

Content Reliability Measurement of Holistic Approach Training Module

Muhamad Afzamiman Aripin¹, Sarimah Ismail¹,
Rohana², Ros Eliana Ahmad Zuki³

¹Department of Technical & Engineering,
Faculty of Education, Universiti
Teknologi Malaysia, Malaysia

² Centre of Modern Language & Human Sciences,
Universiti Malaysia Pahang, Malaysia

³ Department of Technical & Vocational Education,
Institut Pendidikan Guru,
Kampus Temenggung Ibrahim, Malaysia
Email Address: afzamiman@gmail.com

March 27, 2018

Abstract

As part of correctional programme components, the significant role of vocational instructor in prison is not only about delivering skills, but also to improve inmates attitude and perspective about life. Therefore, the Holistic Approach in Correctional Education (HACE) Training Module was developed, to provide awareness and skills in training inmates as a whole among prisons vocational instructors. In addition, this module focused on the integration of God Consciousness element as a compass of life; internally and externally without disturbing the classic objective in vocational training. The development of HACE Module comprised a total of fourteen activities, structured based on modified All Quadrant All Level (AQAL) Model. A pilot study was conducted on a group of 42 prisons vocational

instructors from the Middle Zone of Malaysian Peninsular under Pride. The collected data was analyzed using Rasch Measurement Modelling with Bond & Fox Steps Software to obtain the reliability of participants and the items in this research. The content reliability of the module achieved the alpha Cronbach value of 0.98, person reliability value of 0.94, which was quite high and item reliability value of 0.68, which was at the medium level. Most of the items met the range of infit/outfit Mean Square and were in the range of fit / outfit z-standard. Hence, it was concluded that, the HACE Training Module achieved its objective of development and was thus perceived appropriate to be implemented.

Key Words:Holistic Approach, Module Development, Prison Instructors, Rasch Model, Reliability Analysis

1 INTRODUCTION

Vocational education and training is one of the important elements in rehabilitative process under Malaysian Prison Department (Pride). Similar with the other vocational training programmes, there are hundreds of instructors working behind the wall to facilitate the vocational learning process. It is important to realize that, the prison context is unique, where the working culture is different and not experienced elsewhere⁵. As a vocational instructor in the prison, one should be able to create a positive teaching and learning environment, and to equip inmates, not only with skills-but also improving their personal attitudes and perspectives about life. Hence, a holistic approach in vocational education beyond employability was needed, thus leading the authors to develop a suitable module known as Holistic Approach in Correctional Education (HACE) Training Module for prison vocational instructors. Development of a new training module requires a detailed process to ensure that, the intended objective is achieved ^{11,14}.

Module is an instructional package dealing with a single conceptual unit of the subject matter¹³. There are two types of modules; facilitated module and non-facilitated module²⁰. Development of adult modules, especially in career development, was divided into two other sub categories, namely, on-job training and external

training. Therefore, the nature of HACE Training Module was facilitated module in the form of external training¹⁷. In this study, the module development process was based on modified Rusells Model as shown in figure 1. Since the previous work by Rusell never mentioned about reliability test as part of the module development process, the required modification was done, based on the suggestion made by Sidek & Jamaludin (2008) who considered it a very unusual phenomenon for a validity test to be conducted without reliability measurements¹⁴.

Therefore, this paper introduced the basic concept of content reliability in module development, basic application of Rasch Model in doing the analysis and the discussion on its application.

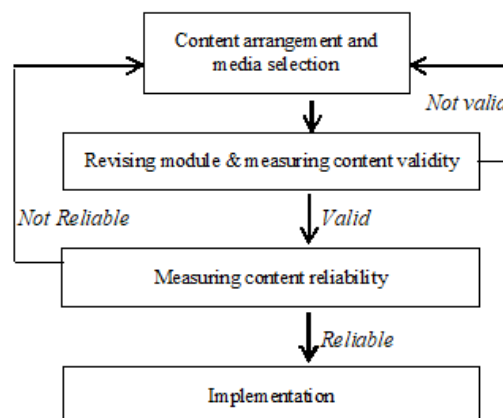


Fig.1. Module development process

2 LITERATURE REVIEW

The literature review gave an overview of the HACE Training Module, Module Reliability & Rasch Modelling.

