

Item Analysis for the Adapted Motivation Scale Using Rasch Model

Ng Sar Ee¹, Kee Jiar Yeo², Azlina bt Mohd Kosnin³

¹ Language Department, Institut Pendidikan Guru Kampus Temenggong Ibrahim, Johor, Malaysia

^{2,3} School of Education, Universiti Teknologi Malaysia, Malaysia

Article Info

Article history:

Received Aug 4, 2018

Revised Sept 20, 2018

Accepted Oct 5, 2018

Keyword:

Item analysis

Construct validity

Rasch model

Integrative motivation

Instrumental motivation

ABSTRACT

This study measured validity and reliability of an adapted motivation scale used in second language learning. The Integrative and Instrumental motivation (IIM) was developed based on two orientations of motivation, which were integrative motivation and instrumental motivation. IIM was used to identify students' level of motivation towards learning Chinese language as a second language (CLSL) in national school. The study presented empirical evidence of validity and reliability using the Rasch Model. Pilot study was conducted in a national school which has students enrolled for the CLSL class. There were 29 students in this pilot study. Data was analyzed using WINSTEPS version 3.72.3. The result showed the reliability of the motivation scale is 0.87 based on Cronbach's alpha. The construct validity was determined by PT-measure correlation value (PMC) range from 0.71 to 0.86, infit and outfit MNSQ between 0.53 to 1.43 and ZSTD range from -1.1 to 1.2. In unidimensional measure, raw variance explained was at 66.4% with the unexplained variance in the first factor was at 13.6%. Finally the fit statistics showed that person separation index, 2.79 was considered good and item separation index, 1.81 was in acceptable range [6]. Person and item reliability were at 0.89 and 0.77 respectively. The result indicated that the new instrument with 15 items-IIM after eliminated unfit items was a reliable and valid instrument to measure the motivation in learning CLSL in the context of national school.

Copyright © 2018 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Ng Sar Ee,

Language Department,

Institut Pendidikan Guru Kampus Temenggong Ibrahim, Johor

Datin Halimah Street, 80350 Johor Bahru, Johor, Malaysia.

Email: ngsaree@ipgm.edu.my

1. INTRODUCTION

Instruments for research purpose is a measurement device in the form of survey, test, questionnaire, etc which is developed, adopted or adapted by researchers to obtain data or responses aligned with research objectives. Instruments play an important role in collecting information or data to solve problems that arise or describe a phenomenon of society. However, a similar instrument administered in different settings may produce varied responses due to variations across geographical areas, culture, belief system, and personal factors of the respondents. Therefore, in order to ensure that instruments is useful, measurement instruments to be used in research must be valid and reliable [1].

Validity refers to the accuracy of an instrument to measure what should be measured. The validity of an instrument refers to whether appropriate inferences can be derived from the scores for a specific group of people with specific purpose. The usual validity to be considered in a research instrument is content validity, construct validity and criterion-related validity. Content validity can be obtained from the experts of relevant field [2, 3]. They assess in terms of content by taking into account the location of the study, the

samples used in the study as well as the objectives of the study. Construct validity can be obtained in several ways such as factor analysis, correlation and item response theory model, which commonly used Rasch Measurement Model (RMM) [4, 5]. Criterion-related validity measures the correlation coefficient between two measures. In this study, RMM was used to test the construct validity of the motivation scale.

Reliability of the instrument refers to the extent to which the instrument can measure what should be measured consistently. To determine the reliability of instrument, Cronbach's alpha is the most common used by the researchers. However, Rasch model in the interpretation of person reliability actually also equivalent to Cronbach's alpha coefficient [6]. Therefore, in this study, researcher reported the reliability and validity of motivation scale using Rasch Model.

The objective of this study was to examine how RMM was used to determine the validity and reliability of the data obtained from motivation scale. The motivation scale developed was initially used in the research for measuring student's motivation towards English as second language in Malaysia context. It was adapted in this research to investigate the motivation of students from national schools in Malaysia who study Chinese language as second language. In order to obtain the good result of the instrument, researchers have listed the following objectives: a) to analyse the validity of instrument based on value of point measure correlation value (PMC), infit and outfit mean square (MNSQ), infit and outfit standardised z value (ZSTD), unidimensionality and person and item separation index. b) To analyse the reliability of the motivation scale (IIM) using Cronbach's alpha coefficient.

