A tale of two construct validation analysis: Rasch model and exploratory factor analysis approach for Three-Factor Eating Questionnaire (TFEQ-R21) among Malaysian male workers

Rosnah Ismail, DrPH*, Mohd Zali Mohd Noh, PhD, Noor Hassim Ismail, MPH*, Azmi Mohd Tamil, MPH***

*Community Health Department, Faculty of Medicine, UKM Medical Center, Kuala Lumpur, **Faculty of Computer Science & Information Technology, Universiti Putra Malaysia, Serdang.

SUMMARY

Introduction: This study aims for construct validation using two approaches, i.e., exploratory factor analysis and Rasch Model.

Methods: A cross sectional of 313 male workers from multiple worksites had completed self-administered Malay translated version of Three-Factor Eating Questionnaire-R21. Data quality was assessed by misfit person criteria, dimensionality, summary statistic, item measure and rating (partial credit) scale followed by exploratory factor analysis and internal consistency reliability assessment.

Results: The dual approaches of construct validation analysis were complement to each other. Rasch analysis supported the theoretical constructs of three eating behaviour dimensions among respondents. In contrary to exploratory factor analysis, it did show presence of a newfound factor (α=0.04) came up from the separation of the cognitive restrain and uncontrolled eating however, the correlation between the two respective sub-factors were fair (r=0.39) and weak (r= -0.08). Both analyses had detected three problematic items but those items were psychometrically fit for used for current study setting. The data had adequate psychometric properties. Cronbach's alpha for cognitive restraint, uncontrolled eating and emotional eating were 0.66, 0.79 and 0.87 respectively. Rating scale quality was conformed to standard criteria.

Conclusion: Malay version TFEQ-R21 with promising psychometric properties and valid measures for eating behaviour dimensions among male workers aged between 20 to 60 years old is now available. Further development should focus on the items in relation to Malaysian cultural adaptation before its use for daily practice in future setting.

KEY WORDS:
Three-Factor Eating Questionnaire-R21, male, psychometrics, factor analysis, workplace

INTRODUCTION

Although factor analysis using IBM Statistical Product and Service Solutions Statistics (IBM SPSS Statistics) and Rasch Model analysis has been long linked to data analysis in various research fields, construct validation using both approaches has rarely been used by medical researchers. The type of approach selected depends on the norm practice, availability of statistical tools in research facility, analysis justification, and presence of expertise in assisting data analysis and interpretation. The uniqueness of this article is to assess construct validity from two approaches for a set of translated version questionnaire. In a great hope, this dual approach may be able to give comprehensive evidence for determining suitability of the items among Malay population. Each approach have unique features to analyse data suited to the analysis objective but the former displayed lack of ability for Item Response Theory to measure underlying traits, such as attribute, proficiency, ability or skill, which are reflected in the endorsed responses to the study questionnaires.[1] This lacking can be complemented by the latter, in addition to handling Likert-style category response, identifying item difficulty and person ability that are intangible to measure.[2]

In comparison to the traditional way of construct validation provided by IBM SPSS Statistics, Winsteps gives more psychometric information on both respondents and items on logits scale ruler (natural log probability of an event).[2] Responses to Likert-style questionnaire items can infer the extent to which a psychological attribute is possessed by the questionnaire respondent, i.e., "How likely a person with attribute A agrees to endorse at an item difficulty D?".[3] Therefore, identifying items with difficulty levels across the range of respondents' ability may heighten the construct explanation psychometrically in the presence of adequate test items and test participants.[2]

In order to describe the dual approaches, Three-Factor Eating Questionnaire R21 (TFEQ-R21) was chosen in this study. The original Three-Factor Eating Questionnaire (TFEQ) known as Stunkard-Messick Eating Questionnaire or Eating Inventory was first constructed in 1985 to measure three dimensions of human eating behaviour in an English population,[4] namely cognitive restraint of eating, disinhibition of control eating, and susceptibility to hunger.[5] TFEQ has been validated for normal adult population, adolescents, and different race of a population.[6] Following construct validation study among