2.1 Overview on HACE Training Module

The HACE Training Module was developed by the researchers with the collaboration of Malaysian Prison Department to enhance the

teaching quality of vocational instructors. The purpose of the module was to set a holistic way of thinking which would allow the vocational instructors to relate their teaching with the correctional objective as a whole. By implementing modified AQAL (All Quadrant All Level) Model by Gehring and Puffer (2006) as a foundation, the module was elaborated and adjusted to suit the Malaysian context. The model emphasized on the manner, the internal factors and spirituality of an instructor could affect external elements in a system, such as, the objectives of teaching and learning, the system sociology, and making the professional thinking a culture. The content of the module was explored from the works of Tessa West (1995) and Lynne Rogers et al., (2014) on the holistic character of prison instructors, Landrum & Gardner (2012) on holistic approach in firms, Haigh (2013) on holistic pedagogical research, and local scholars, such as, Syed Muhammad Naquib Al-Attas (1990), Rosnani (2004), Hamdan & Najmuddin (2013), Ros Eliana (2014) on the holistic education concepts in Malaysia.

2.2 Module Reliability and Rasch Modeling

Educational research was always concerned with the connection of items level and the performance of groups involved in the research¹⁸. Sometimes, the level of performance might be much lower in one group of persons than the other, rendering the same items not being suitable for everyone. Seemingly, there is a relationship between a persons ability and an items difficulty, even though, other factors may influence the results. In most cases, it is expected that, the same items can be used to evaluate impact of treatment on people of the same group. Application of the same items required the researchers to evaluate the reliability of items of the instruments. In the module development process, the instrument for the data collection was the module itself¹⁴. Therefore, every activity was transformed into a question to measure the possibilities of its implementation by vocational instructors. Reliability referred to the consistency of the scores obtained. Reliability estimates are usually expressed as another application of the correlational coefficient known as reliability coefficient. Reliability coefficient must range from 0.00 to 1.006. In a module development, content reliability should be done after content validity. Two ways of testing content

reliability for a module, namely, by activities and by objectives¹⁴. However, the measurement through activities is considered as the most accurate way in determining the reliability coefficient.

One of the ways of measuring reliability coefficient is by finding the alpha Cronbach value using Rasch analysis. Rasch analysis was one of the modern psychometric techniques that formed part of Item Response Theory (IRT)⁹. In fact, it is one of the most widely used IRT models in various IRT applications. One of the properties of Rasch Model is known as invariance. It allows researchers to model all person ability and item difficulty in a continuum^{10,18}. The person reliability index indicates the replicability of a person ordering we could expect if this sample of persons were given another parallel set of items measuring the same construct. The item reliability index indicates the replicability of item placements along the pathway if these same items were given to another same-sized sample of persons who behaved on the same way¹. With these definitions, reliability analysis in the module development process was a procedure to determine, if the designed module activities were suitable with the ability of the sample. This would help to give us confidence in our estimation ability when implementing the same module to the other groups of vocational instructors. The value of alpha Cronbach would represent one of these assumptions; $\alpha < 0.60$ =poor, $0.60 < \alpha < 0.7$ =adequate, $0.70 < \alpha < 0.8$ =good, $0.8 < \alpha < 0.90$ =very good, ≥ 0.9 =perhaps too good¹⁸. However, the assumptions of the Rasch Model were violated, if there were other factors taken into consideration, other than person ability and items difficulty¹⁹. These factors might include guessing possibility (guessing can occur especially in a difficult multi-choices questionnaire), item dependency (probability of success of an activity depends on the response of other activities), differential item functions (different groups of participants respond to the items for an activity in different ways), and other traits (a person may be good in few activities and not in others). The extent to which the Rasch Model assumptions were violated could be measured using fit statistics¹⁹. It was a form of quality control for items and persons¹. The indication of data being fit and misfit the t (z- standardized value) falls outside -2.0 and +2.0 with sample sizes between 30 to 3001,¹⁹. The assumptions made are that, locations on the right ($> +2.0$) are too erratic to be useful and the locations on the left ($< -$

