Journal of Fundamental and Applied Sciences

ISSN 1112-9867

Available online at http://www.jfas.info

Research Article

Special Issue

DEVELOPMENT AND VALIDATION OF AN INSTRUMENT TO MEASURE UNDERSTANDING IN DAILY LESSON PLAN

N. H. C. M. Ghazali*, N. M. Rabi, N. A. Wahab, N. A. A. Rohaizad, N. A. Rahman

Faculty of Education and Human Development, Sultan Idris Education University, Tanjong Malim, Perak

Published online: 24 February 2018

ABSTRACT

A valid, reliable and practical instrument is needed in measuring any kind of understanding. The purpose of this study is to develop and validate an instrument to measure the understanding in the daily lesson plan (DLP). The whole population is chosen as samples which involves 241 students enrolling in the bachelor degree of education in the final semester of the 2015/2016 session in the Malaysian institution of higher learning. The instrument (K-RPH) is developed based on the principles in the Madeline Hunter Model of Mastery Learning and The ASSURE Model. The instrument consists of 32 items with 160 rubrics with four main constructs namely analysing DLP, determining the objective of DLP, preparing learning experiences and evaluating and reflecting. Content validity is conducted by two experts in the field from the university. The instrument validity and reliability have been assessed using the Rasch Measurement Model by identifying the Rasch fit statistics, item difficulty, unidimensionality, item dispersion and reliability as well as the item distribution map. The Rasch analysis shows that the item reliability is 0.99 while the Cronbach Alpha is 0.96. All the items fit the model as their MNSQ values are between 0.75 and 1.49. The dispersion of items from K-RPH data is 10.29 which indicates the existence of 11 to 12 item strata group. No item shows a negative point measure correlation or less than 0.2, and this generally indicated that the item discrimination is very good.

Author Correspondence, e-mail: adinsah@fmpp.ispu.ude.ym

doi: http://dx.doi.org/10.4314/jfas.v10i3s.22

The data shows that the mean for person is 1.38 logits with a standard deviation of 1.67 logits, while the item mean is zero with a standard deviation of 1.02. This indicates that the position of item and person do not fully match, and thus shows a medium difficulty. The overall item quality is good and all 32 items of K-RPH are retained.

Keywords: Instrument development, construct validity, reliability, Daily Lesson Plan, Rasch Model

1. INTRODUCTION

Teachers have to plan their teachings carefully and understand the needs of each components in the DLP if they were to teach effectively. If the teaching is not carefully planned, the teaching would not be in order, boring and yet could not produce a useful human capital. Even, if we look at the Malaysian Educational Development Plan 2011, it says that the teaching and learning process in a classroom is a main indicator of the future success of a country. Furthermore, the issues of truancy is becoming increasingly worrying. Some of the factors that contribute to truancy is the students are not interested with the teaching and learning session in schools (Suhaizah, 2008). This suggests that teachers have to do something to make class more cheerful and fun. Even, planning teaching should be an important task of a teacher (Moonsri & Pattanajak, 2013). Teachings are boring due to the weakness in planning the teachings and also the fact that teachers do not really understand the lesson plan to teach effectively. In addition, long time ago, Tyler (1949) stated that if teachers do not plan teachings thoroughly, how can they were able to teach in an effective way and later could change the students' behavior? An effective teaching comes from an effective and a good quality planning.

An effective teachings involve two issues; those enable students to gain skills, knowledge and values that we expect and, the second is those involving the active engagement of students. So, a good quality planning is very important and very much needed so that every second in this classroom is not being overlooked. Furthermore, the purpose of planning is to enable assignments being done effectively (Akhiar and Shamsina, 2010). And, planning also helps in giving guidance in the future, helping teachers to come out with great ideas, help in time-management, helps in finding the problem in teaching and also to make teaching become more structured.

