VALIDITY OF VOCATIONAL PEDAGOGY CONSTRUCTS USING THE RASCH MEASUREMENT MODEL

Md Yunos, J., Ibrahim Mukhtar, M., Alias, M., Lee, M. F., Tee, T.K., Rubani, S.N.K., Hamid, H., Yunus, F.A.N., Sulaiman, J., & Sumarwati, S.

Faculty of Technical and Vocational Education,
Universiti Tun Hussein Onn Malaysia, Johor, Malaysia

Correspondence author email: jailani@uthm.edu.my

Received April 19th, 2017; Accepted November 19th , 2017

ABSTRACT

This study was conducted to generate empirical evidence on the validity and reliability of the items of a new vocational pedagogy instrument. Two stage processes were used to identify and validate constructs. The first stage involves the identification of constructs using exploratory factor analysis method and the second stage was construct validation stage using the Rasch measurement model. Construct validity was examined by analyzing the point-measure correlation index (PTMEA), infit and outfit MNSQ values; while the reliability was examined by analyzing item reliability index. The instrument was distributed to 183 teachers from Malaysian and Indonesian vocational institutions. EFA indicate the presence of construct from 2 sections (Learning Content and Pedagogical Decision). Further analysis using the Rasch measurement model indicates that the reliabilities were between 0.77 to 0.91 and 0.97 to 10.13 for the Learning Content and Pedagogy Decision constructs respectively. PTMEA values are positive for Learning Content constructs indicating that all items are able to distinguish between abilities of respondents but 3 items from Pedagogical Constructs shows there were not moving in line with other items to measure the constructs. However that item was retained with refinement. In conclusion, this study has established valid constructs for development of instrument for assessing vocational pedagogy.

Keywords: *Vocational pedagogy, Rasch model, construct validity.*

1. INTRODUCTION

The technical and vocational education and training (TVET) is globally accepted as having a key role in promoting both economic and socio-economic growth, increasing productivity, empowering citizens and alleviating poverty. Yet the quality of TVET in terms of learner outcomes and teaching inputs are of varying quality. In some countries this unhelpful inconsistency is being addressed through the use of accountability regimes to validate the quality of provision, while in others quality is improved through increased professionalization and training of the TVET workforce.

As developing nations with fast growing population, it is of utmost importance for Malaysia and Indonesia to provide a comprehensive and up-to-date TVET for their citizens. The success of TVET however, depends to a large extent on the ability of its trainers, instructors and teachers to apply the appropriate pedagogy or specifically vocational pedagogy in achieving TVET goals. It is therefore, important for TVET providers to ensure that vocational pedagogy related knowledge and skills are adequate for Malaysians and Indonesians to serve as the effective technical and vocational trainers, instructors and teachers that they require.

To date, it is not known as to what extent TVET teachers have the vocational pedagogy mastery to create appropriate learning environments that can equip students to compete in the work place. In contrast to general education, students in TVET institutions must learn the high-level technical skills that are expected for positions in their field as well as the transferable skills that will allow them to keep these positions or advance to better ones (Echternacht & Wen, 1997). However, teachers tend to stick to a small number of methods that they feel comfortable with (Lucas and Claxton, 2013) which may not be appropriate for TVET content that they supposed to be addressing. Furthermore, the effects of vocational pedagogies are under-researched and under theorized (Cedefop, 2015). Thus, more studies are needed on vocational pedagogy.

Vocational pedagogy is a relatively new term which is referred to by Lucas and Claxton (2013) "…as the science, art, and craft of teaching that prepares people for working lives. It is critically shaped by the decisions that are taken by teachers – both high-level strategies, and day to day 'in-the-moment' ones – and the values that inform all interactions with students". It does clear that TVET needs to be taught in the context of practical problem-solving and that high-quality TVET almost always involves a blend of methods. But how much it has the blend of method been implemented? To gauge the extent to which vocational pedagogy has been implemented, there's a need to establish a new instrument by using new terminology but still using old practices in vocational pedagogy to improve the status and quality of TVET which requires a paragon of virtue, knowledge and skill. To establish such an instrument, empirical evidence on the validity and reliability of drafted items are necessary.