2.0) are too good to be true. If fit mean square statistics is used, the expectations of the statistic is one, that is, when the data fits the model, the distribution of items fit-mean square values should be scattered close to one. Recommendations on the range of acceptable values for Rasch Model as in Table 12. However the fit statistics should not be used blindly to reject items¹⁹.

Table.1. Scale of Quality for Rasch Model Measurement

Respondent/Item	Scale
Acceptable Point Measure Correlation	$0.4 < x < 0.8$
In-fit/Out-fit Mean Square	$0.5 < x < 1.5$
In-fit/Out-fit z -standardized value	$-2.0 < x < 2.0$

3 OBJECTIVE

The aim of the study was to measure content reliability of HACE Training Module for prison vocational instructors using Rasch Modelling.

4 LOCATION

The administration of staff training by Pride for correctional and detaining institutions is divided into two zones under Malaysian Peninsular region; Middle Zone and East Zone. The pilot study was conducted under the Middle Zone which consists of institutions from Johore, Malacca, Negeri Sembilan and Selangor and took place at Institut Koresional Malaysia, Melaka (IKOM).

5 METHODOLOGY

A quantitative method was used as the main methodology in this study. There were a total of 42 participants or respondents for the content reliability assessment, selected among vocational instructors from 14 different institutions in the Middle Zone of Malaysian Peninsular under Pride. This group was chosen for the pilot study based on two factors as follows;

a) They were vocational instructors appointed as Assistance Officers of Vocational Training (PPLV) and Trade Warders in prison.

b) Involvement of prisons vocational instructors was made based on zone selection to minimize the threat of internal validity in the real implementation. In addition, vocational instructors from one zone are having occasional interactions with their peers from other zones. Therefore, the possibilities of information exchange could be minimized.

Content reliability could be measured through activity-based instrument^{11,14}. In this research, a total of 14 activities within HACE Training Module were transformed into a survey instrument. The instrument consisted of 69 items and was validated by experts prior to the commencement of the study.

Table.2. Abbreviation of items on person-item map

Activity	Sub Module	Abbreviation
1	Ice-Breaking	P1-P6
2	Rationality	BR1-BR4
3		MS1-MS3
4	Knowing purposes	SV1-SV5
5		KV1-KV3
6		DS1-DS6
7		PI1-PI4
8	Journey of Hoopoe	KA1-KA7
9		KB1-KB5
10	Application	OB1-OB8
11		LA1-LA6
12		LB1-LB3
13	Culturing	TK1-TK4
14		HS1-HS5

6 RESULTS

6.1 Module Reliability

Using Rasch measurement analysis with Bond & Fox Step, data collected from the 42 questionnaires was analyzed for the content reliability and the results obtained were as shown in Table 3 below;

Table.3. Summary of Person Reliability

	RAW SCORE	COUNT	MKASURK	MOOKL ERROR	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MKAN	301.5	69.0	4.20	.40				
S.D.	26.3	.0	2.36	.41				
MAX.	345.0	69.0	10.06	1.83				
MIN.	245.0	69.0	.33	-.21				
RRAL RMSK	.58	ADJ. SD	2.28	SEPARATION	3.96	Person	RELIABILITY	.94
MOOKL RMSK	.57	ADJ. SD	2.29	SEPARATION	4.03	Person	RELIABILITY	.94
S.E. OF Person MKAN = .37								