Several researches have been conducted on teacher literacy in writing DLP but there is none focusing on the students' understanding in DLP. In general, most teachers know how to write DLP, but do they really understand the rationale behind each component in DLP? A research

done by Jiar and Siti Sara (2010) on student teachers from UTM found that the pedagogic level among them are high. There are four main constructs involved which are planning, classroom management, measurement and evaluation and communication skills. However, this study is only focusing on knowledge aspects and not on their understanding in planning teaching. Another study is a qualitative study by Harris and Hofer (2009). They do research related to producing an instructional plan base on activity. Learning activity influences the development of an instructional plan and its preparation must focus on students' standard, curriculum, the learning outcome and the existing technology. However, their study is focusing on teaching plan technique and teachers' knowledge in planning and not on understanding of each component in the DLP document. So, there is a need to do this research by producing an instrument in determining the level of understanding in implementing DLP among student teachers.

An instrument is valid when it is measuring what it is supposed to measure (Muijs, 2011). Or, in other words, when an instrument accurately measures any prescribed variable it is considered a valid instrument for that particular variable. There are four types of validity; face validity, criterion validity, content validity or construct validity (Jackson, 2003 and Muijs, 2011). Reliability on the other hand is defined as 'the extent to which test scores are free from measurement error' (Muijs, 2011, p.61). It is a measure of stability or internal consistency of an instrument in measuring certain concepts (Jackson, 2003). There are two types of reliability, person reliability and item reliability. Person reliability index indicates the replicability of person ordering we could expect of this sample of persons were given another a parallel set of items measuring the same construct and item reliability indicates the replicability of item placements along the pathway if these items were given another same-sized sample of persons who behaved in the same way (Wright & Masters, 1982). There is a relationship between validity and reliability. Any instrument can be reliable but not valid however, it cannot be valid if it is not reliable (Jackson, 2003). In other words, if an instrument is valid, it must be reliable. And, in general, checking for validity of an instrument is more difficult than checking for reliability because validity is measuring data related to knowledge whereas reliability only concerns with the consistency of scores. In this study, face validity, content validity and construct validity are measured using Rasch Measurement Model application based on the Rasch fit statistics, item difficulty, unidimensionality, item dispersion and reliability as well as the item distribution map.

The aim of this study is to develop and validate an instrument to measure the understanding in the daily lesson plan (DLP). The Rasch analysis is conducted using six steps: the Rasch fit

statistics, item difficulty measurement, item polarity, unidimensionality, dispersion and reliability, and the item distribution map.

2. LITERATURE REVIEW

There are various definitions of DLP. According to Fredericks (2011), DLP is an instructional guideline for a teacher in a classroom. It is also one way for teachers to know what the teacher wants the students to study and how they are going to deliver it. DLP is seen as a planning to make decisions (Orstein & Lasley, 2000). Planning teaching involves two things; the knowledge about the subject he or she were about to teach and the feedback system. Teachers have to focus on teaching activities such as diagnosing, classifying, organizing and evaluating students. DLP has several functions. First, it is used to help teachers in teaching and learning sessions especially in controlling the class (Akhiar & Shamsina, 2000). DLP is also a medium which prepare instructions to help students achieve their desired level of understanding. It is also to ensure that students learn what they need to know (Fredericks, 2011). It could also be for the purpose of preparing substitute teacher with an orderly instructional plan if the real teacher could not make it. In addition, DLP also makes teachers feel prepared and it could help them to avoid nervousness during teaching (Fernandez & Yoshida, 2004). DLP also prepares some teachers to face differences between individuals in the classroom (Fredericks, 2011).

In order to plan their teaching, student teachers have to prepare themselves with pedagogic knowledge and also DLP writing model which are suitable to their teaching methods. In this study, two teaching models are used as a reference in writing DLP. The models are the ASSURE Model and the Madeline Hunter Model of Mastery Learning. The ASSURE teaching model is a model which gives direction for teachers to plan their teaching using media and technology (Akhiar Pardi & Shamsina, 2010). It is developed by Heinich, Molenda, Russell and Samldino from the Instructional Media and Technologies for Learning in 1982. This model is a very structured teaching model whereby the teaching objectives has been set up earlier explicitly. There are six steps involved: Analyze Learners (A), State Objective (S), Select Method, Media and Material (S), Utilize Media and Material (U), Require Learners Participation (R) and Evaluate and Revise (E). Instructional media is a media which brings message for teaching purpose aiming at helping teachers to communicate, stimulate thinking and also motivate students (Heinich, 2005). The second model is a model of mastery learning by Madeline Hunter (Burns, 2005). A model is based on mastery learning. According to Hunter, Lesson Design should have eight main elements no matter what the