1.1 Purpose of study

The purposes of this study are to explore the psychometric properties and to examine the validity and reliability of the newly developed vocational pedagogy instrument.

1.2 Objectives

(i) To identify the constructs of vocational pedagogy as practiced by teachers in Malaysia and Indonesia.
(ii) To validate the identified constructs of vocational pedagogy for Malaysia and Indonesia.
(iii) To validate an integrated Malaysian-Indonesian vocational pedagogy constructs.

2. RESEARCH METHODOLOGY

This study utilised both the qualitative and the quantitative research methods. The qualitative method involved in-depth interviews and document analysis while the quantitative method involved data collection through the newly developed instrument. However, this paper will only focus on the validation stage (quantitative data) of the research instrument development process which involves a survey on 183 TVET teachers from Malaysia and Indonesia. The polythomous data (Likert) were collected and analyzed using the Rasch Model with the aid of computer application software, WINSTEPS version 3.72.3.

This quantitative phase involves two stage processes to identify and validate the constructs. The first stage involves identification of constructs through literature review and supported by Exploratory Factor Analysis (EFA). The EFA indicated the presence of constructs from two sections (Learning Content and Pedagogical Decision). The validation process was taken further using the Rasch Measurement Model for polythomous data aided by Winstep software version 3.68. The construct validity was examined by analyzing the point-measure correlation index (PTMEA), infit and outfit MNSQ values; while instrument reliability was examined by analyzing item reliability index. A survey technique was used as the major method with the instrument on 183 teachers from TVET institutions in Malaysia and Indonesia.

3. FINDINGS AND DISCUSSIONS

Data were generated from responses on two sub-scales of the vocational pedagogy instrument, Learning Content and Pedagogical Decisions and subsequently analysed using the Rasch measurement model method. The item functionality inspection covered aspects of reliability and items suitability/fit.

3.1 Items and person reliability

The item and person reliability and separation indices show the extent to which the items are compatible (conform to fit) with the Rasch Measurement Model. Table 2 and table 3 show the summary of the item separation index and person separation index, the item reliability and person reliability. The results of the analysis show that items for the 10 Learning Content constructs have a range of reliability between 0.77-0.91 (refer to table 2) and items for the six Pedagogical Decision constructs have a range of reliability between 0.97-10.13 (table 3). The reliability values of more than 0.8 are acceptable, while values between 0.6 - 0.8 are less acceptable and values less than 0.6 are not acceptable (Bond & Fox, 2007). Table 2 shows that there is one construct from the Learning Content that has a low item reliability index which is

Set (0.77). Although this value does not conform to high reliability index, it is adequate and is of acceptable level (Pallant, 2011). The items reliability index can be further enhanced if the misfit items are given special attention. Meanwhile table 3 shows that all Pedagogical Decisions constructs have the acceptable reliability values.

Person separation is used to classify people. Low person separation (< 2, person reliability < 0.8) with a relevant person sample implies that the instrument may not be sensitive enough to distinguish between high and low performers. More items may be needed. Item separation is used to verify the item hierarchy. Low item separation (< 3 = high, medium, low item difficulties, item reliability < 0.9) implies that the person sample is not large enough to confirm the item difficulty hierarchy (construct validity) of the instrument. Separation Index is the separation of items and person. The items and person separation value which is more than 2 is good (Fox & Jones, 1998; Linacre, 2005; Bond & Fox, 2007). Table 1 shows the rating scale of instrument quality criteria

Table 1: Rating scale instrument quality criteria

Criteria	Poor	Fair	Good	Very Good	Excellent
Person & Item Reliability	<0.67	0.67-0.8	0.81-0.9	0.91-0.94	>0.94

