person reliability has value of 0.81 and the separation index of 2.08. Some misfits items were modified accordingly based on the fit statistics suggestions. The Rasch model output has provided statistical evidence for MBI for future purposes.

Keywords

Mathematics, Beliefs, Rasch Model, Reliability

1. Introduction

One of important aspects to be studied in mathematics education is belief. Mathematics beliefs play a role in shaping teachers' practices and create their attitude towards teaching (Swan, 2002). Being successful teachers would derive beliefs system effectively through teaching process (Wilkins, 2008). Although there are increasing numbers of researches on beliefs, the focus is mainly on its relationship with other variables. The beliefs system has a dynamic structure which can be related to achievement, teaching practices and knowledge (Muijs & Reynolds, 2002).

However, mathematics beliefs are limited to empirical measurement only since the focus is more on its relationship with other variables (Swan, 2002) such as the effect on instructional practices (Beswick, 2005, 2007). As a result, by testing the chosen the instrument of mathematics beliefs would lead to a reliable validated meas-

urement of beliefs. The justification of testing the psychometric aspect of the mathematics beliefs is to provide empirical evidence in providing a credible measurement. A meaningful measurement of any variables produces standardized instrument that can be replicated to a different study (Saidfudin et al., 2010). The importance of measuring mathematics beliefs in terms of psychometric aspects is to get the knowledge of the validity, reliability, scores, format and scale of the instrument (Ali et al., 2014). Some selected recent studies like Head (2012), Ali et al. (2014), Çakiroğlu & Işıksal (2009) and Pampaka & Williams (2010) have tested mathematics beliefs instrument using Rasch model. The Rasch measurement model refers to a technique to measure any latent traits (Abdul Aziz et al., 2013). It provides information on validity of instrument which reflects to the quality of the items in the questionnaire. Despite the differences in terms of samples and mathematics beliefs constructs, the findings are found to be common in terms of instrument validity and reliability. The research gap of this study is based on samples as well as the mathematics beliefs instrument. Therefore, by providing the psychometric evaluation on testing and validation of the said instrument would fulfil the research objective.

2. Methodology

Using a quantitative approach, 254 of secondary mathematics school teachers were selected randomly to be the sample of this study. There were 51 and 203 male and female teachers respectively. A set of Mathematics Belief Instrument (MBI) which adapted from Evans (2003) used 5-point Likert scale ranging from "1" as "strongly disagree" to "5" as' strongly agree'. Initially, there were three constructs that include mathematics beliefs towards nature (11 items), beliefs about mathematics teaching (12 items) and beliefs about mathematics learning (13 items).

However, after exploratory factor analysis, some of the items have to be deleted due to low loading factor value. It was revealed that the reliability value of 0.81indicates that the internal consistency for MBI is acceptable (Hair et al., 2010). In addition, each construct of MBI has reliability value which ranging 0.71 to 0.81. The collected data were analysed using the Racsh model in order to convert the data in logit value.

Using Winsteps software, the data was analysed in order to determine the validity and reliability of the MBI. **Table 1** shows the item reliability and construct validity of 254 measured persons. The person reliability which is equivalent to KR20 has acceptable value of 0.81 (Sekaran, 2003). This indicates that consistency level of person arrangement in answering the same constructs exists despite using different set of items (Wright & Masters, 1982). Based on the separation index value of 2.07, it shows that there are 3 levels of respondent ability in terms of their mathematics beliefs.

The item reliability analysis is shown in **Table 2** which indicates a reliability value of 0.98. It is an acceptable value for which the instrument to be consistent if given to another sample that has the same and near features (Arasinah, Bakar, Ramlah, Soaib, & Zaliza, 2015); while the separation index shows a value of 6.27 which indicates the scale of MBI can be divided into 6 levels of difficulty. The six levels are categorized according to the respondents' ability in answering the items of the questionnaire.