teacher's style, grade level, subject matter or economic background of the students. These elements are suggested if teachers were to plan an effective instruction. The elements are an anticipatory set which is a short activity or prompt that focuses the students' attention before the actual lesson begins, the purpose of today's lesson, input (vocabulary, skills, and concepts), modelling, a guided practice, checking for understanding using various questioning techniques, an independent practice and lastly is a closure which is a review or a wrap-up of the lesson.

If we compare these two models, there are few similarities and differences between them. Both models are focusing on the analysis of students by teachers to ensure that teachers know their ability and existing knowledge. Next, both models suggest teachers to state specific teaching objectives which suits the students. Both models also stress on the need to evaluate and reflect. However, they also have few differences. The ASSURE Teaching Model plans teaching based on media and materials whereas Hunter is based on mastery learning. Next, in terms of planning for anticipatory set such as induction. It could be a hand-out given to students at the door, review questions written on the board, two short problems presented on a transparency on the overhead or an agenda for the lesson written on the chalkboard. This is stressed by Hunter but not really by the ASSURE Model. Lastly, the ASSURE Model stresses more on the usage of tools and teaching aids to increase students' interest and understanding, whereas Hunter is stressing more on demonstration and guided practice followed by independent practice. All in all, there is a critique made towards Hunter in Dewey (1929). Dewey states that the model is only focusing on mastery learning and does not encourage students to think critically and creatively. Thinking during learning has to relate to methods, content and the learning objectives. Hunter does not mention about encouraging cultivated thought as a main purpose of teaching. The researcher also feels that ASSURE Model has its limitations. ASSURE assumes teachers to have known the basic thing to plan teaching. And, in the second step of the model which is 'to state objective', it does not include stating the learning outcome which suits the students' needs and teaching objectives. It also does not include planning for assessment activities in achieving the learning outcome which has been stated earlier.

From the analysis done on both models plus theoretical principles which act as the basis of the models such as constructivisme, cognitivisme, behaviorisme, Gardner multiple intelligence theory and cooperative learning, the researcher suggests a model named M3P Model (as in Figure 1) for teachers to plan their teachings. Based on literature reviews, the model is also being improved by adding up learning theories like humanisme theory, problem-based

learning, project-based learning and evaluate teaching based on the CIPP Model. Finally, the model is constructed with four main constructs – analysing DLP, determining objectives, organizing learning experiences and evaluating and reflecting.

First is analysing teaching date, time and the students involved (Yulian, 2010). This is supported by Wile and Shouppe (2011) who believe that learning is improved if students learn at the right time. This contradicts with Hunter and ASSURE Model which focus only on analysing students. In addition, constrctivisme theory believes that a student is not an empty tin when he enters classroom but with a different background, experiences and existing knowledge. Each student also have his own learning style (Jarvis, 2005) so analysing students is important. Secondly, to determine learning objective. Teachers have to decide the learning objectives, the learning outcome, materials for assessment and evaluation and also teaching aids (Mager, 2009). Learning should start with abstract experiences, followed by iconic experiences and then enactive experiences (Dale, 1946). So, this is where materials are needed during teaching and learning. Thirdly, teachers have to arrange learning experiences which is more to student-centered. Arrangement must be concordance with thinking development, thinking level and topic. This is concordance with cognitive theory (Fosnot & Perry, 1996), constructivism theory (Lutz & Huitt, 2004) and cooperative learning (Felder & Brent, 2007). In order to arrange the learning experiences, teacher could use problem-based learning (Ferreira, 2012) or project-based learning (Endo, 2015). Lastly, after the teaching session, teachers should do evaluation and reflection. Not only evaluating teachings but also to evaluate each level of teaching, the assessment materials and also the teaching aids. In this study, evaluation is using the CIPP Model by Daniel Stufflebeam which evaluates the context, input, process and product dimension (Stufflebeam, 2012). From this model, the elements of the constructs of this study are developed.