The result in table 2 showed the item separation index to be between the values of 1.85 to 3.23. Statistically speaking, these items can be divided into 1 strata or levels of agreement. Table 2 shows there are 2 items separation index for the 10 constructs in Learning Content have below than 2.0, while the items separation index for the 6 constructs in Pedagogical Decisions (Table 3) shows the separation index between 0.89 to 5.46 and these items can be divided into 1 to 5 strata or levels of agreement. Separation item value which is less than 2 is less accepted and this suggests that real difference related to the ability of respondents is hard to distinguish for the items. The higher the value of the separation indexes of the items, the better the instrument because the items are separated by levels of varying difficulty. The separation index will increase if the reliability of items is increased and misfit items are detected and removed from the analysis. This indices show that the person difference or separation can measure the ability of persons using the measuring instrument (Wright & Master 1982; Bond & Fox, 2007).

Table 2: The reliability of Learning Content constructs

No	Construct	Total Item	Item Reliability	
			Item	Separation
1	Set	3	0.77	1.85
2	Perception	6	0.80	1.99
3	Organization	6	0.81	2.09
4	Responding	18	0.82	2.15
5	Complex Over Response	6	0.83	2.20
6	Understanding	6	0.84	2.26
7	Application	9	0.85	2.37
8	Characterization By Value Set	9	0.87	2.63
9	Knowledge	9	0.89	2.84
10	Guided Response	12	0.91	3.23
	TOTAL	84		

Table 3: The reliability of Pedagogical Decision construct

No	Construct	Total Item	Item Reliability	
			Item	Separation
1	Nature of Activities	11	0.97	5.46
2	Approach to Task	2	2.76	0.89
3	Role of Learner	2	3.53	0.93
4	Role of Teacher	13	3.56	0.93
5	Visibility of Process	2	8.04	0.98
6	Teaching and Learning Activities	3	10.13	0.99
	TOTAL	33		

3.2 Polarity of items that measure the constructs

This analysis was conducted to determine the same item is generated that measure the construct to be measured. If the point-measure correlation (PTMEA Corr) value is high, it shows that the items are able to distinguish between respondents' ability (Bond & Fox, 2007). If the value of PTMEA Correlation is lower than 0.30, it means that the items do not fulfill the criteria set.

3.2.1 Polarity of items that measure the constructs of Learning Contents

PTMEA correlation value is to test whether all items are moving in one direction with the construct. Table 4 shows the summary of the 10 items for Learning Content indicating that network response to items or the respondent is inconsistent with the constructs. In other words, the items are not moving in line with other items to measure the constructs but they will be retained after revisions and refinements because the items fulfil the requirement of the Learning Content.

Table 4: PTMEA value of items of Learning Content

No	Item	PT-Measure Correlation	Construct
1	B2a	0.16	Guided Response – Theory
2	B2b	0.19	Guided Response – Practical
3	B2c	0.27	Guided Response – Technical Drawing
4	B23a	0.29	Application – Theory
5	B16a	0.29	Understanding – Theory
6	B22a	0.29	Characterization By Values – Theory
7	B26a	0.31	Guided Response – Theory
8	B3b	0.31	Guided Response – Practical
9	B1a	0.32	Perception – Theory
10	B28a	0.32	Perception – Theory

3.2.2 Polarity of items that measure the constructs of Pedagogical Decisions

Table 5 shows that there are 10 items indicating that network response to items or the respondents are inconsistent with the constructs. In other words, the items are not moving in line with other items to measure the constructs. The table shows a summary of the items and constructs that do not measure the constructs that are supposed to be measured. However those items will be retained after revisions and refinements because they fulfil the requirements. Items C9a should be omitted because of lack of fit to the model. These items are measuring something other than the intended content and construct.