This article was accepted: 4 May 2015
Corresponding Author: Rosnah Ismail, Department of Community Health, Faculty of Medicine, UKM Medical Centre, Jalan Yaacob Latif, Bandar Tun Razak, 56000 Cheras Kuala Lumpur, Malaysia Email: drrose@ppukm.ukm.edu.my

Swedish obese men and women, the original 51-item TFEQ has been revised into 18 items, which comprise cognitive restraint (6 items), uncontrolled eating (9 items), and emotional eating domains (3 items).[5,7] Another three items have been added to the emotional eating scale to avoid floor and ceiling effects.[8] They briefly described the differences between TFEQ-R18 and TFEQ-R21. The revised version has been translated into about 30 languages.[6-11]

Construct validation is always used for validating a new developed questionnaire that has been constructed based on a theory and a hypothesis. Construct validation as stated in the title is misnomer. Construct verification is more accurate name than construct validation. Following systematic translation process of the TFEQ-R21 into Malay version,[12] verification of the construct is another way to detect poor translated items in terms of wording or sentence suitability in a smaller sampled population. So, the test user is able to improve the quality of the problematic items. Therefore, the potential unwanted disastrous consequences are controlled at the very initial stage.[13] Often, adequacy of sample size is debatable. Perhaps, a total of 30 to 50 respondents are comfortable enough to generate item analysis such as item level descriptive statistics, item internal consistency, and factor analysis[14] and able to generate Rasch analysis to illustrate psychometric features of an instrument.[15] A group of researcher did a simulation of exploratory analysis study and found that up to 8 factors recovery can be reliable with sample size below than 50 if the factor loadings are higher than 0.8.[16] However, larger sample size is used in this analysis because the lowest expected factor loading is 0.4. The estimated sample size is 234 respondents if 24 items are to be tested and three factor recoveries are expected.16 In this study, repeated exploratory factor analysis shows the number of factor recovery and the factor loading values are more stable in sample size of 250 and more.

Insight about construct validation

A construct is an attribute, proficiency, ability, or skill that happens in the human brain and is defined by established theories, whereas validity is defined as the degree to which a test measures what it claims to be measuring.[17] Construct validation is an evaluation of a measurement instrument that consists of two fundamental elements, i.e., reliability and validity.[18] Element of validity refers to how well an instrument measures what it is supposed to measure to give a true measure,[19] whereas element of reliability refers to consistency (regardless of correctness).[14] It is about how consistent an instrument is to get the same results if it is used to measure repeatedly to the same person under the same circumstances.[20] An instrument cannot be valid unless it is reliable; however, the reliability of an instrument does not depend on its validity.[18] Any threats to the reliability of a test are also threats to its validity.[17]

Reliability of an instrument can be assessed by Cronbach's alpha, factor analysis and standard error of measurement. Cronbach's alpha is also called as internal consistency reliability is normally measured for each construct[14] when multiple-item measures of a construct are employed. It describes the extent of the items interrelatedness within a construct which should be tested on enough spread of sample variability by giving the average amount of all possible split-halves items correlations.[20] The acceptable value is at least 0.60 up to 0.90.[18,20] High value of Cronbach's alpha does not always mean a high degree of internal consistency. The value is affected by a few factors such as the number of items in a construct and the respondent's characteristics (such as reading ability, proficiency with the items format, familiarity with the content domain under investigation).[18] It is also affected by presence of multi-constructs and styles of responding.[21] Hence, accepting and interpreting an alpha value for each construct must be done carefully. Factor analysis describes homogeneity of a set of items in a test by observing which items tend to clump together and not to the other items, is called a factor.[20] The correlation of an individual item with a factor is called factor loading of at least above +0.30 or below -0.30 to indicate that the item contributes meaningfully to a factor.20 Factor loading of less than 0.2 is considered highly problematic.[14] Lastly, in order to express reliability of a set of measurements standard error of measurement should be obtained. It is calculated from standard deviations of repeated measurements of an attribute for a single person;[19] however, it is not discussed in depth in this article.