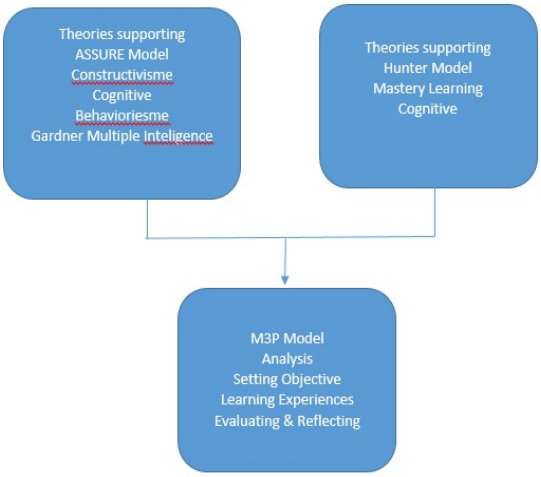

Fig.1. M3P Model

3. METHODOLOGY

This study is a quantitative approach study which involves the collection of data using questionnaire. The questionnaire was administered to 241 student teachers at higher education institutions in Malaysia. A census is used as they are all the ones who registered that semester. The instrument consists of 32 items with 160 rubrics and it is a double-layered type of questionnaire. Respondents were asked to write the score following the scale 0=none, 1=partly and 2=fully. A scale of 5 point are used after the summation for each item which consists of 5 items in it. Rubric score will be changed to ordinal scale as in Table 1.

Table 1. Rubric Scoring is changed to ordinal scale

Rubric Scoring	Ordinal Scale
0 – 2	Very Low = 1
3 – 4	Low = 2
5 – 6	Medium = 3
7 – 8	High = 4
9 - 10	Very high = 5

The instrument is used to gauge 4 constructs which covers students' understanding; i) to analyse DLP (11 items); ii) to determine the objective of DLP (5 items); iii) to organize learning experiences (10 items); and iv) to evaluate and reflect (6 items). Items are analysed using WINSTEP software based on the Rasch Model. According to Green and Frantom

(2002), Rasch analysis requires a sample of 100 respondents and 20 items for the data to be considered stable, so this study is suitable enough for that.

The instrument is developed in several stages as stated below

1. Determining the main constructs from few models on lesson plan such as Madeline Hunter Model of Mastery Learning and the ASSURE Teaching Model. These steps are done in order to gain the rationale of writing each component of the DLP. Based on the models, the researcher has developed a model known as M3P Model (as shown in Figure 1) using the principles of both models.
2. Interview protocols are conducted with five respondents. This is to know the real thing that happen when teaching plan is conducted. Some of the questions asked are: 'What do you understand about DLP?', 'Why do you have to record your DLP?' or 'What is the function of DLP?'.
3. Next, based on literature and the interviews, an Item Determination Table is formed which have constructs, sub-constructs, variables, questions, items and scales. Using that table, a detailed rubric is developed consisting instructions, statements and scale.

4. RESULT AND DISCUSSION

4.1 Fit Statistics

Item fit statistics (infit MNSQ and outfit MNSQ statistics) are examined to ensure that the items are contributing meaningfully to the measurement of the construct (Linacre, 2011). The recommended acceptable range for infit MNSQ and outfit MNSQ fit statistics for rating scale is MNSQ $\geq$ 0.50 to MNSQ $\leq$ 1.50. Items within this range are considered productive (Bond and Fox, 2007). If the individual item does not fill the requirements, then the item will be eliminated. Table 1 reveals that all items show good overall fit of the data to Rasch Model. Only one item (S30) shows poor fit (INFIT < 0.50 logit), and only one item (S31) has OUTFIT > 1.5 logit. However, all the items are retained. The infit MNSQ laid between 0.75 and 1.49 and according to Bond and Fox (2003), the data fits the model if the infit MNSQ is between 0.6 and 1.4.

