INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN BUSINESS & SOCIAL SCIENCES

Vol 10, Issue 1, (2020) E-ISSN: 2222-6990

The Reliability Study of Personal Wellness Questionnaire (PWQ) to Measure Self-Changes among Malaysian Low-Performing Public Service Officers

Abu Yazid Abu Bakar, Hazaila Hassan & Salleh Amat

Faculty of Education, Universiti Kebangsaan Malaysia

Email: yazid3338@ukm.edu.my

To Link this Article: http://dx.doi.org/10.6007/IJARBSS/v10-i1/6855 DOI:10.6007/IJARBSS/v10-i1/6855

Published Date: 30 January 2020

Abstract

This pilot study aimed to identify the reliability of the Personal Wellness Questionnaire (PWQ) which is used as instrument to measure self-changes among Malaysian low-performing public service officers. This instrument consisted of 75 items divided into six sections; section A was demographics data, whereas sections B, C, D, E and F consisted of five sub-constructs of self-changes namely emotional, psycho-spiritual, social, cognitive, and behavioural adjustment. A total of 30 low-performing public service officers at a particular ministry in Putrajaya were involved in this pilot study. The Rasch Model version 3.72.3 was used to analyse the PWQ items, in which value of 0.89 was obtained for item reliability, and value of 0.95 was obtained for respondent reliability. These findings indicated that PWQ items were very good, in effective condition with a high level of consistency, and can be used in actual research. Several items were dropped because they did not match the correct constructs and did not comply with the criteria set by the researchers. The final instrument comprised of 51 appropriate items for measuring the five self-changes sub-constructs of the research target population.

Keywords: Personal Wellness Questionnaire (PWQ), Pilot Study, Public Service, Malaysia

Introduction

Civil servants in Malaysia faced various issues in terms of human development which leads to a low level of commitment among some of them. Nowadays, current changes in society such as higher income and living rates, highly educated societies, and diverse customer demands, subsequently urging the public sector to provide better quality services in terms of broader options and flexibility (Marsidi & Abdul, 2007). Therefore, it was important for counseling services to be established in the workplace. Bakar (2014) stated that among the core goals of counseling services is to encourage changes in client behavior, help client make decisions, form clients' coping skills, rationalize client's minds and help clients improve relationships with others. Circular Letter No. 4/1998 had been issued by the Malaysian Public Service

Department stating that psychological and counseling intervention services were highly emphasized and given much attention in order to improve the service quality of the public service officer. Therefore, the need to implement this intervention in the workplace requires support and involvement of management at all levels.

Emotional stability, psycho-spiritual, social skills, cognitive and behavioral adjustment, if unbalanced, could affect the quality of service of an employee. Therefore, they needed to be improved to enhance their work performance (Bokti & Talib, 2010; Tenney, Poole & Diener, 2016; Milliman et al., 2000; Querstret et al., 2015). In Malaysian Public Service Department (PSD), self-changes of low-performing civil servants in these five elements were measured using the Personal Wellbeing Questionnaire (PWQ). This adapted instrument, however, had never been validated for its' reliability to be used in local context. For that reason, the main objective of study was to test the reliability of this questionnaire in order to see the suitability and to detect any weaknesses in items used. Through this validation study, the researcher performs the functionality check on the items as a whole and each individual item from the aspect of reliability.

Methodology

This pilot study aimed to obtain the reliability of the instruments. There were 75 items in this instrument that were divided into six sections, namely section A for demographic data which contained nine items and sections B, C, D, E and F which were further divided into 5 sub-constructs of self-change, which were emotional stability, psycho-spiritual, social skills, cognitive and behavioral adjustments. The instrument used was a questionnaire adapted by researchers from Psychology Management Division, Public Service Department. Thirty people involved were participants of the Personal Wellbeing Program organized by a ministry in Putrajaya in which the respondents had the same characteristics as the actual respondents chosen by the researcher that were those with Annual Performance Score Report of 60% and below.