2. RESEARCH METHOD

Researcher conducted the pilot study consisted of 29 students who have enrolled in Chinese language as second language (CLSL) class in a national school. The samples answered the questionnaires with assistance from their Chinese language teachers.

2.1. Instrument

Instrument of motivation was modified by researcher based on the 2 orientation of motivation which is integrative motivation and instrumental motivation (IIM) by Syed Mohd Kamil [7]. It consists of 24 items and divided to 12 items respectively. All items were measured using a five point Likert scale ranging from 1=strongly disagree to 5=strongly agree.

2.2. Data Analysis

In order to determine the reliability and validity of the instruments, researchers have conducted the item analysis using WINSTEPS version 3.72.3 based on Item Response Theory using Rasch Measurement Model to examine the items in the questionnaire in this study. The adapted instruments for the pilot study were computed to determine the Cronbach's alpha by using IBM SPSS Statistics version 23. The internal consistency results were accepted in most of the social science research if the scales were above 0.70, it shown consistently reliable [8]. To identified the data fit the model, PMC, infit and outfit MNSQ and infit and outfit ZSTD were being examined and compared sequentially within the range of acceptable fit indices.

PMC calculates the index of the item discrimination where the item with greater value might be too good to other items. The acceptable region for PMC is between 0.28 and 0.86 [9,10]. Infit and outfit of MNSQ and ZSTD are used to identify the relationship between the item difficulty and individual ability level. The acceptable region for both infit and outfit MNSQ is between 0.50 and 1.20 [11]. Infit and outfit ZSTD will only be accepted within ± 2.0 [12,13]. Only when all these three above mentioned controls meet, otherwise the item will be consider as misfit.

A person separation index of 1.5 or a person reliability coefficient of 0.7 present an acceptable level of separation and is considered the minimum required to divide the sample into two distinct strata, while a person separation index of 2.0 and a person reliability of 0.8 represent a good level of separation and are considered the minimum preferable values. Item separation index and item reliability are interpreted using the same criteria [6].

According to Souza et. al [14], Principal Component analysis (PCA) of the residuals used to examine unidimensionality. In unidimensional measures, it is expected that the observed variance explained by the measures roughly matches the expected variance in the model. The "first contrast" is the component that explains the largest possible amount of variance in the residuals. If the unexplained variance found in the first contrast is up to 2.0 Eigenvalue, the biggest possible secondary dimension has the strength of less than 2 items. Raw variance explained by measures minimum acceptable value is 40%, where as unexplained variance in the first factor should be not more than 15% [15,16].

3. RESULTS AND DISCUSSION

Table 1 presented the result of Cronbach's alpha using IBM SPSS Statistics version 23. Cronbach's alpha is used to measure the reliability of a set of scale. The values of reliability coefficient range from 0 to 1.0. The recommended minimum α coefficient between 0.5 and 0.8. The 15 items - IIM scale showed the reliability coefficient $\alpha = 0.87$ is acceptable. Based on this value it is indicated the instrument have high reliability [17].

Table 1: Reliability Statistics on 15 items-IIM

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.872	.880	15

3.1 Rasch Measurement Model(RMM) Item Analysis on Integrative Motivation and Instrumental Motivation (IIM)

Table 2 shows the item fit analysis for 24 item IIM. Any item which do not in the range of PMC ($0.28 < x < 0.86$), infit or outfit MNSQ ($0.5 < x < 1.5$) or ZSTD ($> \pm 2$), researcher eliminate the unfit items from the 24 items -IIM. The evaluation of "fit" items to the Rasch Model assured the quality of a measurement instrument [18].