4.2 Item difficulty measurement

Item difficulty can be defined as a state of variable continuum from easy to more difficult and it is measured using logits. The item validity is defined via the assessment of item difficulty whereby all of the items are arranged in a hierarchical position to define each construct. The

organization of items difficulty is shown in Table 1. The instrument validity according to the Rasch model is the construct validity or the order of the items (Wright and Stone, 1979). In Rasch model, the mean of an item is normally considered as zero (Bond and Fox, 2001). If the item measure and the ability of an individual match closely, then the item would provide a lot of information about the individual, and this is known as a latent trait. If the mean of an individual falls in the range of 2 standard deviations from the mean, then the instrument can be 'targeted'. The 'target' of this instrument are sufficient as the highest measurement (item 30) is 4.88 (in the range of 5 standard deviations) while the lowest measurement in 5 standard deviations. All items fit the model as their MNSQ values are between 0.75 and 1.49.

Table 2. Item measure (INFIT, OUTFIT) MNSQ and Point Measure Correlation

Item	Measure	Std Error	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PT MEA CORR
S30R	4.88	0.04	0.25	-0.30	0.93	-0.30	0.36
S15M	-0.1	0.10	1.49	4.80	1.57	5.20	0.57
S1A	-0.37	0.09	1.29	3.00	1.57	4.60	0.65
S14M	0.16	0.09	1.34	3.40	1.27	2.40	0.66
S4A	-0.14	0.10	1.18	1.90	1.19	1.80	0.68
S3A	0.80	0.09	1.24	2.60	1.21	1.80	0.69
S19S	-0.05	-0.05	1.11	1.30	1.14	1.40	0.69
S27R	0.37	0.37	1.05	0.60	0.98	-0.20	0.70
S5A	-0.14	-0.14	1.09	1.10	1.07	0.80	0.70
S31R	-0.76	-0.76	0.97	-0.30	2.16	6.90	0.72
S29R	-0.18	-0.18	1.05	0.60	1.10	0.80	0.72
S26S	-0.80	-0.80	0.97	-0.30	0.94	-0.40	0.72
S2A	-0.74	-0.74	0.97	-0.30	0.97	-0.20	0.73
S7A	-0.52	-0.52	1.04	0.50	1.00	0.00	0.73
S6A	0.10	0.10	1.05	0.50	1.02	0.20	0.73
S9A	0.04	0.04	0.91	-0.90	0.83	-1.20	0.73
S12M	-0.67	-0.67	0.94	-0.60	0.92	-0.70	0.74
S21S	0.37	0.37	0.98	-0.20	0.86	-1.10	0.74
S17S	0.08	0.08	0.88	-1.40	0.86	-1.00	0.74

S11A	0.84	0.84	1.00	0.00	0.93	-0.70	0.74
S28R	-0.43	-0.43	0.97	-0.30	0.94	-0.50	0.74
S13M	-0.18	-0.18	0.96	-0.40	0.91	-0.80	0.74
S25S	-0.55	-0.55	0.91	-1.00	0.83	-1.50	0.75
S32S	-0.50	-0.50	0.88	-1.30	0.82	-1.10	0.75
S10A	-0.23	-0.23	0.90	-1.10	0.88	-1.10	0.76
S18S	-0.09	-0.09	0.86	-1.70	0.90	-0.80	0.76
S8A	-0.67	-0.67	0.84	-1.80	0.80	-1.90	0.77
S16M	-0.83	-0.83	0.81	-2.20	0.74	-2.00	0.77
S22S	-0.63	-0.63	0.83	-1.90	0.79	-2.20	0.77
S20S	0.80	0.80	0.82	-2.10	0.76	-2.50	0.78
S24S	0.42	0.42	0.78	-2.60	0.68	-2.60	0.78
S23S	0.42	0.42	0.75	-3.10	0.68	-2.90	0.79
Mean	**0.00**	**0.00**	**0.97**	**-0.10**	**1.01**	**0.00**	
SD	**0.98**	**0.98**	**0.21**	**1.80**	**0.29**	**2.20**	

