A RASCH ANALYSIS OF THE INTEGRATIVE STEM TEACHING INTENTION QUESTIONNAIRE AMONG PRE-SERVICE TEACHERS IN SABAH

Analisis Rasch Terhadap Soal Selidik Integrative Stem Teaching Intention Questionnaire dalam Kalangan Guru Praperkhidmatan di Sabah

Chiam Sun May,PhD
Ng Lee Fong, PhD,
Tan Seok Kiang
Ronald Yusri Batahon, PhD
Institut Pendidikan Guru Kampus Kent
sunmayc@yahoo.com

ABSTRACT

The purpose of this study is to investigate the reliability and validity of the Integrative STEM Teaching Intention Questionnaire developed by Lin and Willians (2015) in determining the pre-service teacher's STEM teaching intention in Sabah, Malaysia after being adapted and translated into Malay language. Data were obtained from a total of 40 respondents via Google form. The survey questionnaire consists of two parts. Part A consists of three items, related to the demography of the respondents, while Part B consists of 30 close-ended items using five-point Likert scale to measure the perceptions of the respondents on the six aspects namely knowledge, value, attitude, norm, behaviour and desire. Responses from the questionnaire were analysed quantitatively using Rasch analysis. The findings reveal that all the items fulfilled the quality criteria but the overall spread of the items measure logits across the Wright Map continuum were at the lower position as compared to the person measure logits. This study concluded that the instrument is reliable, valid and suitable to be employed to collect data to investigate the integrative STEM teaching intention among the pre-service teachers in Sabah and generally in Malaysia

Keywords: rasch analysis, integrative STEM, pre-service teachers

ABSTRAK

Tujuan kajian ini adalah untuk mengkaji kebolehpercayaan dan kesahan Integrative STEM Teaching Intention Questionnaire yang dibangunkan oleh Lin dan Willians (2015) dalam menentukan niat mengajar STEM integratif dalam kalangan guru praperkhidmatan di Sabah, Malaysia. Instumen ini diadaptasi dan diterjemah ke dalam bahasa Melayu. Data diperoleh daripada 40 responden melalui Google form. Soal selidik ini terdiri daripada dua bahagian. Bahagian A terdiri daripada tiga item, yang berkaitan dengan demografi responden, sementara bahagian B terdiri daripada 30 item soalan tertutup yang menggunakan skala Likert lima mata untuk mengukur persepsi responden terhadap enam aspek iaitu pengetahuan, nilai, sikap, norma, sikap dan keinginan. Respon soal selidik dianalisis secara kuantitatif menggunakan analisis Rasch. Hasil kajian menunjukkan bahawa semua item memenuhi kriteria kualiti yang diperlukan tetapi penyebaran keseluruhan logits ukuran item (item measure logits) pada kontinum Peta Wright adalah berada pada kedudukan yang lebih rendah berbanding dengan logits ukuran orang (person measure logits). Kesimpulan kajian ini ialah intrumen ini mempunyai kebolehpercayan dan kesahan yang baik serta sesuai untuk digunakan untuk mengumpul data dalam kajian niat guru praperkhidamatan mengajar STEM integratif di Sabah dan amnya Malaysia.

Kata kunci: analisis rasch, STEM integratif, guru praperkhidmatan

INTRODUCTION

Ministry of Education Malaysia launched Science Technology Engineering Mathematics (STEM) education initiative that stated in Malaysia Education Blueprint (2013 – 2025), namely to increase the interest of students and teachers awareness of STEM education. The aim of STEM education initiative is to prepare students with the skill to meet the challenges of science and technology and to ensure Malaysia has sufficient number of qualified STEM graduates to fit in the industry revolution 4.0. Teachers play a very important role to achieve the objectives of the Education ministry. Teachers are expected to have interdisciplinary teaching skills that allow students to master the concepts of science, technologies, engineering and mathematics. In addition they are able to integrate the knowledge of the four disciplines to solve problems in daily life. STEM integration in Malaysia was emphasis in curriculum by KSSM and KSSR revision in 2017. This shows that the Ministry of Education Malaysia was takes a step to change. The curriculum has been put STEM agenda approach as a core element in the construction and implementation which emphasis pedagogical teaching in-depth approach to teaching and learning based on higher order thinking skills (HOTS), inquiry-based learning, problem solving, project-based learning and related skills with the integration of STEM (Suraya, Norsalawati & Nasir, 2017). Therefore, research studies are suggesting the need for validate an instrument to measure pre-service teachers' intentions for integrative STEM teaching.

