INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN PROGRESSIVE EDUCATION & DEVELOPMENT

OPEN ACCESS JOURNAL

Vol 12, Issue 4, (2023) E-ISSN: 2226-6348

Measuring Reliability and Validity of Questionnaire on Teacher Digital Professional Development: A Rasch Model Analysis

Goh Kok Ming

Sultan Idris Education University

Email: kokming888@gmail.com

To Link this Article: http://dx.doi.org/10.6007/IJARPED/v12-i4/19662 DOI:10.6007/IJARPED/v12-i4/19662

Published Online: 22 November, 2023

Abstract

In this study, the validity and reliability of a questionnaire on Teacher Digital Professional Development (TDPD) were empirically examined using Rasch model analysis. The questionnaire was filled out by 125 primary school teachers in total. For this particular survey, the questionnaire was adapted. Three experts who specialise in educational technology have validated the content. The TDPD has great reliability within six constructs, based on the data analysis performed using WINSTEP software version 3.71.01. It was found that the TDPD is reliable and can be accepted based on the person reliability of 0.95 and the item reliability of 0.81. However, one element had to be eliminated after being subjected to the Rasch model analysis, while the remaining items could be maintained. Researchers will be able to use or adapt this high-quality instrument in their research if reliable and valid instrumentation is assured.

Keywords: Reliability, Validity, Rasch Model Analysis, Teacher Digital Professional Development, Questionnaire

Introduction

Digital technology brings great potential in the effort to transform teacher learning and delivery methods of teacher professionalism development activities. Digital professional development has become a popular choice for teachers to receive continuous professionalism training (World Bank, 2021). There is a increasingly need to explore professional development more deeply when it has gone beyond a face-to-face context to a technology-oriented digital context. However, there is lack of related questionnaires to examine teacher digital professional development. In that regard, this Teacher Digital Professional Development (TDPD) questionanaire was adapted from previous studies and tested with Rasch Model Analysis. To determine the reliability of a questionnaire, Crobanch's Alpha coefficient and correlation or factor analysis values are the typical methods used by most researchers (Che Lah & Tasir, 2018). Not many, however, have applied Rasch Model Analysis to measure the reliability and validity of their instruments especially questionnaires (Che Lah & Tasir, 2018). Therefore, researcher chose to use Rasch Model Analysis on the TPDP questionnaire in this study to examine its validity and reliability.

Rasch Model Analysis is a well-known item-response theory (IRT), and it poses as a strategic method for instrument measurement, as it can reveal the quality of the items used concomitantly (Kean et al., 2018). It can be used either when developing new instruments or when evaluating and revising existing ones (Bortolotti et al., 2013). According to Che Lah and Tasir (2018), there are seven types of Rasch Model Analysis, which includes: (i) summary statistics, (ii) item fit, (iii) infit (MNSQ) and infit (ZSTD) values, (iv) unidimensionality, (v) differential item functioning, (vi) validity of scale rating, and, (vii) person-item. However in this study, six specific types were tested: (i) infit (MNSQ) and infit (ZSTD) values, (ii) item polarity, (ii) measure value, (iii) standardized residual correlation, (iv) unidimensionality, and (v) summary statistics. To validate the content of questionnaire, three experts, from Universiti Teknologi Malaysia, Johor who are exceptional in educational technology courses in the educational field, had been selected before the pilot study was conducted (Gay & Air Asian, 2003; Hair et al., 2018). After that, the TPDP questionnaire was distributed to 125 primary teachers. After the results were collected, Rasch Model Measurement Analysis was utilized to measure its validity and reliability.

Methodology

Purposive sampling was used in the study to collect samples from 125 primary school teachers. The questionnaire distributed in this study had been adapted from the study by (Evers et al., 2015). However, in this present study, innovation construct was decided to add in as another construct of TDPD because today's discussions regarding the potential of digital technologies in education increasingly position the topic as part of a broader strategy for educational innovation (OECD, 2016). Despite the fact that educational systems and institutions are not innately resistant to transformation, there appear to be very strong impediments standing in the way of digital technologies' full potential in educational settings and in terms of teaching and learning practises. Transformation requires innovation among teachers to remain revelant so that teachers could be able to adapt learning and societal needs. Thus, we conducted a functional assessment of the reliability, item polarity and fitness items as well as the correlation of the standardized residuals.

Instrumentation

The questionnaire adapted was divided into Part A and Part B.

Demographic Data

Demographic data, such as gender, teaching experience, and educational background, were included in Part A. Each item had a relevance to the respondent's background.