4.3 Item Polarity

All items show positive item discrimination and a pattern which showed a high validity via a positive correlation point size value. Point Measure Correlation (PMC) is a statistical item showing the correlation results between one points (a response choice) with a continuous variable (scores for all candidates in a test). In Rasch statistics, the mean square value of the residual item which is sensitive to the items which have failed to relate to the test scores and point-biserial items with very large values is considered. It means the correlation point size in Rasch statistics is sensitive to the interaction of items, which do not follow a certain model in the calibration sample (Wright and Stone, 1979). The acceptable critical point measure correlation of an item is 0.2 or more (Pray and Popovich, 1985). In addition, a discrimination index of less than 0.2 is weak and more than 0.4 is good (Masey, 1995). From Table 1, a lowest value for PMC is 0.36. No items shows a negative PMC or less than 0.2. All items (except for one item) show values more than 0.4. This indicates that the item discrimination is very good.

4.4 Unidimensionality

Unidimensionality is critical in determining an instrument which is measuring in one dimension. An instrument which is not exact in measuring what it supposed to measure could give a confusing outcome. According to Azrilah *et al*. (2013), Rasch analysis needs at least 40 % of raw variance explained by measurement as an indicator of a good unidimensionality, and noise level must be less than 15%. From Table 2, it shows that the raw variance explained by measurement is 97.4% (which is more than 40%) and noise level is 7.9% (which is less than 15%). This indicates that all the items are clear and not confusing.

Table 3. Standard residual variance (in Eigenvalue)

	Empirical (in %)	%	Model (%)
Number of raw variance in observation	1218.4 (100.0)		100.0
Raw variance explained by measurement	1186.4 (97.4)		97.4
Raw variance explained by the students	535.7 (43.9)		43.9
Raw variance explained by the items	650.7 (53.4)		53.4
Number of unexplained raw variance	32.0 (2.6)	100	2.6
Unexplained variance in the first contrast	2.5 (7.9)		7.9

4.5 Dispersion and Reliability

Fit statistics also enable the researcher to detect whether each item contributes to the measure of each construct. The value of item reliability indicates whether the items or cluster of items interact well with one another in describing the same attributes (Wright and Stone, 1979). A person reliability is explained on a scale of 0 to 1, and this provides meaning just like the Cronbach alpha value. Dispersed items and people are calibrated. The dispersed item, people and reliability are used to assess the arte of dispersion across the trait continuum. It measures the dispersion of items and people in standard unit. The dispersion of instrument to be used should reach the value of 1 (a high dispersion value shows there is item and person dispersion further along the continuum; a low dispersion value shows there might be an overlapping items and less person variability in the trait). So, in short, dispersion is used to describe how a strata of latent traits can be found using item measurement (Full agreement=2, some=1 and none=0).Expected dispersion should reach the value of 2.0 to describe all the three strata. Table 3 shows the dispersion of items from the instrument is 10.29, which indicates the existence of 10 to 11 item strata while the dispersion of person is 5.20 which indicates the existence of 5 to 6 people strata. Dispersion indicates reliability. Dispersion reliability for people generally will be similar to Cronbach alpha value which shows an instrument's internal consistency reliability. Item reliability is 0.99 and students' reliability is 0.96. A reliability value of more than 0.8 is considered a very good reliability value (Linacre, 2007).