LITERATURE REVIEWS

The interest in examining effects of the integrative STEM teaching intention on the students' perception was intensively studied in Turky (Sari, Alici & Sen, 2017). Their findings revealed that the students' attitudes towards STEM disciplines, STEM career interests in STEM-related occupations significantly increased after the integrative teaching with STEM implemented in their learning system. The students stated that integrated STEM teaching was effective in their learning and also it helps to develop their skills of the 21st century. Inot her side of view, the findings in Thailand from Pimthong and Williams (2018) stated that it is important of promoting pre-service teachers' understanding of the integrated of STEM and how these four disciplines were integrated. Therefore the researcher suggested the STEM teachers' preparation program should be different from science, technology or mathematics teachers because it represents the integration of the four disciplines to enhance their teaching intention. According to Sujarwanto, Madlazim, and Ibrahim (2019) learning design that connects interdisciplinary in STEM in education is a complex activity and it require teachers that able to make their students in understand how to implement STEM in real-world problems. Therefore, teachers those teach with intention are able to teach more effectively.

The Integrative STEM Teaching Intention Questionnaire developed by Lin and Willians (2015) was adapted and revalidated by Hacıömeroğlu and Bulut, (2016). This questionnaire was adapted into Turkish context to examine elementary pre-service teachers STEM teaching intentions level. The same instrument was further used by Doğan and Benzer (2019) and Koçak, Aslan and Capellaro (2019) to investigate the integration of STEM in teaching context. By the literature review, this instrument was not yet in used for Malaysia context. However, there have been studies conducted on STEM education in Malaysia (Kamisah & Rohaida, 2014; Suraya, Norsalawati & Nasir, 2017) but not research on the intention of the primary pre-service teachers' STEM teaching intention. Therefore, this study is the first, to the authors' knowledge, to validate the instrument Integrative STEM Teaching Intention Questionnaire (InSTEMtiq) adapted from Lin & Williams (2015) for Sabah pre-service teacher in Malaysian context.

Messick (1993) stated that instruments should always be piloted for new settings and new populations as "existing validity evidence becomes enhanced (or contravened) by new findings".

There are six aspects to be evaluated in the InSTEMtiq namely Knowledge (K), Value (V), Attitude (A), Norm (N), Behaviour (B) and Desire (D). Knowledge of STEM aspect referred to interdisciplinary knowledge of understanding and experience about science, technology, engineering and mathematics, which might influence their professional education. In other words, teachers should be knowledgeable in their chosen discipline to effectively synthesize this knowledge with other STEM disciplines particularly in some cases such as arts, humanities, and social sciences (Lin & Williams, 2015).

In the study of Sujarwanto et al. (2019), STEM knowledge is referred as all the information that the teacher has about the characteristics of STEM, the advantages, and the disadvantage of STEM and the integration between disciplines in STEM. The value aspect concerns the implementation and on the manner in which resulting in positive or negative remarks when teaching in classroom. It emphasizes on any changes in an individual's set of criteria regarding STEM teaching and the way these affect one's self-evaluation and evaluation of students' remarks concerning its practice (Lin & Williams, 2015). Attitude aspect is focussed on the interest in STEM teaching of the person (Lin & Williams, 2015). While another researcher defines the attitude towards STEM as the teacher's level of curiosity towards STEM Education, the teacher's agreement to apply the integration of STEM during learning, and what the teacher thinks about STEM Education.