The Rasch Model approach is used to determine the reliability of an instrument. In this pilot study, the researchers used the Rasch Measurement Model to test the reliability of items and respondents and for the removal of inappropriate items in the study. However, for this paper, Rasch's model measurement approach was also used to examine the reliability of questionnaire instrument developed through quantitative data collection in the pilot study. Normally, the reliability of an item was only seen through Alpha Cronbach (α) value for the entire instrument.

Results and Discussion

A total of 30 respondents answered this questionnaire, those who were involved in the Personal Wellbeing Program conducted for three days and two nights, similar to the actual program which would be conducted for 20 hours. After the data were collected, the data were analyzed descriptively and the minimum value used in this analysis was the Rasch Measurement Model approach, researchers perform item functionality checks in term of reliability and item-respondents differentiation and removal of items. The explanation for each item functionality check was described in Table 1 as follows:

Table 1
Interpretation of Alpha-Cronbach (α) Scores (Bond & Fox 2007)

Alpha-Cronbach(α) Score	Reliability
0.9 – 1.0	Very good and effective with high degree of consistency
0.7 – 0.8	Good and acceptable
0.6 – 0.7	Acceptable
< 0.6	Item need to be repaired
< 0.5	Item needs to be removed

In order to determine item reliability for instruments, Rasch measurement model approach was used by referring to the reliability and differentiation of items. The findings of the analysis showed that the reliability value obtained based on Alpha Cronbach (α) value was 0.95 as shown in Table 2. This clearly demonstrated that the instruments were very good and effective with a high level of consistency and thus could be used in the actual research.

Table 2
The Reliability Value (Alpha Cronbach (α)) for the Pilot Study

PERSON RAW SCORE-TO-MEASURE CORRELATION = 1.00
CRONBACH ALPHA (KR-20) PERSON RAW SCORE RELIABILITY = 0.95

An analysis of the instrument was also performed on the whole by looking at the reliability and differentiation of items and respondents. Table 3 showed the reliability and differentiation of items in which the item's reliability value was 0.89, while the item separation value was 2.78 when rounded-up became 3.0. Based on item reliability, the value of 0.87 indicated that it was in good condition and acceptable (Bond & Fox 2007). Whereas the separation value of the item was 2.62 and if rounded up, it was equal to 3.0. According to Linacre (2005), the value of good separation index was greater than 2.0.

Table 3
Reliability and Differentiation Value of Items for the entire Instrument Constructs

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	125.6	30.0	.00	.40	1.00	.0	.97	-.1
S.D	7.4	.0	1.19	.02	.27	1.0	.34	1.1
MAX.	143.0	30.0	2.27	.47	1.47	1.6	1.74	2.2
MIN.	110.0	30.0	-2.96	.34	.52	-.23	.39	-.23
REAL RMSE	.42 TRUE SD		1.11 SEPARATION 2.62		ITEM		RELIABILITY .87	
MODEL RMSE	.40 TRUE SD		1.12 SEPARATION 2.78		ITEM		RELIABILITY .89	
S.E OF ITEM MEAN .17								

Meanwhile, based on Table 4, the reliability value of the respondent was 0.95 and the respondent's separation value was 4.15. This showed the reliability of the respondents was very high and it was good because Bond and Fox (2007) stated that the confidence value exceeded 0.80 was good and strong. While the separation value of the respondents showed a good value for the degree of difficulty of the item, which corresponded to the statement of Linacre (2005) which considered the separation value exceeding 2.0 was a good value.

Table 4
Reliability and Differentiation Value of Respondents for the Overall Instrument Constructs

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	213.5	51.0	2.38	.31	.99	-.2	.97	-.3
S.D	14.4	.0	1.32	.02	.48	2.1	.50	1.9
MAX.	244.0	51.0	5.39	.37	2.25	3.8	2.31	3.7
MIN.	175.0	51.0	-.68	.24	.36	-3.4	.33	-.3.3
REAL RMSE	.33 TRUE SD		1.27 3.81	SEPARATION	PERSON		RELIABILITY .94	
MODEL RMSE	.31 TRUE SD		1.28 4.15	SEPARATION	PERSON		RELIABILITY .95	
S.E OF ITEM MEAN .24								