Table 4. Summary of person and item measure

	Measurement	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
Mean	1.38	0.98	-0.2	1.01	-0.3
SD	1.67	0.49	2.0	0.75	2.0
Dispersion	5.20				
Reliability	0.96				
Mean	0.00	0.98	-0.1	1.03	0.1
SD	1.02	0.20	1.8	0.36	2.4
Dispersion	10.29				
Reliability	0.99				

4.6 Item Distribution Map

The item distribution map shows the distribution of person and item on the same measurement scale. The scale measures constructs vertically with the most capable person and the most difficult item are placed at the top. The column on the left in Table 3 shows the measure of the person's capability in logits. Table 1 also shows that the item distribution map and it enables the researcher to observe the item function and students' overall capability measurement. To assess item distribution, items need to be less than -2 logits to +2 logits. If the empty value between items is more than 0.15 logits, the items differ between one another and if the empty value is more than 0.30 logits, then it has to be filled with other items. Table 3 shows the standard error between items is 0.15 and this indicates that item dispersion exists and items differ from one another. The mean for person is 1.38 logits with a standard deviation of 1.67 logits, while the mean for item is 0.00 logits with a standard deviation of 1.02 logits. This indicates that the position of person and item does not fully match, and thus shows a medium difficulty. As in Table 1, a total of 12 items are above the mean and these items are considered difficult to understand, in relative. Item 30 is the most difficult item

whereby the students appeared to have difficulties in understanding. A total of 20 items are below the mean and these items are considered easy to understand, in relative.

5. CONCLUSION

This study aims to develop and validate an instrument to assess understanding in the DLP amongst students. The development of instrument goes through several processes. The validation uses Rasch measurement model. The analysis revealed that all the items fit the model as their MSNQ items are between 0.7 and 1.35. The item with the lowest point measure correlation is 0.35. No item showed a negative point measure correlation or less than 0.2. This indicates that the item discrimination is very good. Cronbach Alpha value of 0.96 also indicates a clear uni-dimensionality. Item reliability is 0.99 and teacher reliability is 0.96 that a value of more than 0.8 showed very good reliability. The data shows that the mean for person is 1.38 logits with a standard deviation of 1.67 logits, while the item mean is zero with a standard deviation of 1.02. This indicates that the position of item and person do not fully match, and thus shows a medium difficulty. The overall item quality is good and all 32 items of K-RPH are retained.

REFERENCES

Akhiar, P. & Shamsina, S. (2010). Rancangan Pengajaran Harian. Puchong: Penerbitan Multimedia Sdn. Bhd.

Azrilah, A. A., Mohd Saifuddin, M. and Azami, Z. (2013). Asas Model Pengukuran Rasch. Bangi: Penerbit UKM.

Bond, T. G. and Fox, C. M. (2003). Applying the Rasch Model: Fundamental Measurement in the Human Sciences, Journal of Educational Measurement, 40(2): 185-187.

Bond, T.G. and Fox, C. M. (2001). Applying the Rasch model: Fundamental measurement in the human sciences. Mahwah, N.J.: Lawrence Erlbaum.

Bond, T. G. and Fox, C. M. (2007). Applying the Rasch Model: Fundamental Measurement in the Human Sciences. Mahwah, New Jersey: Lawrence Erlbaum.

Burns, A. C. (2005). Teaching experientially with the Madeline Hunter Method: An Application in a Marketing Research Course, Developments in Business Simulation and Experiential Learning, 32: 47-53.

Dale, E. (1946). Audio-visual methods in teaching. New York: The Dryden Press.

Dewey, J. (1929). The Quest for Certainty: A Study of the Relation of Knowledge and Action. New York: Minton, Balch and Company.

Felder, R. M. and Brent, R. (2007). Coorperative Learning in Active Learning: Models from the analytical sciences, ACS Symposium Series, 970: 34-53.

Fernandez, C. and Yoshida, M. (2004). Lesson Study: An Introduction June 2014 from www.globaledresources.com

Ferreira, M. (2012). Invited Discussion of Spatial quantile multiple regression using the asymmetric Laplace process by K. Lum and A. E. Gelfand, Bayesian Analysis, vol. 7: 235–276.

Fosnot, C. T. and Perry, R. S. (1996). Constructivism: A Psychological Theory of Learning, Constructivism: Theory, perspectives and practice, 2: 8-33.

Fredericks, A. D. (2011). The Teacher's Handbook New Delhi: Rowman & Littlefield Education Lanham.