Both values of the reliability person and item respectively have fulfilled the cut-off point as suggested by Bond & Fox (2007). In addition, the separation index for both item and person exceeds 2.0 which are considered good

Table 1. Person reliability: mathematics beliefs.

	Total Score	Count	Measure	Model Error	Infit		Outfit	
					MNSQ	ZSTD	MNSQ	ZSTD
Mean	137.3	36.0	.84	.23	1.07	−.1	1.04	−.1
S.D.	11.8	.0	.61	.03	.74	2.3	.73	2.3
Max.	170.0	36.0	3.20	.36	4.27	6.7	4.50	7.6
Min.	93.0	36.0	−.79	.17	.17	−4.6	.16	−4.5
Real RMSE	.27	True SD	.55	Separation	2.08		Person Reliability	.81
Model RMSE	.23	True SD	.57	Separation	2.50		Person Reliability	.86
	S.E of Person Mean = .04							

Summary of 254 measured person.

Table 2. Item reliability: mathematics beliefs.

	Total Score	Count	Measure	Model Error	Infit		Outfit	
					MNSQ	ZSTD	MNSQ	ZSTD
Mean	968.4	254.0	.00	.09	1.00	−.1	1.04	.5
S.D.	123.2	.1	.57	.01	.15	1.6	.21	2.0
Max.	1112.0	254.0	2.06	.11	1.55	5.3	1.66	5.9
Min.	502.0	254.0	−.73	.06	.82	−1.5	.78	−2.0
Real RMSE	.09	True SD	.56	Separation	6.27		Item Reliability	.98
Model RMSE	.09	True SD	.56	Separation	6.38		Item Reliability	.98
		S.E of Person Mean = .10						

Summary of 36 measured item.

(Fisher Jr., 2007). Person and item separation indices are recommended to be good between 2 to 3 and good at 5 (Fisher Jr., 2007). The more levels of difficulty the better the measurement reflecting different kinds of person's ability.

3. Result and Discussion

Figure 1 shows the Person-Item Map Distribution (PIDM) which indicates the person ability on the left side and the item difficulty on the right side. The top left is allocated for person with high ability in terms of agreeable level. While the low ability person who has less agreeable level is placed at the lower part of the scale. Based on the item difficulty, the most difficult item is K6 which is followed by K11and K26; and the least difficult items are K24 and K34. The statement for item K6 is given by "In quadratic equation, the enhancement should be given to the skill of using formula in getting correct answer". While the easiest items in the instrument include item K24 which states "Mathematics learning is an active process" and item K34 states "In learning mathematics, understanding mathematics concept, principal and strategy is important". The overlapping items at different level of difficulty are considered identical in terms of measuring the same construct (Bond & Fox, 2001). Overlapping items include K10 and K22; K18, K20, K23 and K32, K13; K17, K19, K25, K30 and K7; K16, K21, K28 and K3; K12, K33, K8 and K9; K2, K35, K36, K4 and K5; K14, K15 and K27; and K24 and K34.

The item difficulty is not equally distributed among respondents of the study. Most items are considered to be at easy level which is not match with the respondents' ability. The distribution of the person ability is towards high ability. Three respondents are found to be at the top scale of person ability and all of the respondents exceed the mean of the person ability.

Next, the polarity of items is shown by **Figure 2** which is used to validate the construct validity of the instrument. All items have positive values of Point Measure Correlation (PTMEA CORR) that are more than 0.30 except item K6, K26, K1, K5 and K27. However, positive values of PTMEA designate that all items measure the required construct of mathematics beliefs. It is an indicator whether the items are in line with the measured construct.

Figure 2 is a screenshot of Rasch model output showing the Infit MNSQ and Outfit MNSQ values of the MBI. Item K6 has to be removed since the value Infit and Outfit value respectively is more than the suggested value of 1.44 (Bond & Fox, 2007). The recommended range of these values is 0.60 to 1.40 if the instrument uses Likert scale type (Bond & Fox, 2007). An item with low fit index value shows that they are overlapping with other items and high value indicates nonhomogeneous with the other. The most difficult item of K6 has the highest Infit/Outfit value which contradicts to the easiest items of K24 and K34 which has the lowest index mentioned.