Teacher Digital Professional Development (TDPD)

Part B comprised TDPD-related items that measure digital professional development among teachers. A total of 30 items were formed by six key constructs, which are: (1) latest knowledge (5 items), (2) experimentation (5 items), (3) reflection and feedback (5 items), (4) teaching and learning collaboration (5 items), (5) school improvement collaboration (5 items), and (6) teacher innovation (5 items). Table 1 shows the constructs and items utilized in this study. The 5-point likert points were applied in the current study's responses from respondents: Strongly Disagree (1), Disagree (2), Somewhat Disagree (3), Agree (4), and Strongly Agree (5). Table 2 displays the level of agreement score for Part B.

Table 1
Constructs and Items in Part B

Construct	Item No.	Total Item
Latest Knowledge	1,2,3,4,5	5
Experimentation	6,7,8,9,10	5
Reflection And Feedback	11,12,13,14,15	5
Teaching And Learning Collaboration	16,17,18,19,20	5
School Improvement Collaboration	21,22,23,24,25	5
Teacher Innovation	26,27,28,29,30	5

Table 2
Level of Agreement Score

Strongly Disagree	Disagree	Somewhat Disagree	Agree	Strongly Agree
1	2	3	4	5

Findings and Discussion

This study employed a five-point Likert scale for the TDPD. Before the item analysis was done, person
analysis was performed and the result revealed that 16 students needed to be dropped while 109 others were retained. The results will be discussed accordingly with the six spectific types of Rasch Measurement Model Analysis that we decided.

Item Fit

Each item should contribute in a meaningful way to the construct or concept being tested (Bond & Fox, 2015). Researchers must monitor the value of this index to determine whether the item developed is appropriate (item fit) to measure a latent variable or construct (Nurulhuda et al., 2018). Item polarity and infit scores (MNSQ and ZSTD) were obtained to measure item fit. Infit values check was employed as the instrument's Cronbach's alpha values alone did not assure the instrument' ability alone (Che Lah & Tasir, 2018). Table 3 reveals the item fit analysis related to the MNSQ and ZSTD scores along with the infit scores. Items can be kept if their infit score (MNSQ) is between 0.4 and 1.5 (Linacre, 2005), and their infit score (ZSTD) should be between -2 and +2. (Bond & Fox, 2007). Therefore, according to the results displayed in Table 3, entry numbers 24, referring to items Q24, needed to be dropped due to non-adherence to the criteria stated.

Infit (MNSQ) and Infit (ZSTD) Values

Table 3.
Item Fit

ENTRY NUMBER	INFIT (MNSQ)	INFIT (ZSTD)	ITEM	RESULT
4	1.06	0.50	Q4	Retained
24	1.48	3.00	Q24	Dropped
10	1.10	0.80	Q10	Retained
18	1.32	2.00	Q18	Retained
30	1.25	1.60	Q30	Retained
2	1.18	1.40	Q2	Retained
11	1.24	1.60	Q11	Retained
1	1.07	0.60	Q1	Retained
26	0.99	0.00	Q26	Retained
20	1.07	0.60	Q20	Retained
25	1.10	0.70	Q25	Retained
22	0.86	-1.10	Q22	Retained
3	1.07	0.50	Q3	Retained
28	0.92	-0.05	Q28	Retained
19	1.00	0.10	Q19	Retained
9	0.90	-0.70	Q9	Retained
15	1.01	0.10	Q15	Retained
7	0.99	0.00	Q7	Retained
14	0.92	-0.50	Q14	Retained
13	0.97	-0.20	Q13	Retained
12	0.96	-0.20	Q12	Retained
17	0.90	-0.60	Q17	Retained
27	0.89	-0.70	Q27	Retained
6	0.87	-0.70	Q6	Retained
5	0.82	-1.30	Q5	Retained
21	0.82	-1.30	Q21	Retained
16	0.82	-1.30	Q16	Retained
23	0.82	-1.40	Q23	Retained
8	0.77	-1.60	Q8	Retained
29	0.76	-1.90	Q29	Retained

Item Polarity

Item fit can be determined based on the polarity of the item by calculating the PT MEASURE CORR value. If all of the items in the collection measure the same construct, then this value refers to a collection of those items (Bond & Fox, 2007). Each item in the TDPD has been subjected to item polarity check, where the Point Measure Correlation value must be positive. The PT MEASURE CORR value achieved in this study was between 0.51 and 0.76, which was within the minimal value of 0.3 (Wu & Adam, 2007).With item polarity, this type of analysis ensures that each measured item matches its target. Any element that contains a negative value requires that the corresponding element be deleted. Based on the analysis, the positive PTMEASURE CORR score indicated that the retained items could contribute to the instrument's psychometric features, thus allowing all items to be retained. Table 4 shows the results of the item measurement.