Green, K. E. and Frantom, C. G. (2002). Survey development and validation with the Rasch Model. In International Conference on Questionnaire Development, Evaluation and Testing, Charleston, USA.

Harris, J. and Hofer, M. (2009). Instructional Planning Activity Types as Vehicles for Curriculum-Based TPACK Development Introduction: TPACK. Information Technology in Teacher Education, 2: 99-108.

Heinich, R. (2005). Instructional Media and Technologies For Learning, New Jersey: Merrill Prentice-Hall.

Jackson, S. L. (2003) Research Methods and Statistics, A Critical Thinking Approach. USA: Thomson Wadsworth.

Jarvis, M. (2005). The Psychology of Effective Learning and Teaching. Cheltenham: Nelson Thornes, Ltd.

Jiar, Y. K. and Siti Sara, A. H. (2010) Tahap Pengetahuan Pedagogi Pelajar Tahun Akhir, Fakulti Pendidikan Dalam Pengajaran Dan Pembelajaran. Tahap Pengetahuan Pedagogi Pelajar Tahun Akhir, Fakulti Pendidikan Dalam Pengajaran Dan Pembelajaran. pp. 1-8. (Unpublished)

Linacre, J. M. (2007). Standard errors and reliabilities: Rasch and raw score. Rasch Measurement Transactions, 20(4): 1086.

Linacre, J. M. (2011). A User′s Guide to Winstep. Ministep Rasch-Model Computer Programs. Program Manual 3.73.0.

Linacre, J. M. (2007). A users' guide to WINSTEPS Rasch Model computer programs.

Linacre J. M & Fisher W. P. (2012). Harvey Goldstein's objections to Rasch measurement: A response from Linacre and Fisher. Rasch Measurement Transactions, 26(3): 1383-1389.

Lutz, S., and Huitt, W. (2004). Information processing and memory: Theory and applications. Educational Psychology Interactive. Valdosta, GA: Retrieved January 2010, from http://www.edpsycinteractive.org/brilstar/chapters/infoproc.pdf

Mager, R. F. (2009). Menyediakan Objektif Pengajaran. Kuala Lumpur: Institut Terjemahan Negara Malaysia Berhad & Penerbit UPSI.

Masey, A. (1995). Evaluation and analysis of examination data: Some guidelines for reporting and interpretation. UCLES Internal report, Cambridge, UK.

Moonsri, A. and Pattanajak, A. (2013). Lesson Planning in Primary School Using Lesson Study and Open Approach. Scientific Research, 4(12), 1064-1068 dari http://doi.org/10.4236/psych.2013.412155

Muijs, D. (2011). Doing Quantitative Research in Education with SPSS. London: SAGE Publications Ltd.

Orstein, A. C. and Lasley, T. J. (2000). Strategies for Effective Teaching. USA: The McGraw-Hill Companies.

Pray, W. S. and Popovich, N. G. (1985). The development of a standardized competency examination for doctor of pharmacy students. American Journal of Pharmaceutical Education, 49(1): 1-9.

Stufflebeam, D. L. (2001) Evaluation Models: New Directions for Evaluation, 7-98, John Wiley and Sons.

Tyler, R. W. (1949). Prinsip Asas Kurikulum dan Pengajaran. Kuala Lumpur: Dewan Bahasa dan Pustaka.

Wile, A. J. and Shouppe, G. A. (2011). Does Time-of-Day of Instruction Impact Class Achievement? Studies, 13.

Wright, B. D. and Masters, G. N. (1982). Rating scale analysis. Chicago: MESA Press.

Wright, B. D. and Stone, M. H. (1979). Best Test Design (Rasch Measurement Series), MESA Press, Chicago, USA.

Yulian, Y. (2010). Searching for the Monday Blues in Order Fulfilment and its Cure. Social Science Research Network.

How to cite this article:
Ghazali N H C M, Rabi N M, Wahab N A, Rohaizad N A A, Rahman N A. Development and validation of an instrument to measure understanding in daily lesson plan. J. Fundam. Appl. Sci., 2018, *10(3S), 268-283.*