4. Conclusion

The importance of this study is to provide psychometrical characteristics of mathematics beliefs instrument. The findings of this analysis provide credibility of the MBI using Rasch Model. Although there are different levels
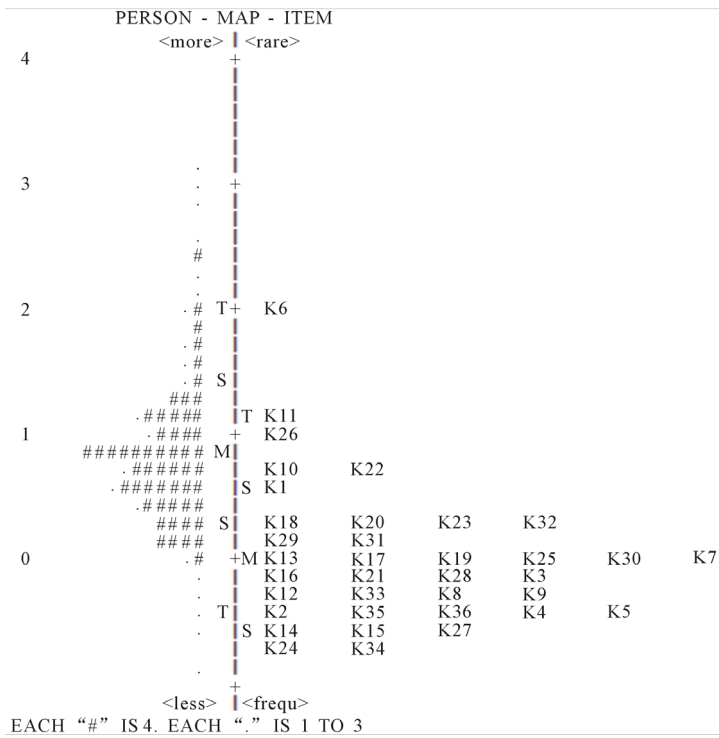
```
                         PERSON - MAP - ITEM
                       <more> | <rare>
          4                   +
                              |
                              |
                              |
                              |
                         .    |
          3               .   +
                          .   |
                              |
                          .   |
                          #   |
                          .   |
                          .   |
          2              .#  T+  K6
                          #   |
                         .#   |
                         .#   |
                         .#  S|
                        ###   |
                      .#####  |T K11
                       .####  +  K26
          1      ########### M|
                      .#####   |   K10        K22
                     .#######  |S K1
                       .#####  |
                        ####  S|  K18    K20      K23      K32
                        ####   |  K29    K31
          0               .#  +M K13    K17      K19      K25    K30    K7
                          .    |  K16    K21      K28      K3
                               |  K12    K33      K8       K9
                          .   T|  K2     K35      K36      K4     K5
                          .    |S K14    K15      K27
                               |  K24    K34
                          .    |
                               +
                       <less> | <frequ>
          EACH "#" IS 4. EACH "." IS 1 TO 3
```

Figure 1. The person-item map distribution conclusion.