Therefore, in this study pre-service teachers' attitude toward STEM embodies their personal willingness to apply STEM in teaching. On the other hand, norms aspects referred to the impressions of reference groups from school about their support of or opposition to STEM teaching implementation and the individual's degree of compliance with these norms. According to Lin and Williams (2015) perceived behavioural control with respect to STEM teaching entails the prospect of pre-service teachers able to controlling proper resources and resolving difficulties successively. Hence, behaviour control may affect the decision to execute or prevent certain actions based on the availability of resources. The last aspect discussed in this study named as desire, which an individual gives intention to adopt STEM teaching in future as teaching career. Therefore, intention is a major component of the psyche. The effectiveness of teaching and learning also linked to the teaching intention of teachers. In other word intention is the will directs attention, the guide behind the making of choices, propensities and dispositions to act, which are components of affect (Manson, 2010).

OBJECTIVE OF THE STUDY

The objectives of this study were to investigate the reliability and validity of the InSTEMtiq which was adapted from Lin and William (2015), to measure the STEM teaching intention among the pre-service teachers in Malaysian teachers' institution in Sabah. The reliability and validity was conducted using Rasch analysis.

RESEARCH METHODS

This study was a pilot test to the initial set of InSTEMtiq which was translated into Malay language for the used in Malaysian context. Each item from all aspects were analysed using Rasch analysis in order to identify if there is any issues with the items or scales before it administered.

Instrument

The respondents were being briefed about the purpose of the study regarding to the integrative STEM Teaching Intention. The InSTEMtiq consists of 30 items which were categorised into 2 parts. Part A

consists of three items related to demography of the respondents. Part B consists of 30 items with five point Likert scale to measure the attainment of the respondents on the six aspects. There were four items on knowledge aspect, six items on value aspect, five items on attitude aspect, norm aspect, behaviour aspect and desire aspect respectively as in Table 1.

Table 1
Aspects and Items in Questionnaire

Part	Aspect	Number of Items	Item
A	Demography	3	Dm1, Dm2, Dm3
	Total	3	
B	Knowledge (K)	4	K1, K2, K3, K4
	Value (V)	6	V1, V2, V3, V4, V5, V6
	Attitude (A)	5	A1, A2, A3, A4, A5
	Norm (N)	5	N1, N2, N3, N4, N5
	Behaviour (B)	5	B1, B2, B3, B4, B5
	Desire (D)	5	D1, D2, D3, D4, D5
	Total	30	
Grand Total		33	

Target Sample

This study focused on pre-service teachers as they are part of the future implementers of STEM program. Respondents were invited to response to the questionnaire through online using Google form. There were 40 respondents who responded to the questionnaire. The demography of the respondents is shown in Table 2.

Table 2
Demography of the Respondents

	Frequencies (n)	Percent (%)
Gender (Dm1)		
Male	14	35.0
Female	26	65.0
Option (Dm2)		
PISMP Science Major	18	45.0
PISMP Mathematics Major	18	45.0
PISMP Science Elective	4	10.0
Year (Dm 3)		
Year 1 Semester 2	24	60.0
Year 3 Semester 2	7	17.5
Year 4 Semester 2	9	22.5

Data Analysis

Data were analysed using Winsteps (Linacre, 2016), an analysis programme based on Rasch model (Alagumalai, Curtis, & Hungi, 2005) and according to the clusters. Rasch measurement is a technique where measurement scales such as surveys can be constructed and used in research. The technique is based upon

performing a common measurement instruments such as a measuring stick that measures in the same way from group to group, and there is a clear understanding of what is being measured by the measuring stick. The Rasch measurement can be accepted as allowing for the construction of robust instruments to measure human traits.