Rasch Model overview

This article gently introduces Rasch Model analysis for construct validation based on a philosophy that the recorded endorsements on response categories are reflections of a single underlying construct.[2] The most magnificent feature of the model is the total score of both person and item are mapped side by side on a logits ruler scale (natural log probability of an event) where the midpoint of the scale is zero. Thus the ruler measures respondents' ability to endorse responses to the different hierarchy of items difficulty in different magnitude of agreement at zero. The model also provides adequate psychometric information on person and item reliability. High person reliability means enough spread of ability among persons in the sampled population towards line of inquiry items. High item reliability indicates enough items spread along the continuum in which some items are more difficult and some items are easier to be endorsed accordingly. It provides fit statistics of item or person (based on point-measure correlation, mean square (MNSQ) and Z-standard), unidimensionality and rating scale quality as explained further below.

MATERIALS AND METHODS

This study obtained ethical approval from the Ethics Committee of Universiti Kebangsaan Malaysia (project code FF-288-2012). All participants gave informed consent for participation following short briefing about the study objectives.

Participants

A total of 316 male workers from multi-worksites were recruited purposively from August 2012 to January 2013. They were selected by the executive officers of the respective organisations based on the inclusion criteria, i.e., male workers who had been in service for at least a year, age

between 20 to 60 years old, and free of cardiovascular disease. For those who could not read well, the TFEQ-R21 questionnaires were read to them. If they found the items confusing and difficult, they were allowed to ask the researcher for clarification. Their answers were thoroughly checked to ensure there were no missing responses on any of the items.

Instrument

Translated Three-Factor Eating Questionnaire R21 (TFEQ-R21) in Malay version was used as an instrument to collect data from respondents. The ten steps of translation process as shown in Figure 1 had been described in details somewhere else[12] based on International Society for Pharmacoeconomics and Outcome Research (ISPOR) report.[22] A total of twenty one items were used including cognitive restraint, CR (6 items); uncontrolled eating, UE (9 items); and emotional eating, EE (6 items).[23] Twenty items had four Likert-scale and one vertical scale of 8-point numerical rating scale. Question number 1 to 16 had similar response categories (1=Definitely agree, 2=Mostly agree, 3=Mostly disagree, 4=Definitely disagree) while the remaining four had response categories based on the nature of the questions. Items 1 to 16 were reverse coding. Endorsement item 21 was recorded as 1 for precorded item value 1-2, 2 for precorded item value 3-4, 3 for precorded item value 5-6, and 4 for precorded item value 7-8. Higher scores indicated more uncontrolled, restraint or emotional eating respectively. The uncontrolled eating scale assessed the tendency to lose control over eating when feeling hungry or when exposed to stimuli; cognitive restraint assessed control over food intake to influence body weight and body shape; and emotional eating measured the propensity to overeat in relation to negative mood states, e.g., when feeling lonely, anxious, or depressed.[8]

Analysis

a. Rasch Model analysis

Raw SPSS data of TFEQ R21 were converted into Winsteps construct that was fitted to the Rasch model using Winsteps software version 3.72.3. The data were first cleaned based on misfit person diagnosis. The data were diagnosed based on 1) point measure correlation negative value, 2) outfit MNSQ>2, and 3) Z-standard value>2. Those persons who fulfilled all three misfit criteria were removed from the analysis. Each construct was examined for summary statistics, dimensionality, good rating scale and item fit criteria. The objectives were to obtain statistical information about the tested items and the persons affirming the items, and to test how well the observed item fit the expectations of the measurement model.