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	ITEM	G
6	502	254	2.06	.07	1.55	5.3	1.66	5.9	.12	.50	36.2	40.1	K6	0
11	689	254	1.17	.06	1.27	3.2	1.40	4.5	.30	.51	31.9	36.5	K11	0
26	758	254	.95	.06	1.31	3.8	1.47	5.3	.27	.49	34.6	35.1	K26	0
10	840	254	.67	.07	1.04	.5	1.15	1.7	.39	.43	44.5	42.4	K10	0
22	863	254	.67	.07	1.19	2.0	1.27	2.7	.30	.45	31.1	41.5	K22	0
1	846	254	.60	.07	1.30	3.4	1.44	4.7	.22	.46	32.7	38.7	K1	0
18	986	254	.30	.09	.96	-.4	.98	-.2	.42	.37	60.6	57.9	K18	0
20	922	254	.28	.08	1.04	.4	1.13	1.3	.38	.41	49.2	50.2	K20	0
32	1012	254	.26	.09	.92	-.8	.95	-.5	.45	.37	63.0	59.6	K32	0
23	891	254	.24	.08	1.01	.2	1.02	.3	.41	.42	41.7	44.1	K23	0
29	953	254	.13	.08	1.03	.3	1.07	.7	.40	.42	50.4	49.1	K29	0
31	949	254	.13	.08	.96	-.3	1.04	.4	.46	.41	53.5	51.3	K31	0
13	957	254	.04	.08	1.00	.1	1.02	.3	.42	.40	53.9	50.1	K13	0
25	914	254	.03	.07	1.05	.6	1.14	1.6	.37	.43	42.1	43.3	K25	0
17	962	254	.00	.08	.87	-1.5	.89	-1.2	.53	.42	52.8	48.5	K17	0
30	963	254	-.02	.08	.90	-1.0	.92	-.8	.51	.39	54.7	53.9	K30	0
19	975	254	-.04	.10	.89	-.9	.86	-1.2	.45	.34	66.5	64.5	K19	0
7	962	254	-.07	.09	.92	-.7	.94	-.6	.48	.37	56.3	57.9	K7	0
28	992	254	-.15	.08	.89	-1.2	.93	-.7	.50	.40	50.0	49.9	K28	0
3	1009	254	-.18	.10	.91	-.6	.94	-.4	.45	.34	68.9	67.6	K3	0
21	993	254	-.20	.09	.90	-.9	.92	-.8	.47	.38	60.2	57.4	K21	0
16	1031	254	-.21	.08	1.05	.5	1.16	1.3	.34	.37	63.0	59.4	K16	0
8	1038	254	-.27	.11	.83	-1.2	.78	-1.7	.52	.33	72.8	69.8	K8	0
33	1040	254	-.34	.09	.89	-.9	.95	-.4	.47	.36	61.4	59.1	K33	0
12	1084	254	-.35	.10	.92	-.5	.90	-.7	.41	.34	65.0	60.8	K12	0
9	991	254	-.35	.09	.95	-.4	.97	-.2	.43	.37	64.2	60.1	K9	0
4	1059	254	-.36	.10	.90	-.7	.92	-.6	.44	.33	69.3	65.9	K4	0
5	1110	254	-.37	.09	1.09	.7	1.41	2.7	.27	.35	57.5	54.3	K5	0
35	1055	254	-.39	.11	.99	-.1	.98	-.2	.35	.33	65.4	64.0	K35	0
2	1018	254	-.49	.08	1.00	.0	1.02	.2	.40	.39	54.3	53.2	K2	0
36	1050	254	-.49	.11	.94	-.4	.88	-.9	.39	.32	67.7	69.0	K36	0
27	1073	254	-.59	.10	1.02	.2	1.16	1.4	.29	.33	62.2	61.5	K27	0
15	1083	254	-.62	.11	.87	-1.0	.84	-1.4	.47	.32	66.1	62.6	K15	0
14	1085	254	-.64	.10	.92	-.6	.91	-.8	.44	.34	62.6	59.2	K14	0
24	1096	254	-.66	.10	.82	-1.5	.78	-2.0	.51	.32	65.7	59.7	K24	0
34	1112	254	-.73	.11	.85	-1.5	.82	-2.0	.50	.30	66.5	60.7	K34	0
MEAN	968.4	254.0	.00	.09	1.00	.1	1.04	.5			55.5	54.4		
S.D.	123.2	.1	.57	.01	.15	1.6	.21	2.0			11.7	9.4		

Figure 2. Screen shots of Rasch model output showing item distribution according to difficulty level in mathematics beliefs.

of difficulty and ability, high consistency should be justified based on the reliability of person and item respectively. The findings have become the statistical evidence in validating the instrument which should be consistent with theoretical expectations (Ariffin & Abdul Hamid, 2009). Misleading items can be either replaced or omitted from being included in the instrument. For future work, the study can be discussed further based on gender and years of teaching experience. The discussion of the mathematics beliefs instrument can be extended to misfit behaviours, differential item functioning, and other related item analysis using Rasch Model. The Rasch output has created a paradigm shift in measuring respondent's perception using instrument. A meaningful data has been obtained through Rasch model output. Individual item checking can be done which leads to a quality instrument.