Table 4

Item Polarity

ENTRY NUMBER	PT MEASURE CORR	ITEM	RESULT
18	0.51	Q18	Retained
2	0.54	Q2	Retained
1	0.56	Q1	Retained
4	0.56	Q4	Retained
10	0.58	Q10	Retained
30	0.60	Q30	Retained
3	0.60	Q3	Retained
11	0.61	Q11	Retained
20	0.64	Q20	Retained
7	0.64	Q7	Retained
26	0.66	Q26	Retained
15	0.66	Q15	Retained
19	0.66	Q19	Retained
13	0.66	Q13	Retained
9	0.67	Q9	Retained
12	0.67	Q12	Retained
17	0.68	Q17	Retained
6	0.68	Q6	Retained
25	0.68	Q25	Retained
5	0.68	Q5	Retained
28	0.69	Q28	Retained
22	0.71	Q22	Retained
27	0.71	Q27	Retained
16	0.71	Q16	Retained
14	0.72	Q14	Retained
23	0.72	Q23	Retained
8	0.74	Q8	Retained
21	0.74	Q21	Retained
29	0.76	Q29	Retained

Measure Value

According to Aziz et al (2015), if two or more measure values of the items are the same within the same construct, all except one item need to be dropped. Decision can be appropriately made based on the infit (MNSQ) value close to 1 and the infit (ZSTD) value close to 0 in order to retain them (Bond & Fox, 2015). As a result, all of the items were retained based on the data in Table 5 as there are no redundant items and none of the items in the construct shared the same value of measure.

Table 5

Measure Value

ENTRY NUMBER	MEASURE	ITEM	RESULT
2	0.67	Q2	Retained

19	0.55	Q19	Retained
21	0.48	Q21	Retained
1	0.45	Q1	Retained
25	0.41	Q25	Retained
22	0.39	Q22	Retained
14	0.34	Q14	Retained
29	0.20	Q29	Retained
30	0.15	Q30	Retained
16	0.14	Q16	Retained
17	0.11	Q17	Retained
11	0.03	Q11	Retained
12	0.02	Q12	Retained
15	-0.01	Q15	Retained
20	-0.03	Q20	Retained
13	-0.11	Q13	Retained
23	-0.11	Q23	Retained
28	-0.12	Q28	Retained
9	-0.20	Q9	Retained
5	-0.21	Q5	Retained
26	-0.22	Q26	Retained
27	-0.25	Q27	Retained
18	-0.28	Q18	Retained
7	-0.44	Q7	Retained
10	-0.46	Q10	Retained
8	-0.53	Q8	Retained
3	-0.54	Q3	Retained
4	-0.58	Q4	Retained
6	-0.62	Q6	Retained

Standardized Residual Correlation

According to Aziz et al (2015), the item with correlation value < 0.7 needs to be retained, whereas the item with a value of > 0.7 allows one item to be retained and one item dropped by selecting accordingly using the infit (MNSQ) value near to 1 and infit (ZSTD) value near 0. In Table 6, the values of all correlation were found to be < 0.7, thus all items could be retained.

Table 6

Standardized Residual Correlation

CORRELATION	ENTRY NUMBER	ITEM	ENTRY NUMBER	ITEM	RESULT

0.64	19	Q19	20	Q20	Retained
0.62	12	Q12	13	Q13	Retained
0.61	11	Q11	12	Q12	Retained
0.61	26	Q26	27	Q27	Retained
0.60	16	Q16	17	Q17	Retained
0.59	7	Q7	8	Q8	Retained
0.58	6	Q6	7	Q7	Retained
0.57	28	Q28	29	Q29	Retained
0.57	24	Q24	25	Q25	Retained
0.49	28	Q28	30	Q30	Retained

Unidimensionality

As per Bond and Fox (2015), the principles of unidimensionality require that the analytical techniques include indicators of how closely the persons and items resemble the concept of the ideal unidimensional line. Its inclusion will confirm that the instrument's objectives are accurate and achievable, which typically involves Residual Principal Component Analysis (PCA). Fisher (2007) stated that the range of the raw variance explained by measures of > 40% and unexplained variance in the 1st contrast of < 15% are good and acceptable average values. In Table 7, the raw variance explained by measures is 48.1%, while the unexplained variance in the 1st contrast is 7.9%. With that, the raw variance explained by measures revealed a good value which is more than 40%, was sufficient in comparison to the minimal value (Linacre, 2012) and the unexplained variance in the 1st contrast is also a high and good value which was < 15%. It has been shown that the TDPD contains items that are reliable indicators of teachers' digital professional development.