Table 5: PTMEA value of items Of Pedagogical Decision

No	Item	PT-Measure Correlation	Construct
1	C5a	-0.3	Approach to Task
2	C7a	-0.2	Teaching & Learning Activities
3	C11a	-0.9	Visibility of processes
4	C10a	0.12	Organisation of space
5	C5b	0.13	Approach to Task
6	C10b	0.13	Teaching & Learning Activities
7	C9a	0.14	Role of Learner
8	C1a	0.18	Proximity to teacher
9	C3b	0.20	Role of Teacher
10	C4a	0.20	Role of Teacher

3.3 Suitability or fit of items in measuring constructs

The appropriateness of items in measuring the constructs can be seen in the total mean square Infit and mean square Outfit of each item and respondent. For polytomous data (Likert scale) the acceptable range of fit items for Likert scale is between 0.6 logits to 1.4 logits (Bond & Fox, 2007). If the items are out of the range, it must be separated, modified or rephrased (Linacre, 2005). Revision must be made to a problematic item because the suitability of an item can affect and influence the reliability and validity of the instrument.

3.3.1 Suitability or fit of items in measuring constructs of learning content

The appropriateness of items in measuring the constructs can be seen in the total mean square Infit and mean square Outfit of each item and the respondent. For polytomous data (Likert scale) the acceptable range of fit items for Likert scale is between 0.6 logits to 1.4 logits (Bond & Fox, 2007). If the items are out of the range, it must be separated in order to make modifications or rephrase (Linacre, 2005). This is due to the matter that suitability of an item will affect and influence the reliability and validity of the instrument.

Table 6 shows the measurement of misfit items or items that do not fit the Rasch measurement model for certain construct in Learning Content. A value that are higher than 1.4

indicate that the items are not homogeneous with other items in a measurement scale and value that is lower than 0.6 indicates redundancy with other items. Table 5 shows a total of 8 misfit items out of 84 items of Learning Content based on Outfit/Infit MNSQ index. The constructs and the total items are Guided Response (3), Internalizes Values (1), Organization (1), Knowledge (1) and Perception (2).

Table 6: Misfit items of Learning Content

No	Item	Outfit – Meansquare (MNSQ)	Construct
1	B2c	2.01	Guided Response – Technical Drawing
2	B26c	1.98	Guided Response – Technical Drawing
3	B20c	1.82	Guided Response – Technical Drawing
4	B22c	1.78	Characterization By Value – Technical Drawing
5	B27c	1.76	Organization – Technical Drawing
6	B25c	1.76	Knowledge - Technical Drawing
7	B28c	1.64	Perception - Technical Drawing
8	B1b	0.39	Perception – Practical

3.3.2 Suitability or fit of items in measuring constructs of pedagogical decision

Table 7 shows the measurement of misfit items or items that do not fit the Rasch measurement model for certain construct in Learning Content. A value that is higher than 1.4 indicates that the item is not homogeneous with other items in a measurement scale. Table 6 shows a total of 3 misfit items out of 33 items on Pedagogical Decision based on Outfit/Infit MNSQ index. The constructs and the total items are Role of Learner (1), Visibility of Processes (1) and Nature of Activities (1).

Table 7: Misfit items of Pedagogical Decision

No	Item	Outfit – Mean square (MNSQ)	Construct
1	C9a	1.83	Role of Learner
2	C11a	1.69	Visibility of processes
3	C10c	0.36	Nature of Activities

3.4 Discussion on the final instruments of learning content

Table 8 shows a summary of the items that need to be repaired or removed and the number of items that remain. A total of 3 items of Learning Content constructs have been modified, 0 items need to be dropped and 81 items are maintained.