Table 2. Item measure analysis for 24 items -IIM

ITEM	INFIT		OUTFIT		PT-MEASURE	
	MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.
23	1.17	0.7	1.65	2.0	0.46	0.67
18	1.56	1.9	1.45	1.4	0.54	0.66
5	1.48	1.7	1.35	1.1	0.55	0.66
17	1.29	1.1	1.19	0.7	0.63	0.66
20	1.26	1.0	1.10	0.4	0.71	0.64
15	1.19	0.8	1.11	0.4	0.42	0.62
2	1.18	0.8	1.04	0.2	0.67	0.66
9	1.10	0.5	1.11	0.5	0.68	0.65
12	1.11	0.5	1.05	0.3	0.60	0.61
13	1.11	0.5	1.05	0.3	0.60	0.61
19	1.09	0.4	1.02	0.2	0.69	0.65
21	1.07	0.3	0.97	0.0	0.75	0.66
10	1.05	0.3	0.94	-0.1	0.58	0.63
16	1.03	0.2	1.01	0.1	0.59	0.62
8	0.96	-0.1	0.83	-0.4	0.58	0.61
22	0.88	-0.3	0.80	-0.4	0.63	0.56
11	0.73	-1.0	0.77	-0.6	0.65	0.61
14	0.77	-0.9	0.68	-1.0	0.76	0.64
1	0.72	-1.1	0.76	-0.6	0.68	0.61
4	0.72	-1.1	0.76	-0.6	0.68	0.61
6	0.68	-1.2	0.63	-1.0	0.70	0.58
3	0.58	-1.8	0.63	-1.2	0.71	0.64
7	0.58	-1.8	0.63	-1.2	0.71	0.64
24	0.56	-1.9	0.60	-1.5	0.70	0.66
Mean	0.99	0.0	0.96	0.0		
S.D.	0.27	1.1	0.27	0.8		

According to Boone [19], it is best to remove the item or replaced with new item if an item was found clearly unfit. Table 3 provides a summary of the result after the elimination of the unfit items. There were remain 15 items fall in the acceptable range with PMC from 0.71 to 0.86, the infit and outfit of MNSQ were between 0.53 to 1.43 and ZSTD ranged from -1.1 to 1.2.

Table. Item measure analysis for 15 items-IIM

ITEM	INFIT		OUTFIT		PT-MEASURE	
	MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.
20	1.43	1.2	1.31	0.9	0.80	0.79
12	1.29	0.8	1.18	0.5	0.71	0.74
13	1.29	0.8	1.18	0.5	0.71	0.74
14	1.15	0.5	1.20	0.6	0.78	0.80
3	1.06	0.3	1.03	0.2	0.77	0.78
7	1.06	0.3	1.03	0.2	0.77	0.78
2	1.03	0.2	1.00	0.1	0.84	0.85
19	1.02	0.2	0.98	0.1	0.81	0.81
24	0.94	0.0	1.01	0.2	0.86	0.80
21	0.93	-0.1	0.90	-0.2	0.86	0.83
11	0.90	-0.1	0.66	-0.5	0.76	0.72
1	0.72	-0.5	0.52	-0.8	0.76	0.70
4	0.72	-0.5	0.52	-0.8	0.76	0.70
22	0.70	-0.6	0.66	-0.3	0.73	0.67
6	0.53	-1.1	0.54	-0.7	0.77	0.70
Mean	0.98	0.1	0.92	0.0		
S.D.	0.24	0.6	0.26	0.5		

Table 4 shows the PCA analysis of the residuals in Rasch showed the raw variance explained by measures was at 66.4% , with the unexplained variance in the first factor was at 13.6%. The results were fulfilled the requirement of a satisfactory dimensionality which showed more than 40% in raw variance explained by measures and less than 15% in unexplained variance in 1st contrast [20]. Thus, the dimensionality test demonstrated that the 15 items IIM was unidimensional and it was a good instrument in terms of its construct validity.

Table 4. Standard residual variance (in Eigenvalue units) for 15 items-IIM

		Empirical	Modeled
Total raw variance in observations =	44.6	100.0%	100.0%
Raw variance explained by measures =	29.6	66.4%	66.7%
Raw variance explained by persons =	18.0	40.3%	40.5%
Raw variance explained by items =	11.6	26.0%	26.2%
Raw unexplained variance (total) =	15.0	33.6%	100.0%
Unexplained variance in 1st contrast =	6.1	13.6%	40.5%
Unexplained variance in 2nd contrast =	2.3	5.1%	15.0%
Unexplained variance in 3rd contrast =	1.7	3.9%	11.5%
Unexplained variance in 4th contrast =	1.6	3.5%	10.4%
Unexplained variance in 5th contrast =	1.2	2.7%	8.1%

Table 5 shows the fit statistics that the person separation index, 2.79 (good) and item separation index, 1.81 (acceptable) and person reliability and item reliability were at 0.89 and 0.77 respectively. The finding was in lined with recommendations by Linarce [6] and supported by the researches [21,22]. This results indicated that the new instrument, 15 item-IIM was a reliable and valid instrument to measure the motivation in the context of national school students.