Person RAW SCORE-TO-MEASURE CORRELATION = .95
 CRONBACH ALPHA (KR-20) Person RAM SCORE RELIABILITY = .98

Table.4. Summary of Item Reliability

SUMMARY OF 69 MEASURED (NON-SKIPPED) Items								
	RAW SCORE	COUNT	MKASURK	MOOKL ERROR	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MKAN	168.5	39.0	.00	-.36	.99	-.1	1.02	-.1
S.D.	5.3	.0	.67	.01	.37	1.4	.55	1.5
MAX.	184.0	39.0	1.99	.41	2.32	4.0	3.38	5.0
MIN.	151.0	39.0	-2.16	-.30	.43	-2.8	.38	-2.6
RRAL RMSK	.38	ADJ. SD	.56	SEPARATION	1.47	Item	RELIABILITY	.68
MOOKL RMSK	.36	ADJ. SD	.57	SEPARATION	1.60	Item	RELIABILITY	.72
S.E. OF Item MKAN = .08								
LACKING RESPONSES: 2 Items								
UNKNOWN USUAL=1.000								
Item RAW SCORE-TO-MKASURK CORRELATION = -1.00								
2691 DATA POINTS. APPROXIMATE LOG-LIKELIHOOD CHI-SQUARE: 3256.59								

The content reliability of the module achieved the alpha Cronbach value of 0.98, with the person reliability value of 0.94 and item reliability of 0.68. The person reliability value was very high, while the item reliability value was moderate. 0.68 was an adequate reliability coefficient value and could still thus be accepted¹⁸. The maximum item was located at +1.99 logits, while the maximum for person was at +10.06 logits. The results showed that, most of the items were distributed under the moderate and lower levels. Most of the respondents achieved higher than +1.99 logits, indicating that they were in the group of those with a high ability. The minimum item was located at -2.16 logits, while the minimum for person was at +0.33 logits, above the level of items difficulties.

6.2 Item Fit

Generally, there were seven mis-fit items detected using Rasch Model. Table 5 shows that items 5, 37, 1, 13, 26, 35 and 32 were misfits in many ways. Most of the items were having z-std value >+2, while for item number 32, the MNSQ value was greater than 1.5.

Table.5. Item mis-fit

ENTRY	RAW		MODEL	INFIT	DIFFFIT	IPWRA	ROOCT	WATC	
NUMBER	SCORE	COUNT	MEASURE	S.E.	(MNSQ)	ZSTD(MNSQ)	ZSTD(CORN.)	ORIG	ROOFL
1	5	151	39	1.99	-.3012.37	3.413.08	5.01A	.551	56.4
1	37	175	39	-.56	-.3811.19	-.913.06	3.71B	-.431	74.4
1	1	170	39	-.17	-.3611.71	2.412.45	3.21C	-.391	59.0
1	13	172	39	-.43	-.3812.20	4.012.12	2.51D	-.431	61.5
1	26	165	39	-.46	-.3511.50	1.812.09	2.91E	-.311	76.9
1	35	160	39	1.05	-.3411.85	2.412.03	2.81F	-.571	64.1
1	32	170	39	-.96	-.3711.20	-.911.88	1.81G	-.431	69.2
1	MEAN	168.5	39.0	.00	-.361.99	-.111.02	-.11	1	74.5
1	S.D.	5.3	.0	.87	.011.37	1.41.55	1.51	1	8.4

Before we proceed with the discussion on how we are going to treat the data, this paper will present the distribution of module reliability for persons and items as in figure 2.