The Point Measure Correlation (PTMEA CORR) value is meant to detect the polarity of the item was intended to test the extent to which construction of the constructs achieved its goals. If the value found in the PTMEA CORR part was a positive (+) value, it indicated that the

item measured the constructs as intended (Bond & Fox, 2007). Conversely, if the value was negative (-) the developed item did not measure the constructs as intended. Therefore, the item needed to be removed or revised as the item did not point to the question or was difficult to answer by the respondent. Based on Table 5, there were three items that had negative values of B1, E58 and F65. For the rest, the PTMEA CORR value was positive and it showed that the item measured the constructs you want to measure. Thus, there were three items needed to be removed from the entire 75 items in the questionnaire (PWQ). While the value of PTMEA CORR was positive, there were five lowest positive values for B2 (0.05), B10 (0.05), D33 (0.06), F62 (0.04) and F69 (0.05). This value should also be noted because it was likely that the item was difficult to answer by the respondent (Azman, 2011). Therefore, the items needed to be revised. The findings showed that positive items in the questionnaire were moving in one direction with constructs and able to measure constructs and did not conflict with the constructs to be measured. If the value of PTMEA CORR was high, then the item was able to differentiate the ability between respondents who answered this questionnaire.

Table 5
Point Measure Correlation (PTMEA CORR) Value

Entry Number	***Point Measure Corr***	***Item***	***Entry Number***	***Point Measure Corr***	***Item***	***Entry Number***	***Point Measure Corr***	***Item***
1	-0.2	B1	26	0.69	C26	51	0.74	E51
2	0.05	B2	27	0.32	C27	52	0.73	E52
3	0.41	B3	28	0.33	C28	53	0.71	E53
4	0.12	B4	29	0.17	C29	54	0.52	E54
5	0.39	B5	30	0.52	C30	55	0.75	E55
6	0.48	B6	31	0.19	D31	56	0.72	E56
7	0.3	B7	32	0.37	D32	57	0.69	E57
8	0.61	B8	33	0.06	D33	58	-0.14	E58
9	0.41	B9	34	0.45	D34	59	0.12	E59
10	0.05	B10	35	0.38	D35	60	0.09	E60
11	0.42	B11	36	0.53	D36	61	0.32	F61
12	0.36	B12	37	0.63	D37	62	0.04	F62
13	0.22	B13	38	0.67	D38	63	0.4	F63
14	0.38	B14	39	0.48	D39	64	0.27	F64
15	0.52	B15	40	0.64	D40	65	-0.01	F65
16	0.53	C16	41	0.59	D41	66	0.14	F66
17	0.51	C17	42	0.66	D42	67	0.22	F67

18	0.51	C18	43	0.56	D43	68	0.26	F68
19	0.48	C19	44	0.59	D44	69	0.05	F69
20	0.7	C20	45	0.62	D45	70	0.3	F70
21	0.43	C21	46	0.6	E46	71	0.22	F71
22	0.68	C22	47	0.71	E47	72	0.21	F72
23	0.38	C23	48	0.77	E48	73	0.64	F73
24	0.68	C24	49	0.68	E49	74	0.38	F74
25	0.61	C25	50	0.59	E50	75	0.43	F75

In addition, the suitability (fit) of items in measuring constructs could also be seen through the values of MNSQ infit and MNSQ outfit. MNSQ's outfit and infit value should be within a range of 0.6 to 1.4 in order to ensure the built items were suitable for measuring the constructs. The MNSQ value should be between 0.6 and 1.4, if the logit value exceeds 1.4 it means the item was misleading and needed to be viewed again. If MNSQ value was less than 0.6, this means that the item was too easily expected by the respondent (Linacre & Ph, 2014). In addition, the value of ZSTD outfit and infit should be between -2 and +2 (Bond & Fox, 2007), however, if the value of the MNSQ outfit and infit was acceptable, then the ZSTD index might be ignored (Linacre & Ph, 2014; Abazeed, 2018). Table 6 showed the misfit order which displayed items that had MNSQ highest and MNSQ lowest values from the statistical item analysis of misfit order.