Table 5. Fit statistics for 15 items-IIM

	Separation	Reliability
Person	2.79	0.89
Item	1.81	0.77

The outcome showed that the new instrument of IIM suitable used to measure the motivation of students learning CLSL in national school. There are other analyses such as factor analysis; item-total correlations and item discrimination power are calculated to analyze the validity of scale [23]. However, Rasch Models is widely used by researchers because the model is able to define the constructs of valid items and provide a clear definition of the measurable constructs that are consistent with theoretical expectations [24].

4. CONCLUSION

This study aimed to examine the validity and reliability of the motivation scale, IIM using Rasch Model. The findings of the result indicated that the objectives of the study were fulfilled. Drawing from the current study, it is suggested that to assure the validity and reliability of the research instrument, Rasch model analysis is highly recommended. In conclusion, the analysis proven the appropriateness of the IIM instrument for the measurement of students' motivation toward CLSL in national schools in Malaysia.

ACKNOWLEDGEMENTS

We would like to thank the school for providing me an opportunity to do my reseach in the school. We would also like to show my gratitude to all those students who had participated in answering the questionnaires. In addition, thank you to all those who are directly and indirectly giving us all support and guidance which made me complete the writing duly.

REFERENCES

- [1] Kimberlin CL, Winterstein AG. "Validity and reliability of measurement instruments used in research". *Am J Health Syst Pharm.* 1:65(23):2276-84; 2008.
- [2] Yaghmale, F., "Content validity and its estimation". *Journal of Medical Education.* 3(1); 2003.
- [3] Vahid Zamanzadeh, Akram Ghahramanian, Maryam Rassouli, Abbas Abbaszadeh, Hamid Alavi-Majd, Ali-Reza Nikanfar., "Design and Implementation Content Validity Study: Development of an instrument for measuring Patient-Centered Communication". *J Caring Sci.* 4(2): 165–178; 2015. Published online 2015 Jun 1. doi: 10.15171/jcs.2015.017
- [4] Wright, B., "Equitable Test equating". *Rasch Measurement Transctoins,* 7(2), 298-308; 1993.
- [5] Alatawi, Abdullah Jabr E. "Using Rasch to Investigate Messick's Validity of a Multiple-Choice Test". *Imperial Journal of Interdisciplinary Research (IJIR)* Vol-3, Issue-4, 2017.
- [6] Linacre, JM., *Winsteps Rasch Measurement Computer Program User's Guide.* Beaverton: Winsteps.com.2012
- [7] Syed Mohd Kamil Syed Alias. *Secondary School Students' Motivation and Attitude towards English* (Unpublished master dissertation), Universiti Teknologi Malaysia, Johor, Malaysia. 2012
- [8] Gorsuch, R. L. Factor analysis. Hillsdale: Lawrence Erlbaum Associates. 1983.
- [9] Nurhazirah, H., Nor Irnoni, M.L., Wan Edura, W.R. and M. Saidfudin, M., Personality Traits, Work-Family Conflict and Job Satisfactin: Items Validity using Rasch Measurement Approach. *Procedia-Social and Behavioural Science,* 2012, 65, 1013-1019. In Teng, K. Y. (2016), *Contributions of Inclusive Educational Programme, Teachers' Attitudes and Efficacy on Social Skills among Students with Autism Spectrum Disorder.* Doctor Philosophy, Universiti Teknologi Malaysia, Skudai.
- [10] Teng, K.Y. *Contributions of Inclusive Educational Programme, Teachers' Attitudes and Efficacy on Social Skills among Students with Autism Spectrum Disorder.* (Unpublished doctoral thesis), Universiti Teknologi Malaysia, Johor, Malaysia. 2016.
- [11] Sheu C, Chen C, Su Y, Wang W. "Using SAS PROC NLMIXED to fit item response theory models". *Behav Res Methods* 37:202–18; 2005.
- [12] Wright, B.D. and Linacre, J. M., "Understanding Rasch Measurement: Construction of Measures from Many-Faced Data." *Journal of Applied Measurement,* 3(4), 486-512, 2002.
- [13] Schumacker RE. *Rasch measurement: the dichotomous model.* In: Smith EV, Smith RM, eds. *Introduction to Rasch Measurement: Theory, Models and Applications.* Maple Grove: JAM Press, 226–53; 2004.
- [14] Souzal, M. A. P., Coster, W. J. Mancini, M. C., Dutra, F. C. M. S., Kramer, J. Sampaio, R. F, "Rasch analysis of the participation scale (P-scale): usefulness of the P-scale to a rehabilitation services network". *BMC Public Health* (2017) 17:934 DOI 10.1186/s12889-017-4945-9. 2017.
- [15] Fisher, W., "Rating Scale Instrument Quality Criteria". *Rasch Measurement Transaction,* 21, 1095-1098. 2007.
- [16] Conrad, K., Dennis, M. and Funk, R. *Validation of the Substance Problem Scale (SPS) to the Rasch Measurement Model, GAIN Methods Report.* (1.1 ed.) Chicago, I.L.: Chestnut Health Systems. 2011
- [17] Mehmet, K. "Book reading motivation scale: Reliability and validity study". *Educational Research and Reviews.* Vol. 10(3). pp. 300-307; 2015.
- [18] Boone. WJ, Staver. JR and Yale. MS., "Rasch Analysis in the Human Sciences.Dordrecht", Netherlands: Springer; 2014.
- [19] Boone. WJ., "Rasch Analysis for Instrument Development: Why, When, and How?". *CBE Life Sci Educ.* Winter; 15(4): rm4; 2016.
- [20] Samah Ali Mohsen Mofreh, Mohammed Najib Abdul Ghafar, Abdul Hafiz Hj Omar, Monsurat Mosaku1 & Amar Ma'ruf. "Psychometric Properties on Lecturers' Beliefs on Teaching Function: Rasch Model Analysis". *International Education Studies;* Vol. 7, No. 11; 2014.
- [21] Norasmah Othman, Suria Mohd Salleh, Haliza Hussein and Haryaty Ab. Wahid. "Assessing Construct Validity and Reliability of Competitiveness Scale using Rasch Model Approach". *The 2014 WEI International Academic Conference Proceedings.* Bali, Indonesia. 2014.
- [22] Lerdal, A., Opheim, R., Gay, CL., Moum, B., Fagermoen, MS., Kottorp, A., "Psychometric limitations of the 13-item sense of coherence scale assessed by Rasch Analysis". *BMC Psychology.* June 2017.