For summary statistics, Cronbach's alpha, item reliability, item separation, person reliability, person separation (at least value of 1.5 indicates the construct separates respondents into two distinct hierarchical strata), and standard error of item were observed and evaluated. Unidimensionality of each construct was considered violated if 1) the unexplained variance in first contrast was more than 2 Eigenvalue, 2) meaningful pattern (i.e. items that are clustered together vertically with loadings substantially greater than zero) in the residuals factor plot after extracting "Rasch Factors", and 3) the construct displayed Differential Item Functioning (DIF) for waist circumference groups (1=waist circumference ≥102 cm, 2=waist circumference <102 cm). The criteria for negligible DIF analysis were both t-value between -2 and +2 and DIF contrast between -0.50 and +0.50. Each theoretical construct of TFEQ-R21 was analysed separately because overall TFEQ-R21 was found having multidimensionality issue.

Quality of rating scale categories were examined for (a) minimum number of ≥10 responses per category, (b) the category frequencies displayed regular distributions (i.e. uniform, normal, bimodal or slightly skewed), (c) average measures increased monotonically across the rating scale, (d) adjacent threshold distance between 1.4 and 5 logits, (d) distinct probability curve graph on each response category, and (e) outfit MNSQ was less than 2. Outfit MNSQ refers to outlier-sensitive fit which is more sensitive to responses to items with difficulty far from a person, and vice-versa.[15]

Lastly, the item was deemed fit the Rasch model if the point measure correlation was between 0.4 and 0.85, the outfit MNSQ was between 0.5 and 1.5 logits, and the Z-standard was between -2 and +2.[15] The point-measure correlation is a summary of a relationship between both discrete variables with more than two equally spaced values (e.g., Likert-scale).[24] Those misfitting or overfitting items were not removed because the data were analysed from an existing test.

b. Exploratory factor analysis by using SPSS

Exploratory factor analysis and reliability analysis were replicated suggestion of Naing.[14] All TFEQ R-21 items were moved to variables column of Factor Analysis menu. Exploratory factor analysis was extracted using Principle Component Analysis and Varimax rotation with Eigenvalues more than 1; the Factor Loading less than 0.40 was suppressed for analysis. Internal consistency reliability estimation in this study was computed separately for each recovered factor. Analysis was performed with IBM SPSS Statistics version 21.0.

RESULTS

Respondents

A total of 3 respondents were removed from the analysis because they demonstrated misfit pattern at all constructs based on Rasch analysis. Those misfit respondents were comparable and not significantly different to fit respondents in terms of age, service duration, body mass index, and waist circumference. About 54% of the 313 workers were from service sectors while the remainder were from manufacturing sectors. The number of workers was in the range of 3 to 15 per organisations. Majority of the respondents (80.8%) were Malay, had tertiary school education (50.0%), had income between RM2001 and RM4000 (45.0%), and were married (59.7%). The median age (IQR) of the respondents was 32 (26; 38) with range of age between 20 and 58 years old. The median (IQR) waist circumference was 88 (80; 97) with range of between 65 and 164 cm. Only 14.7% of the respondents had waist circumference of 102 cm and above.

Rasch factor analysis

From the Rasch analysis, overall it was found that the TFEQ-R21 was a multidimensional questionnaire with the strength of 4 items, as shown in the Table I. Therefore each construct

Table I: Rasch Model Analysis: Construct validation

			Three-Factor Eating Questionnaire R21 (UPMEAN = 0)		
		Overall (n=313)	Cognitive restraint (n=306)	Uncontrolled eating (n=304)	Emotional eating (n=307)
Summary statistics	Cronbach's α	0.76	0.66	0.79	0.87
	Item reliability	0.99	0.97	0.97	0.93
	Item separation	9.28	5.30	6.12	3.67
	Person reliability	0.72	0.61	0.75	0.75
	Person separation	1.61	1.25	1.73	1.74
	Standard error	0.19	0.24	0.22	0.15
Dimensionality	Measured variance, %	36.6	42.2	41.7	51.0
	Eigenvalue	4	1.7	1.6	1.6
	Meaningful pattern of residuals	Yes	No	No	No
	DIF to waist circumference	No	No	No	No
Rating scale	≥10 responses per category		Yes	Yes	Yes
	Regular distributions		Yes	Yes	Yes
	Average measure increase monotonically		Yes	No	No
			(1.05; 1.07; 1.28)	(1.44; 1.26; 0.74)	(2.67; 2.48; 0.55)
	Distance adjacent threshold		1.95; 2.45	2.93; 1.82	5.23; 2.20
	Outfit MNSQ <2		Yes	Yes	Yes
	Probability curves		Distinct 4 peaks	Distinct 4 peaks	Distinct 4 peaks