References

Abdul Aziz, A., Masodi, M. S., & Zaharim, A. (2013). *Asas Model Pengukuran Rasch: Pembentukan Skala & Struktur Pengukuran.* Bangi: Penerbit Universiti Kebangsaan Malaysia.

Ali, S., Mofreh, M., Najib, M., Ghafar, A., Hafiz, A., Omar, H., & Mosaku, M. (2014). Psychometric Properties on Lecturers' Beliefs on Teaching Function : Rasch Model Analysis. *International Education Studies, 7,* 47-55.

Arasinah, K., Bakar, A. R., Ramlah, H., Soaib, A., & Zaliza, H. (2015). Using Rasch Model and Confirmatory Factor Analysis to Assess Instrument for Clothing Fashion Design Competency. *International Journal of Social Science and Humanity, 5,* 418-421. http://dx.doi.org/10.7763/IJSSH.2015.V5.492

Ariffin, S. R., & Abdul Hamid, N. A. (2009). Critical Thinking Skills Profile between the Science and Non-Science Students. *Education Deans' Council Journal, 3,* 1-22.

Beswick, K. (2005). The Beliefs/Practice Connection in Broadly Defined Contexts. *Mathematics Education Research Journal, 17,* 39-68. http://dx.doi.org/10.1007/BF03217415

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch Model: Fundamental Measurement in the Human Science* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Çakiroğlu, E., & Işıksal, M. (2009). Preservice Elementary Teachers' Attitudes and Self-Efficacy Beliefs toward Mathematics İlköğretim Öğretmen Adaylarının Matematiğe Yönelik Tutum ve Özyeterlik Algıları. *Education and Science, 34,* 132-139.

Evans, D. B. (2003). *Early Childhood (K-5) Preservice Teachers' Beliefs about Mathematics, Teaching Mathematics and Learning Mathematics*. Dissertation of Doctorate in Education. Georgia: Southern University.

Fisher Jr., W. P. (2007). Rating Scale Instrument Quality Criteria. *Rasch Measurement Transaction, 21,* 1095. http://www.rasch.org/rmt/rmt211a.htm

Head, F. C. (2012). *Comparative Analyses of Mathematics Teachers' Efficacy Using Factor Analysis and the Rasch Model*. Ph.D. Thesis, Georgia: Kennesaw State University.

Muijs, D., & Reynolds, D. (2002). Teachers ' Beliefs and Behaviours: What Really Matters? *Journal of Classroom Interaction, 37,* 3-15.

Pampaka, M., & Williams, J. (2010). Measuring Mathematics Self Efficacy of Students at the Beginning of Their Higher Education Studies. *British Congress for Mathematics Education,* 159-166.

Saidfudin, M., Azrilah, A., Rodzo'an, N., Omar, M., Zaharim, A., & Basri, H. (2010). Easier Learning Outcomes Analysis Using Rasch Model in Engineering Education Research. *Latest Trends on Engineering Education,* 442-447.

Swan, M. (2002). Designing and Using Research Instruments to Describe the Beliefs and Practices of Mathematics Teachers. *Research in Education, 75,* 58-70. http://dx.doi.org/10.7227/RIE.75.5

Wilkins, J. L. M. (2008). The Relationship among Elementary Teachers' Content Knowledge, Attitudes, Beliefs, and Practices. *Journal of Mathematics Teacher Education, 11,* 139-164.