In Rasch analysis, the Summary Statistics provided general information on the quality of the respondents and the instrument, as well as the interaction between persons and items. Sumintono and Widhiarso (2014) explain some important indicators in the summary statistics. The person measure which is far from logit 0.00 means that the respondents had the tendencies to agree to the statements in the items. Cronbach Alpha value that reads> 0.70 showed that the reliability score is very good. The person reliability index and item reliability that reads > 0.80 is good. While the infit MNSQ (mean square) and outfit MNSQ which near to 1.00 are good, and the infit ZSTD (standardized as Z score) and outfit ZSTD which near to 0.00 are good. Person separation indicates how efficiently a set of items is able to separate those persons measured (Wright, 1999). It shows that the bigger the person separation index means it is more likely that the respondents will response correctly to the items. While the item separation index indicates how well a sample of people is able to separate those items used in the test (Wright, 1999). The index suggest the spread of items, wider spread means the items are better and fitter. It can be concludes that the higher the value of reliabilities, the better the separation that exists and the more precise the measurement.

The item fit order provides information on the precision of measurement. It tells the difficulties and the qualities of the items. According to Bond and Fox (2015), the outfit MNSQ, the outfit ZSTD and the point measure correlation (Pt Meas Corr) are the criteria that need to be examine in order to improve the item fit. Sumintono and Widhiarso (2014) and Boone, Staver and Yale (2014) give some explanations on the infit, outfit, MNSQ and ZSTD. The infit MNSQ tells the fitness of the items. An item is misfit if the logit is greater than the sum of mean and S.D. The acceptable outfit MNSQ value is $0.50 < MNSQ < 1.50$, outfit ZSTD value is $-2.00 < ZSTD < +2.00$.

The Point-Measure Correlation (Pt Meas Corr) value is $0.40 <$ Pt Meas Corr < 0.85. The Item Polarity is determined by the value of Pt Meas Corr. If the value of the correlation coefficient is positive, the item ability in measuring the the construct level is valid. If the value of Pt Meas Corr is negative or near zero, this showed that the relationships for response items are contradict and not consistent with the construct (Linacre 2002, Bond and Fox 2015). If the value of the correlation coefficient is out of this range but the expected correlation (EXP) value is near to the Pt Meas Corr value, then the items still can be considered to be accepted.

The Item-person Variable Map is Winsteps technique that able to map the conceptual understanding of the respondents against the items with varying level of difficulty. It is a simultaneous evaluation of both items and person expressed in a logit scale displayed.

Validity refers to the extent to which a test measures what it intends to measure. One of the important aspects of construct validity is the trustworthiness of the scores and its interpretation. Investigating the unidimensionality of item response data is an essential component of construct validity. Unidimentionality is the fundalmental requirement in construct validity (Engelhard Jr. 2013). The variance explained by measures needs to meet a minimum requirement of 20% and considered good if it reaches 40%. While the unexplained variance in contrast should not be > 15% (Sumintono & Widhiarso, 2014). The rating scale is able to verify if the rating confused the respondents (Sumintono & Widhiarso, 2014). It refers to the existence of one underlying measurement construct (dimension) that accounts for variation in examinee responses. Violating this assumption could severely bias item and ability parameter estimation (Chong, Popp, DiGangi, Jannasch-Pennell, 2007). Therefore, it is important to establish an unidimensionality structure to provide the evidence of internal consistency.

FINDINGS AND DISCUSSION

The data was analysed based on Rasch Measurement Model for reliability and validity test. The reliability and separation index, item polarity, item fit, item dimensionality and variable maps were analysed.

Reliability and Separation Index

From Table 3, the findings showed that the Person Reliability Index was 0.89, indicated that the Person Reliability Index is good. The Person Separation Index was 2.83, indicated that there were four levels of person ability. While the Item Reliability Index was 0.93, showed that the Item Reliability Index was good, which is acceptable and considered strong evidence for person sample to confirm the item difficulty hierarchy (Linacre & Wrignt, 2012) The Item Separation Index was 3.53, indicated that there were five levels of item difficulty. On the other way of analysis, the Cronbach Alpha KR-20 value was 0.92, showed that the reliability of the items was good. Table 3 Showed the Reliability and Separation Index.