Table 7
Residual PCA

RAW VARIANCE EXPLAINED BY MEASURES	UNEXPLAINED VARIANCE IN 1st CONTRAST
48.1%	7.9%

Summary Statistics

In Table 8, the results indicated that the value of Cronbach's Alpha, α, is 0.84, reflecting the instrument in good range in terms of internal consistency (Nunnally & Bernstein, 1994) based on the interpretation of the internal consistency of the Cronbach's Alpha in Table 9. The value of person reliability obtained was 0.94, whereas the value of item reliability was 0.81; both indicate a good range as per Fisher's recommendation (2007) of 0.81-0.90 for both person and item reliability, and which is an acceptable range (Pallant, 2001; Sekaran, 2003). Separation values of person and item were 4.10 and 2.04 within the acceptable range (Bond & Fox, 2007; Linacre, 2005). Overall, it is a good indication of TDPD's effectiveness to be used in future research.

Table 8
Summary Statistics

Summary Statistic	Value Obtained
Cronbach Alpha	0.96

Person Reliability	0.94
Person Separation	4.10
Item Reliability	0.81
Item Separation	2.04

Table 9
Internal Consistency of Cronbach's Alpha

Cronbach's Alpha	Internal Consistency
< 0.5	Unacceptable
0.5-0.6	Poor
0.6-0.7	Questionable
0.7-0.8	Acceptable
0.8-0.9	Good
0.9-1.0	Excellent

Limitations and Suggestions

The current research had some limitations, which also gave directions for further research. The study's primary limitation is that it was limited to primary school teachers in Perak. Second, caution is advised when applying this instrument to other situations, and more testing with samples from other cultural groups is necessary. Additionally, when extending this instrument to other contexts, it is necessary to investigate differential item functioning in order to draw relevant comparisons. Third, the type of Rasch Model Analysis that was excluded in this study, the validity of structure calibration and scale review. In future research, we will conduct this type of analysis on the questionnaire, and it would be fascinating if the sample size was higher from different states.

Conclusion

In conclusion, the Teacher Digital Professional Development (TDPD) questionnaire is strongly reliable and acceptable based on the findings in this study. However, there is one item was needed to be dropped because it was not within the acceptable range of infit values. With that, this TDPD questionnaire can be used in future research in a wide-scale sample size. This research significantly advances the existing knowledge by providing a validated instrument for assessing Teacher Digital Professional Development. Researchers in the educational field can use this high-quality instrument, thereby contributing to a more rigorous and precise examination of digital professional development's impact on primary school education. It also plays a vital role in improving the quality of teaching and learning in the primary education setting.

Reference

Azrilah, A. A., Masodi, M. S., and Azami, Z. (2013). Asas Pengukuran Rasch, Pembentukan Skala & Struktur Pengukuran, Universiti Kebangsaan Malaysia. Penerbit UKM, Universiti Kebangsaan Malaysia, Bangi Selangor First printing.

Bond, T. G., & Fox, C. M. (2015). Applying The Rasch Model : Fundamental Measurement in the Human Sciences. (3rd. Eds.). London: Mahwah, New Jersey.

Bortolotti, S. L. V., Tezza, R., de Andrade, D. F., Bornia, A. C., and de Sousa Jr., A. F. (2013). Relevance and advantages of using the item response theory. Qual Quant., 47:2341–60. https://doi.org/10.1007/s11135-012-9684-5.

Fisher, W. P. J. (2007). Rating scale instrument quality criteria. Rasch Measurement Transcation 21(1): 1095.

Gay, D., & Asian, A. (2003). The Rating Scale Model. In van der L. W. J. K. & H. R. (Eds.), Handbook of Modern Item Response Theory. (p. 67–84.). New-York: Springer-Verlag.

Linacre, J. M. (2002). What do infit, outfit, mean-square and standardize mean? Rasch Measurement Transactions, 16,878.

Kean, J., Brodke, D. S., Biber, J., Gross, P. (2018). An introduction to Item Response Theory and Rasch Analysis of the Eating Assessment Tool (EAT-10). Brain Impair. 19(1), 91-102.