Based on Table 6, there were 27 items that were within the prescribed range and they needed to be revised or removed. Items exceeding the value of 1.40 in the MNSQ outfit were A10 (3.43), A1 (3.30), A2 (2.82), B18 (2.49), D58 (2.61), E62 (1.76), D59 (1.54), E68 (1.58), A15 (1.48), A5 (1.47), E64 (1.51), E66 (1.45), A9 (1.48), B28 (1.47) and D54 (1.41). Conversely, the value less than 0.6 were D47 (0.49), E73 (0.59), B25 (0.57), C42 (0.57), D53 (0.55), D51 (0.54), B20 (0.52), C43 (0.51), D50 (0.50), E63 (0.49), C40 (0.42), B26 (0.35), C44 (0.44), C45 (0.41), D55 (0.41), and C36 (0.31).

Therefore, with reference to Table 6, a total of 38 items needed to be revised or removed. There were eight items that were not within the PTMEA CORR range. There were 16 items removed because they did not accurately measure the constructs. In addition, 14 items had been revised by looking at the needs of the researchers and expert views. After the analysis, 51 items fulfilled the purpose of constructs to be investigated by researchers.

Table 6
Item Fit Based on MNSQ Value

Entry Number	***INFIT***		***OUT FIT***		***Point Measure Corr.***	***Items***
	MNSQ	*ZSTD*	*MNSQ*	*ZSTD*		
1	2.45	3.2	3.3	4.5	-0.2	B1
2	2.16	4.3	2.82	5	0.05	B2
3	0.8	-0.5	0.84	-0.4	0.41	B3

4	1.18	0.6	1.25	0.8	0.12	B4
5	1.51	1.7	1.47	1.6	0.39	B5
6	0.89	-0.2	0.89	-0.2	0.48	B6
7	0.92	-0.1	1.05	0.3	0.3	B7
8	0.98	0.1	1.06	0.3	0.61	B8
9	1.42	1.2	1.48	1.3	0.41	B9
10	2.7	3.5	3.43	4.5	0.05	B10
11	0.73	-0.7	0.74	-0.7	0.42	B11
12	0.86	-0.3	0.89	-0.2	0.36	B12
13	0.98	0.1	1.02	0.2	0.22	B13
14	1.04	0.3	1.09	0.5	0.38	B14
15	1.58	1.9	1.08	0.4	0.52	B15
16	1.19	0.9	1.08	0.4	0.53	C16
17	0.78	-1	0.67	-1.1	0.51	C17
18	2.68	5.2	2.49	4.6	0.51	C18
19	0.81	-0.9	0.71	-1	0.48	C19
20	0.53	-2.3	0.52	-2.3	0.7	C20
21	0.84	-0.8	0.8	-0.7	0.43	C21
22	0.67	-1.3	0.65	-1.5	0.68	C22
23	1.31	1.5	1.24	1.1	0.38	C23
24	0.8	-0.5	0.68	-1.1	0.68	C24
25	0.58	-1.9	0.57	-1.9	0.61	C25
26	0.45	-1.8	0.35	-2.5	0.69	C26
27	1.07	0.3	1.17	0.6	0.32	C27
28	1.32	0.9	1.47	1.3	0.33	C28
29	1.14	0.5	1.19	0.6	0.17	C29
30	0.69	-0.8	0.69	-0.9	0.52	C30
31	0.59	-1.2	0.65	-1	0.19	D31
32	0.61	-1.3	0.7	-0.9	0.37	D32
33	0.79	-0.6	0.84	-0.4	0.06	D33
34	0.65	-1.1	0.59	-1.3	0.45	D34
35	0.67	-0.9	0.67	-0.9	0.38	D35
36	0.29	-2.7	0.67	-0.9	0.53	D36
37	0.78	-0.6	0.65	-1.1	0.63	D37
38	0.78	-0.7	0.75	-0.8	0.67	D38
39	0.97	0	0.97	0	0.48	D39
40	0.48	-1.6	0.42	-0.2	0.64	D40
41	0.58	-1.3	0.64	-1.1	0.59	D41
42	0.55	-1.4	0.57	-1.4	0.66	D42
43	0.49	-1.6	0.51	-1.6	0.56	D43
44	0.43	-2	0.44	-2	0.59	D44
45	0.41	-2.1	0.41	-2.2	0.62	D45
46	0.67	0	0.67	-0.9	0.6	E46
47	0.61	-1.2	0.49	-1.8	0.71	E47
48	0.8	-0.6	0.78	-0.7	0.77	E48
49	1.02	0.2	1.03	0.2	0.68	E49