Table 8: Summary of functionality examination of Learning Content item

No	Construct	Maintained Item	Total Maintained Item	Improve Item	Total Improved Item
1	Characterization by Value	B19a, B19b, B19c, B22a, B22b, B22c, B24a, B24b, B24c	9	0	0
2	Set	B17a, B17b, B17c	3	0	0
3	Responding	B8a, B8b, B8c, B9a, B9b, B9c, B10a, B10b, B10c, B11a, B11b, B11c, B13a, B13b, B13c, B21a, B21b, B21c	18	0	0
4	Perception	B1a, B1b, B1c, B28a, B28b, B28c	6	0	0
5	Organization	B6a, B6b, B6c, B27a, B27b, B27c	6	0	0
6	Complex Over Response	B4a, B4b, B4c, B5a, B5b, B5c	6	0	0
7	Understanding	B14a, B14b, B14c, B16a, B16b, B16c	6	0	0
8	Application	B12a, B12b, B12c, B18a, B18b, B18c, B23a, B23b, B23c	9	0	0
9	Knowledge	B7a, B7b, B7c, B15a, B15b, B15c, B25a, B25b, B25c	9	0	0
11	Guided Response	B2a, B2b, B2c, B3a, B3b, B3c, B20a, B20b, B20c, B26a, B26b, B26c	9	B2a, B2b, B2c	3

3.5 Discussion on the final instruments of Pedagogical Decision

Table 9 shows a summary of the items that need to be repaired or removed and the number of items that remain. A total of 3 items of Pedagogical Decision constructs have been modified, 1 item need to be dropped and the numbers of maintained items are 28 items.

Table 9: Summary of functionality examination of Pedagogical Decisions items

No	Construct	Maintained Item	Total Maintained Item	Improve Item	Drop Item	Total Improved Item	Total Drop Item
1	Nature of activities	C1c, C2c, C3c, C4c, C5c, C6c, C7c, C8c, C9c, C10c, C11c	10	C10c	0	1	0
2	Role of teacher	C1b, C2a, C2b, C3b, C4a, C4b, C6a, C6b, C8a, C8b, C9b, C11b	13	0	0	0	0
3	Teaching and learning activities	C7b, C10b	2	C7a	0	1	0
4	Approach to task	C5a, C5b	2	0	0	0	0
5	Role of learner	C1a	1	0	C9a	0	1
6	Visibility of processes	C10a	1	C11a	0	1	0

4. CONCLUSION

This study set out to validate the constructs for a newly drafted vocational pedagogy instrument. Item analysis using the Rasch measurement model was used to establish the suitability of items associated with identified constructs. The method enabled instrument designer to determine if responses to items are consistent with theoretical expectations resulting in items that are consistent with the purpose of measurement through the use of two indices, item reliability index and respondent reliability index. The reliability coefficient of scores obtained in the study is high indicating that the items are stable. Separation index for the level of difficulty exceeded the value of 2 indicating that items are strongly accepted. From the findings, threat regarding construct irrelevant-variance was minimum based on the dimensionality test as well as the within-range fit indices.

Although most of the items are moving in a similar direction, there are also a few items that do not contribute meaningfully to the measurement of the desired constructs. A review of the reliability and validity of the content of the instrument indicate that 1 item needs to be dropped from the 117 items, while 6 items need to be revised and improved. Removal of items that are not compatible with the model is necessary to improve the validity and reliability of the vocational pedagogy instrument. Finally, 116 items are found to fit the Rasch model with an indication of unidimensionality.

The construct validity and reliability of the items for the vocational pedagogy instrument are thus established and the instrument can be used to investigate vocational practices among Malaysian and Indonesian TVET teachers in the future. The use of the

instrument by both countries can facilitate future comparative studies that will promote a better understanding of TVET vocational pedagogical practices in both countries in particular and in Asia in general. Data obtained from the established instrument can also be used to develop a transnational framework that can promote greater TVET collaborations between Malaysia and Indonesia. Extending the validation efforts to include TVET teachers from other Asian countries can further enhance the utility of the new instrument in the Asian TVET sector.