Table 3
Reliability and Separation Index

Total Item	Cronbach Alpha	Reliability		Separation Index	
30	0.92	Person	Item	Person	Item
		0.89	0.93	2.83	3.53

Item Polarity

Table 4 showed all the Pt-Meas-Corr values were positive for all the 30 items. However, item D5 (Pt Meas Corr = 0.35, EXP = 0.49), item V4 (Pt Meas Corr = 0.38, EXP = 0.39), item K1 (Pt Meas Corr = 0.36,EXP = 0.45), item V2 (Pt Meas Corr = 0.38, EXP = 0.37) and item V1 (Pt Meas Corr = 0.31, EXP = 0.33), showed that the Pt-Meas-Corr values were less than 0.4 but the EXP value were close to the Pt-Meas-Corr value (0.01 – 0.15). Therefore, all the items can be accepted. Table 4 showed the Item Polarity.

Table 4
Item Polarity

Item	Pt Meas Corr	EXP
D5	**0.35**	**0.49**
K3	0.45	0.63
K2	0.49	0.60
N3	0.52	0.62
V4	**0.38**	**0.39**
D1	0.44	0.50
K1	**0.36**	**0.45**
D3	0.42	0.50
K4	0.46	0.52
V2	**0.38**	**0.37**
V1	**0.31**	**0.33**
A2	0.58	0.54
N4	0.61	0.64
A5	0.54	0.53
N2	0.60	0.64
A4	0.56	0.52

A1	0.63	0.58
N2	0.64	0.64
V3	0.53	0.42
V6	0.41	0.40
D2	0.44	0.45
B5	0.58	0.52
D4	0.58	0.48
A3	0.56	0.56
V5	0.62	0.53
N1	0.69	0.62
B4	0.58	0.51
B3	0.64	0.50
B1	0.71	0.53
B2	0.72	0.52

Item Fit Analysis

The findings from Table 5 showed the fit analysis for the items. The item infit MNSQ was between 0.52 and 1.66, and the item outfit MNSQ was between 0.48 and 2.40. The infit ZSTD was between -2.40 and 2.50, while the outfit ZSTD was between -2.4 and 3.50. There were three items' (H5, P3 and P2) result are having infit MNSQ and Outfit MNSQ which were out of the range of 0.50 < MNSQ <1.50. The same items infit ZSTD and outfit ZSTD values were located outside of the acceptable range of -2.0 < ZSTD < 2.0. These three items should be closely checked and perhaps improved before use with a larger, broader sample. Table 5 showed the Item Fit analysis.

Table 5
Item Fit Analysis

Item	Measure	Model S. E.	Infit		Outfit	
			MNSQ	ZSTD	MNSQ	ZSTD
D5	-0.59	0.32	**1.53**	**2.00**	**2.40**	**3.50**
K3	1.88	0.25	**1.66**	**2.50**	**1.59**	**2.30**
K2	1.28	0.26	**1.59**	**2.30**	**1.50**	**2.00**
N3	1.62	0.26	1.44	1.80	1.34	1.50
V4	-1.73	0.40	1.42	1.40	1.04	0.30
D1	-0.39	0.31	1.20	0.90	1.38	1.30
K1	-1.04	0.35	1.36	1.40	1.19	0.60
D3	-0.39	0.31	1.12	0.60	1.35	1.20
K4	-0.20	0.30	0.92	-0.30	1.30	1.10
V2	-1.90	0.42	1.22	0.80	0.77	-0.30
V1	-2.29	0.47	0.95	0.00	1.18	0.50
A2	0.15	0.29	1.16	0.80	1.12	0.60
N4	2.25	0.25	1.13	0.60	1.12	0.60
A5	-0.02	0.30	1.02	0.20	1.03	0.20
N5	2.13	0.25	0.99	0.00	1.03	0.20
A4	-0.20	0.30	0.89	-0.40	0.96	-0.10
A1	0.93	0.27	0.94	-0.20	-0.95	-0.20
N2	2.07	0.25	0.89	-0.40	0.92	-0.30
V3	-1.43	0.37	0.91	-0.20	0.62	-0.90
V6	-1.57	0.39	0.89	-0.30	0.86	-0.20
D2	-1.04	0.35	0.86	-0.50	0.88	-0.20
B5	-0.11	0.30	0.85	-0.60	0.82	-0.60