- [23] Yesil, R., "Validity and Reliability Analysis for the Teacher Determination for Democratic Behavior ScaleKuram ve Uygulamada Eğitim Bilimleri" /*Educational Sciences: Theory & Practice* 10 (4) • autumn 2010 • 2683-2692.2010.
- [24] Mimi, M.M., Nor Lisa, S., Lai, C.S. and Kahirol, M.S., "Measuring the Validity and Reliability of Research Instruments". *4th World Congress on Technical and Vocational Education and Training (WoCTVET)*, 5th–6th November 2014, Malaysia 2014.

BIOGRAPHIES OF AUTHORS



Sar Ee Ng is a lecturer attached to the Language Department, Institut Pendidikan Guru Kampus Temenggong Ibrahim, Johor, Malaysia. She is experience in the field of teaching Chinese language as first and second language.



Professor Dr. Yeo Kee Jiar is a lecturer attached to the School of Education, Universiti Teknologi Malaysia, Johor, Malaysia. With her vast experience and involvement in the field of education, her expertise and interest covers a wide scope which includes educational psychology, early childhood education, special education and Malay language. Among her notable involvement includes a few national level projects for research and consultation, as Munsyi Dewan under Dewan Bahasa dan Pustaka (Malay for The Institute of Language and Literature), publications in indexed journals as well as books, book chapters, and monographs. She has been a speaker and facilitator in many workshops and conferences centred on her field of specialization and interest.



Associate Professor Dr. Azlina binti Mohd Kosnin has vast experience in teaching, research and administrative work. Currently she holds the position as the Chair of the School of Educaiton, Universiti Teknologi Malaysia, Johor, Malaysia. Her expertise and interest covers the area of Educational Psychology, measurement and testing, and early childhood education. She has been leading a few national level research projects and involved in consultative work related to her field of specialization and have notable publications in both local and international journals.