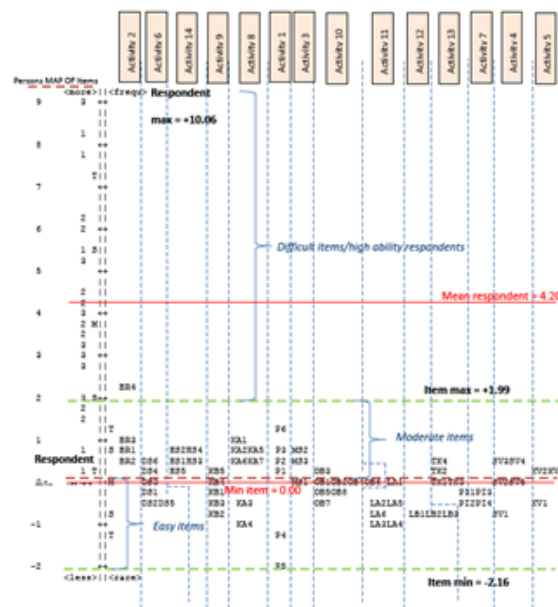


Fig.2. Distribution of Module Reliability

7 DISCUSSION

Is it really matter for the items to be fitted into the model? In the first analysis, there were 7 items that were considered as being miss-fitted. Unfortunately, these items could not simply be removed, because in the module development, each item represented the learning outcomes or ability of the respondents in performing the tasks. There were many assumptions that we could suggest on the reasons it did not fit the Model, one of them being absolutely

different ability of individual or group in performing certain tasks. For example, items 32 (KA1), 35 (KA4) and 37 (KA6) were representing group activities in a few checkpoints. The participants were not expected to perform successive accomplishments in every single task. The activities were purposely designed in such a way that, the best situation was when there was a failure and a successor so that, they could do a great reflection on activity 9. This was also one of the reasons most of the items in activity 9 were located under the easy category. Sometimes, it was not about the activity but the way it was conducted, time allocation and respondents readiness. For example, item number 5 (P5) was actually a verbal activity and it was the easiest item among all. However, in Rasch analysis, the item or activity was too erratic to be considered as useful, due to the z-standardized value higher than +2. In this situation, the possibility of eliminating the activity could be considered, but not before an assessment was done especially, on the mentioned factors that could affect the statistical result. Since the procedure of removing items was not that easy, the other option left was to improve the reliability measurement of the module by eliminating unreliable persons. However, in this research, eliminating the unreliable persons also did not give much impact on the results. For example, by eliminating persons numbered 31, 29, 19, 33, 32 and 24, the item reliability increased to 0.69, but the content reliability remained at 0.98 and the person reliability at 0.94. The removal of persons considered as really, really bad was not so bad after all, because there were only small changes caused (0.01 differences) after the second analysis. Therefore, the best consideration was to report all persons and items in the study. By referring to the distribution of HACE Training Module reliability, the samples in this study were mostly in the category of high ability, while the items were mostly in the continuum of moderate and easy levels. This was a good indicator in module development, because, it showed that, the module was possible to be implemented and it should be very functional. From all 69 items, item number 5 (BR4) was the only item located in the difficult category. BR4 was actually a statement for an item representing I feel ready to learn in activity 2. Only six respondents found that it was hard, or in this study, it meant I am not ready to learn. In a social science study, it is normal to have a minority or even a majority of people with different perceptions and

ideologies, especially at the very beginning of the programme. For example, in HACE training module, activity 2 was actually a mind setting sub-module. At least, in this study, 36 other participants were ready to learn and participate in the programme.

8 CONCLUSION

This article aimed to measure the content reliability of HACE Training Module for prisons vocational instructors. The reliability of persons and items were analyzed using Rasch Model to provide an in-depth understanding before the real implementation took place. Rasch Model analysis was a powerful tool in gaining confidence and in achieving progress scientifically in social science research¹⁹. Nevertheless, in doing the Rasch Model analysis, it must be borne in mind that, measurement, as precisely as in natural sciences, could not be expected to be done, due to the complexity of human behaviours. While there was a hesitation to make any substantial changes either on the activities or the objectives of the activities, the pilot study of HACE Training Module was shown to have demonstrated appropriate content reliability in the analysis. However, a good assessment for the effectiveness of the HACE Training Module was still needed.