References

Abrami, Aslan & Nicolaidou (2007). *Teaching and Learning Strategies Questionnaire.* Centre for the study of learning and performance.

Anthony, R. & Artino, J. (2012). *A Review of the Motivated Strategies for Learning Questionnaire***.** Running head: REVIEW OF MLSQ**,** University of Connecticut.

Asia Pacific Education System Review Series. (2015). *Transversal Skills in TVET: Pedagogies and Assessment.* UNESCO 2015

Bestler, L. (2007) *Student Evaluation of Teaching (SET) Sample Rubric Developed by CELT***.** DEMO-SET Questionnaire 2011 Spring, Iowa State University.

Brown, B.L. (2012). *Clearinghouse on Adult, Career, and Vocational Education.* Learning Styles and Vocational Education Practice*.*

Bond & Fox. (2007). *Applying the Rasch Model*: Fundamental Measurement in the Human Sciences.

Duckett, I., Tatarkowski, M. (2015). *Practical Strategies for Learning and Teaching on Vocational Programmes*. Vocational Learning Support Programme : 16-19. Learning and skills development agency.

Echternacht, & Wen, (1997).*Learning Strategies and Styles in Vocational Education*. Trefort Ágoston Centre for Engineering Education, Óbuda University

Economic and Social Research Council (2002). *Experiences of Teaching & Learning Questionnaire***.** Enhancing Teaching-Learning Environment in Undergraduate Courses : ETL Project.

Fayombo, G. (2015). *Learning Styles, Teaching Strategies and Academic Achievement among some Psychology Undergraduates in Barbados.* Caribbean Educational Research Journal The University of the West Indies : Vol. 3, No. 2, 46-61.

Gargallo, Bernardo, Suarez-Rodriguez, Jesus M. & Perez Cruz. (2009). *An Instrument to Assess The Learning Strategies of University Students:* E-Journal of Educational Research, Assessment and Evaluation.

International Association for the Evaluation of Educational Achievement (IEA) The Netherlands (2009). *Teacher Questionnaire OECD Teaching and Learning International Survey (TALIS).* IEA Dat Processing and Research Centre, Germany Statistics, Canada.

Jane, K., Lartec, Anastacia, M., Belisario, Jamaica, P., Bendanillo, Hanni, K., Binas-o, Novefirst, O., Bucang, Jan Lorie, W. & Cammagay. (2010). *Strategies and Problems Encountered by Teachers in Implementing Mother*, *Tongue - Based Instruction in a Multilingual Classroom*. School of Teacher Education, Saint Louis University, Baguio City, Philippines.

Linacre, J. M. (2005). Test validity, and Rasch measurement: Construct, content, etc. Rasch measurement transactions.

Lucas, B., Spencer, E. & Claxton, G. (2012). *How to teach vocational education:* A theory of vocational pedagogy.

Lysadht, Z. & O'Leary, M. (2016). *An Instrument to Audit Teachers' Use Of assessment For Learning***.** Irish Educational Studies.

Moreno-Murcia, J.A., Torregrosa, Y.S. & Pedreno, N.B. (2015). *Questionnaire Evaluating Teaching Competencies In the University Environment. Evaluation of Teaching Competencies in the University.* New Approaches in Educational Research

OSO Oluwatumbi, S. (2010*). ICT literacy among vocational and technical education teachers in kogi state technical and vocational colleges*: skill gaps Department of Curriculum Studies .Faculty of Education, Ekiti State University, Ado-Ekiti.

Pallant, K. (2011). Transfer of Training: *A Review and Direction for Future Research. Personnel Psychology*, 41(1), 63-105.

PTC Leading Professional Development. (2006). *Instructional Strategies Survey***.** EARCOS.

Strakal , D.J. (2007). *The Use Of Real Life Learning Strategies In Personal and Career Development Situation.* Montana University.
Wright, S. & Master, (1982). *Applying the Rasch Model*: Fundamental Measurement in the Human Sciences.