D4	-0.70	0.33	0.82	-0.70	0.80	-0.60
A3	0.48	0.28	0.74	-1.20	0.78	-1.00
V5	0.07	0.29	0.72	-1.30	0.70	-1.30
N1	1.62	0.26	0.60	-2.00	0.66	-1.70
B4	-0.29	0.31	0.62	-1.80	0.65	-1.40
B3	-0.39	0.31	0.61	-1.90	0.55	-1.80
B1	-0.02	0.30	0.56	-2.20	0.53	-2.20
B2	-0.20	0.30	0.52	-2.40	0.48	-2.40

Item Dimensionality

Table 6 showed the value of raw variance explained by measures was 48.20% which was less than 50% that will be located in poor category (Fisher, 2007). However it still can be accepted as the raw variance explained by measures was more than 40% (Sumintono & Widhiarso, 2014). The unexplained variance in the first contrast was 7.1, which indicates that it has the strength of about 7 items (7.1 rounded to 7) out of 30 items. In addition, the unexplained variance in the first contrast was 12.20% which was less than 15% of accepted value (Sumintono & Widhiarso, 2014). The results suggested the unidimensionality structure of the aspects in inSTEMtiq to provide the evidence of internal consistency. Table 6 showed the Standardized Residual Variance for the items.

Table 6
The Standardized Residual Variance for the items

30 Items	Empirical			Modeled (%)
Total raw variance in observation	57.	100.0%		100.0%
Raw variance explained by measures	27.9	**48.2%**		48.8%
Raw variance explained by persons	11.9	20.6%		20.8%
Raw variance explained by items	16.0	27.7%		28.0%
Raw unexplained variance (total)	30.0	51.8%	100.0%	51.2%
Unexplained variance in 1st contrast	**7.1**	**12.2%**	23.5%	
Unexplained variance in 2nd contrast	3.9	6.7%	13.0%	
Unexplained variance in 3rd contrast	3.1	5.3%	10.2%	
Unexplained variance in 4th contrast	2.3	3.9%	7.5%	
Unexplained variance in 5th contrast	2.1	3.6%	7.0%	

Wright Map

Figure 1 showed the distribution of the position of persons on the left side of the vertical line and items on the right side. The items side covered the range of 2.08 to -2.10 logits in level of difficulty. V1 was the easiest item to agree while V4 and V5 were the most difficult items to agree for the respondents. The highest person measure was 7.00 logits and the lowest person measure was 0.90 logits. The Wright Map showed that 31 person measure logits were above the highest item measure logits and 22 items measure of logits were placed below the lowest person measure logits. This implies that the items are not able to measure the ability of respondents meaningfully. Figure 1 showed the Wright Maps.