ACKNOWLEDGMENT This research was funded by Knowledge Transfer Programme (KTP) grant under KPM with UTM research project vote no. R.K 130000.7806.4L512. It was conducted in collaboration with the Malaysian Prison Department

References

- [1] Bond T.G & Fox C.M (2007) Applying The Rasch Model. Fundamental Measurement in the Human Sciences (2nd edition). Routledge. New York
- [2] Fisher, W.P. (2007). Rating Scale Instrument Quality Criteria. Rasch Measurement Transactions. Vol.21(1), 1095.
- [3] Haigh, M. (2013). AQAL Integral: a holistic framework for pedagogic research. Journal of Geography in Higher Education. Vol. 37, No.2, pp 174-191.

- [4] Hamdan Muhammad & Muhammad Najmuddin Ulwan (2013). Kurikulum dari Langit. Kemilau Publika Sdn Bhd. Selangor
- [5] Hawley, Jo., Murphy, I. & Manuel Souto-Otero (2013). Prison Education and Training in Europe. Current State-of-Play and Challenges. European Commission.
- [6] Jack R. Fraenkel, Norman E. Wallen and Helen H. Hyun (2012). How to Design and Evaluate Research in Education (8th Edition). McGraw Hill. USA.
- [7] Landrum, N.E and Gardner, C.L (2012). An Integral Theory Perspective on the Firm. International Journal of Business Insights & Transformation. Volume 4, Special Issue 3, pp 74-79
- [8] Lynne Rogers, Margeret Simonot, Angela Nartey (2014). Prison Educators: Professionalism Against the Odds. Institute of Education University of London. UK
- [9] Nor Fadila Amin (2013). Pendidikan Reka Cipta Menerusi Sistem Pembelajaran Web Berasaskan Prinsip Andragogi dan Pembelajaran Arah Diri. Phd Thesis. UT
- [10] Rasch G. (1980). Probabilistic Models for Some Intelligence and Attainment Test. The University of Chicago Press. Chicago
- [11] Ros Eliana Ahmad Zuki (2014). Model dan Modul Pengajaran Integrasi Holistik Guru Pelatih Pendidikan Teknik dan Vokasional. PhD Thesis. UTM
- [12] Rosnani Hashim (2013). Educational Dualism in Malaysia. Implications for Theory and Practice (2nd ed.). The Other Press. Selangor
- [13] Rusell, J.D (1974). Modular Instruction: A Guide to the Design, Selection, Utilization and Evaluation of Modular Materials. New York Publishing Company. The United States
- [14] Sidek Mohd Noah & Jamaludin Ahmad (2008). Pembinaan Modul. Bagaimana Membina Modul Latihan dan Modul Akademik. Universiti Putra Malaysia. Serdang

- [15] Syed Muhammad Naquib al-Attas (1990). *The Nature of Man and the Psychological of the Human Soul. A brief outline and a framework for an Islamic Psychology and Epistimology.* International Institute of Islamic Thought and Civilization. IIUM
- [16] Tessa West (1995). *Basic Education in Prisons. Doing things differently: a holistic approach to prison education.* United Nations and UNESCO Institute of Education. USA
- [17] Thom Gehring and Margaret Puffer (2006). *Integral Correctional Education.* AQAL. *Journal of Integral Theory and Practice.* Volume 1.1, pp 183–212
- [18] Tolmie A., Daniel Muijs, Erica McAteer (2011). *Quantitative Methods in Educational and Social Research Using SPSS.* McGraw-Hill Education. United Kingdom.
- [19] Wu, M. & Adams, R. (2007). *Applying the Rasch model to psycho-social measurement: A practical approach.* Educational Measurement Solutions, Melbourne.
- [20] Zafir Mohd Makhbul & Fazilah Mohamad Hasun (2011). *Prinsip Pegurusan Sumber Manusia.* Leeds Publications. Selangor