50	0.48	-2.1	0.5	-2.1	0.59	E50
51	0.54	-1.8	0.54	-1.9	0.74	E51
52	1.03	0.2	0.91	-0.2	0.73	E52
53	0.55	-1.5	0.55	-1.6	0.71	E53
54	1.37	1.1	1.41	1.2	0.52	E54
55	0.41	-2.7	0.41	-2.7	0.75	E55
56	0.78	-0.6	0.77	-0.7	0.72	E56
57	0.71	-1	0.71	-1	0.69	E57
58	2.49	3.2	2.61	3.4	-0.14	E58
59	1.64	1.6	1.54	1.4	0.12	E59
60	1.11	0.5	1.06	0.3	0.09	E60
61	0.81	-0.5	0.82	-0.4	0.32	F61
62	1.6	1.5	1.76	1.9	0.04	F62
63	0.49	-1.6	0.49	-1.7	0.4	F63
64	1.5	1.3	1.51	1.4	0.27	F64
65	0.86	-0.4	0.85	-0.4	-0.01	F65
66	1.49	1.9	1.45	1.7	0.14	F66
67	0.93	-0.2	0.92	-0.2	0.22	F67
68	1.63	1.9	1.58	1.8	0.26	F68
69	1.27	1.1	1.35	1.4	0.05	F69
70	1.07	0.3	1.07	0.3	0.3	F70
71	1.11	0.5	1.2	0.8	0.22	F71
72	1.21	0.5	1.21	0.9	0.21	F72
73	0.61	-1.9	0.59	-2	0.64	F73
74	0.87	-0.6	0.86	-0.5	0.38	F74
75	0.84	-0.6	0.71	-0.7	0.43	F75

Once the data was analysed, all items and instruments underwent revisions in order to achieve the validity and reliability standards of the instruments based on the Rasch Measurement Model. Although all the items were analysed by SPSS version 23, however, the instrument was supported and strengthened by using the Rasch Measurement Model in terms of checking the item reliability, respondents' reliability, respondents' differentiation and item differentiation as well as item removal. Based on data analysis conducted, 24 items did not meet the requirements of the analysis that had been determined and needed to be rejected.

When using the Rasch analysis application, the rating scale worked to form a category. This category could be used for multiple choice questions or Likert scales. In this questionnaire, 5-point Likert scales were used:

1. Strongly disagree
2. Disagree
3. Somewhat disagree
4. Agree
5. Strongly agree

Table 7 showed the 5-point Likert scale of the categories according to the sequence of 1 to 5 that were 1, 8, 63 and 28. Therefore, through the table above, the difference in the structure

calibration between the scale and the range was to be 1.4<y<5. For example, 2 to 3 = none, 3 to 4 = 1.89, and 4 to 5 = 2.82. This means that the scale in this questionnaire was understood and can be maintained using 5-point Likert scales.