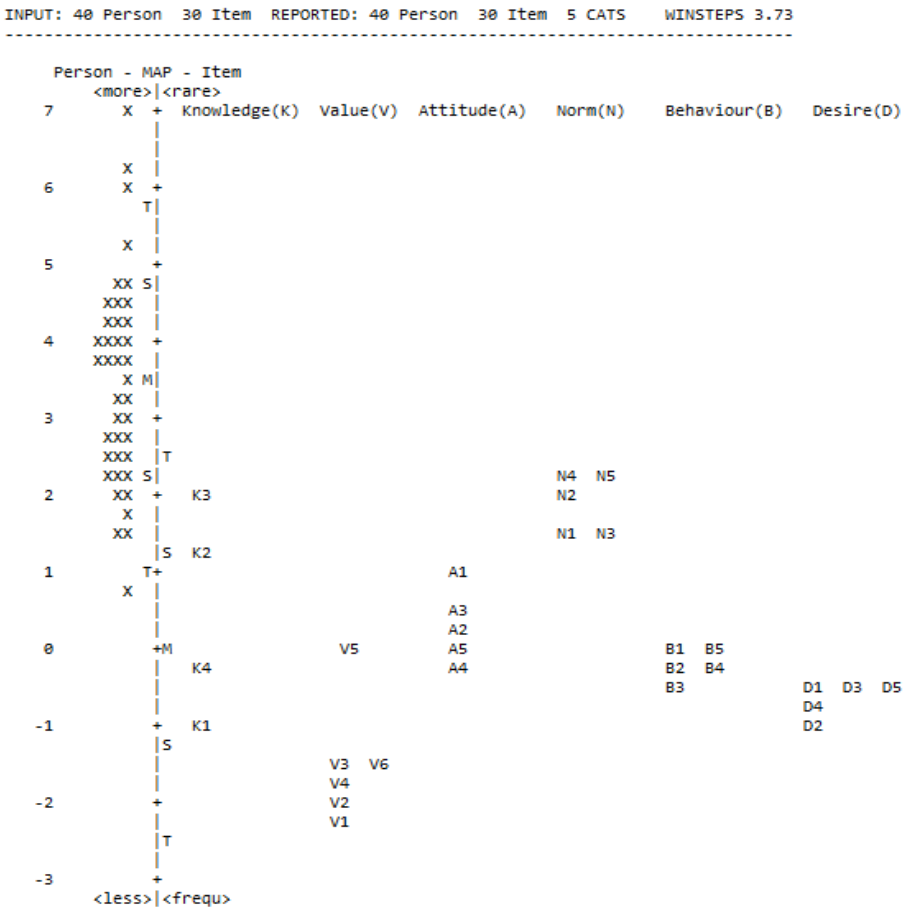

Figure 1. Wright Map

CONCLUSION

The objectives of Rasch analysis in this pilot study were to investigate the items and determine the quality of the items before the main study. The findings from the analysis found that both Person Reliability Index and Items Reliability Index were good. Beside that there were four levels of person ability and five levels of item difficulty indicates that this instrument is able to separate relevant persons by their performance (Wright, 1999).

From the analysis of Item Polarity, all items were good in general, even though there were five items (D5, V4, K1, V2 and V1) which were a little out of the optimum range. However, all items were acceptable. These five items were further examined on Item Fit Analysis. Three items (D5, K3 and K2) were slightly out from the optimum range. K3 was more disputable out of the three items, as the infit and outfit MNSQ, as well as the infit and outfit ZSTD was furthest away from the optimum range. The Item Dimensionality suggested that there were no critical issues concerning the items in the questionnaire and were located in acceptable range (Sumintono & Widhiarso, 2014). The Wright Maps suggested that all respondents were easy to agree with the items in general. The Items were easy to them as the logits of the items were at lower position as compared to the Persons logits. The gaps between the items were small, and the same scenario went to the gaps of the respondents.

As a conclusion, the findings indicate that there is not much improvement needed to be done on the items. The instrument is reliable, valid and suitable to be employed to collect data to investigate the integrative STEM teaching intention among the pre-service teachers in Sabah and generally in Malaysia.

REFERENCES

Alagumalai, S., Curtis, D.D., & Hungi, N. (2005). *Applied Rasch measurement: A book of exemplars*. Springer-Kluwer

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed). New York: Routledge.