Table 7
Summary of Category Structure

CATEGORY LABEL	SCORE	OBSERVED COUNT	OBSERVED AVERAGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATION	CATEGORY MEASURE	
2	2	10	1	-1.30	.87	.65	NONE	(-4.02)	2
3	3	122	8	.40	1.13	1.13	-2.82	-1.88	3
4	4	970	63	1.94	1.02	.99	1.93	1.43	4
5	5	428	28	4.01	.93	.81	3.75	(4.86)	5

Conclusion

Rasch technique had greatly impacted the manner in which social science research made use of tests and surveys. The Rasch Model framework offered procedures for constructing and revising social science measurement instruments and documenting measurement properties of instruments (e.g., reliability, construct validity). Rasch technique also enabled researchers to make critical corrections when using raw test score data or survey data. Specifically, Rasch technique allowed nonlinear raw data to be converted to a linear scale, which then could be evaluated through the use of parametric statistical tests. In addition to the examples provided earlier, there were Rasch steps that could be used to investigate additional important instrumentation issues such as step ordering/step disordering, item reliability, person reliability, differential item functioning, and differential test functioning (Boone, 2016; Sadoughi & Hesampour, 2017).

In a nutshell, this study helped to validate the Personal Wellbeing Questionnaire (PWQ) which is used by the Malaysian Public Service Department (PSD) as one of its' counselling psychology measurement tools. It could be concluded that the validity and reliability were an important aspect that should be emphasized in evaluating an instrument whether it was new or adapted before it was used in the field of real research. Based on the analysis of this validation study, this instrument was good in quality and appropriate to be used by psychological officers in ministries, departments or in the private sector to measure the self-change through five sub-constructs namely the emotional stability, psycho-spiritual, social skills, cognitive and behavioural adjustments for low-performing civil service officers.

References

Abazeed, R. A. M. (2018). The Impact of Talent Management on Organizational Commitment of the employees of telecommunication companies in Jordan: The Mediating Role of Employee Work Engagement, International Journal of Academic Research in Accounting, Finance and Management Sciences 8 (4): 153-162

Bakar, A. Y. A. (2014), Perkhidmatan Kaunseling Pintar dan Berbakat, Doctoral Dissertation, Universiti Kebangsaan Malaysia.

Bokti, N. L. M., & Talib, M. A. (2010). Tekanan kerja, motivasi dan kepuasan kerja Tentera Laut Armada Tentera Laut Diraja Malaysia. *Jurnal Kemanusiaan*, 15, 56–57.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences (2nd ed.)* Mahwah, NJ: Lawrence Erlbaum.

Boone, W. J. (2016). Rasch Analysis for Instrument Development : Why , When , and How ?, 1–7. http://doi.org/10.1187/cbe.16-04-0148.

Hasan, A. (2011). Kesahan Dan Kebolehpercayaan Item Penilaian Pembimbing Dalam Pembelajaran Berasaskan Kerja (PBK) Menggunakan Model Pengukuran Rasch. USM, Psychometrics Centre, MIMOS & Malaysian Examination Syndicate, MOE.

Milliman, J., Czaplewski, A. J., & Ferguson, J. (2000). Empirical Assessment Workplace spirituality and employee work attitudes An exploratory e. http://doi.org/10.1108/09534810310484172

Marsidi, A., & Abdul, L. H. (2007). Faktor-faktor yang mempengaruhi komitmen pekerja di organisasi awam. *Jurnal Kemanusiaan*, 56–64.

Querstret, D., Cropley, M., Kruger, P., Heron, R. (2015). Assessing the effect of a Cognitive Behaviour Therapy (CBT) -based workshop on work-related rumination, fatigue, and sleep, *0643*(October). http://doi.org/10.1080/1359432X.2015.1015516.

Tenney, E. R., Poole, J. M., & Diener, E. (2016). Research in Organizational Behavior Does positivity enhance work performance ?: Why , when , and what we don ' t know. *Research in Organizational Behavior*. http://doi.org/10.1016/j.riob.2016.11.002.

Rabi, N. M., Ghazali, N. H. C. M., Rohaizad, N. A. A., & Zulkefli, M. Y. (2018). Readiness of PreService Teacher to teach Student with Special Needs through Inclusive Education Course. International Journal of Academic Research in Progressive Education and Development, 7(4), 200–210.

Sadoughi, M., & Hesampour, F. (2017). The Relationship between Perceived Social Support and Loneliness among University Students. *Multilingual Academic Journal of Education and Social Sciences*, *5*(1), 1–6.