Chong, H. Y. Popp, S. O. DiGangi, S., & Jannasch-Pennell. A. (2007). Assessing unidimensionality: A comparison of Rasch Modeling, Parallel Analysis, and TETRAD, *Practical Assessment, Research & Evaluation, 12*(14). Retrieved https://pareonline.net/

Doğan, T & Benzer, S. (2019). Investigation of science teacher candidates' opinions towards Science, Technology, Engineering And Math (STEM) Teaching. *Malaysian Online Journal of Educational Sciences, 7* (2), 1-9. http://mojes.um.edu.my/ EISSN: 2289-3024

Engelhard Jr, G. (2013). *Invariant measure: Using Rasch models in the social, behavioural and health sciences*. New York: Routledge.

Haciömeroğlu, G., & Bulut, A. S. (2016). Integrative STEM teaching intention questionnaire: a validity and reliability study of the turkish form. *Journal of Theory and Practice in Education*, *12*(3), 654-669

Kamisah Osman & Rohaida Mohd Saat. (2014) Editorial. Science Technology, Engineering and Mathematics (STEM) education in Malaysia. *Eurasia Journal of Mathematics, Science & Technology Education*, *10*(3), 153-154

Koçak, B., · Aziz Aslan, A., &· Cappellaro, E. (2019). Science, Mathematics and primary pre-service teachers' intention on STEM teaching. *Gönderilme Tarihi*, 168-188.

Linacre, J. M. (2016). *A User's guide to WINSTEPS MINISTEP Rasch-Model computer programs*. Beaverteon, Oregon: Winsteps.com.

Lin, K. Y. & Williams, J. (2015). taiwanese preservice teachers' science, technology, engineering, and mathematics teaching intention. *International Journal of Science and Mathematics Education*, 1-30, doi: 10.1007/s10763-015-9645-2

Linacre, J. M., & Wright, B. D. (2012). *A user's guide to WINSTEPS ministeps Rasch model computer programs*. Chicago: Mesa Press.

Manson, J. (2010). Attention and intention in learning about teaching through teaching. In R. Leikin & R. Zazkis (Eds.) *Learning through teaching mathematics: Development of teachers' knowledge and expertise in practice*. 23-47. New York: Springer

Messick, S. (1993). *Foundations of validity: Meaning and consequences in psychological assessment*. New Jersey: Educational Testing Service

Ministry of Education Malaysia. (2013). *Malaysia education blueprint 2013-2025 (Preschool to Post-Secondary Education)*. Putrajaya: Kementerian Pendidikan Malaysia

Pimthong, P. & Williams, J. (2018). Preservice teachers' understanding of STEM education. *Kasetsart Journal of Social Sciences,* In Press. https://doi.org/10.1016/j.kjss.2018.07.017

Sarı, U., Alıcı, M. & Sen, O. M. (2017). The Effect of STEM instruction on attitude, career perception and career interest in a problem-based learning environment and student opinions. *Electronic Journal of Science Education, 22*(1). Retrieved from http://ejse.southwestern.edu

Sumintono, B. Widhiarso, W. (2014). *Aplikasi model RASCH untuk penelitian ilmu-ilmu sosial.* (Edisi Versi). Cimahi, Indonesia: Trim Komunikata Publishing House. (In Indonesian).

Sujarwanto, E., Madlazim, Ibrahim, M. (2019). Attitude, knowledge, and application of STEM owned by science teachers. *Journal of Physics: Conference Series* 1417, 1-8 doi:10.1088/1742-6596/1417/1/012096

Suraya Bahrum, Norsalawati Wahid, Nasir Ibrahim (2017). Integration of STEM education in malaysia and why to STEAM. *International Journal of Academic Research in Business and Social Sciences,* 7 (6), 645-654

Wright B.D. (1999*).* Fundamental measurement for psychology. In S.E. Embretson & S.L. Hershberger (Eds.), *The new rules of measurement: What every educator and psychologist should know.* Hillsdale, NJ: Lawrence Erlbaum Associates.