Table II: Rasch Factor Analysis (n=306)

Items	Correlation	Misfit type	Factor loadings
Cognitive restraint (n=306)			
CR1: Choose small helpings to control weight	0.74	Fit	0.64
CR3: Consciously restrict eating during meals to avoid gaining weight	0.71	Fit	0.57
CR2: Don't eat some foods because they make me fat	0.70	Fit	0.46
CR6: Restraint eating	0.62	Fit	-0.43
CR5: Probable effort to eat less than what is wanted	0.51	Fit	-0.46
CR4: "Stocking up" frequency on tempting food	0.24	Fit	-0.61
Uncontrolled eating (n=304)			
UE5: When I smell appetizing food or see a delicious dish, I find it very difficult not to eat - even if I have just finished a meal	0.67	Fit	0.05
UE2: Being with someone who is eating, often makes me want to also eat	0.65	Fit	0.16
UE8: Eating binges even though not hungry	0.47	Fit	0.56
UE9: Frequency of hungry related to meal time	0.45	Fit	0.70
UE7: When I see something that looks very delicious, I often get so hungry that I have to eat right away	0.71	Fit	-0.15
UE4: I am always hungry so it's hard for me to stop eating before I finish the food on my plate	0.67	Fit	-0.59
UE6: Hungry enough to eat at any time	0.67	Fit	-0.28
UE3: Often get so hungry - stomach feels like a bottomless pit	0.59	Fit	-0.47
UE1: Seem can't stop when start eating	0.55	Fit	-0.33
Emotional eating (n=307)			
EE3: Eat when feel tense or "wound up"	0.78	Fit	0.47
EE2: Eat too much when feel sad	0.77	Fit	0.61
EE1: Eat when feel anxious	0.76	Fit	0.41
EE6: Eat when feel depressed	0.79	Fit	-0.60
EE4: Eat when feel lonely	0.76	Fit	-0.36
EE5: Eat when feel nervous	0.74	Fit	-0.59

All items are not bias to waist circumference measurement

Table III: TFEQ-R21 Exploratory Factor Analysis (n=313)

Item	Mean (SD)	Factor 1	Factor 2	Factor 3	ITC
CR1	2.78 (0.80)	0.80			0.57
CR2	2.38 (0.83)	0.75			0.48
CR3	2.61 (0.81)	0.80			0.52
CR4	3.03 (0.56)	-0.85			0.06
CR5	2.64 (0.76)	0.44			0.29
CR6	2.68 (0.81)	0.47			0.37
UE1	1.90 (0.75)		0.42		0.43
UE2	2.47 (0.79)		0.58		0.50
UE3	1.79 (0.67)		0.42		0.47
UE4	2.00 (0.75)		0.73		0.56
UE5	2.26 (0.77)		0.69		0.55
UE6	1.95 (0.65)		0.67		0.58
UE7	2.25 (0.78)		0.71		0.60
UE8	2.16 (0.84)		0.53		0.30
UE9	1.81 (0.72)		0.82		0.30
EE1	1.93 (0.73)			0.77	0.65
EE2	1.63 (0.64)			0.81	0.70
EE3	1.82 (0.69)			0.73	0.68
EE4	1.83 (0.67)			0.70	0.63
EE5	1.77 (0.59)			0.69	0.61
EE6	1.74 (0.61)			0.78	0.72
Alpha		0.66	0.79	0.87	

in TFEQ-R21, i.e., cognitive restraint eating, uncontrolled eating, and emotional eating, was analysed separately. Summary statistics, dimensionality, and rating scale quality are also shown in the Table 1. The data had enough spread of item difficulty and person eating behaviour variability. Significant difference of mean and its standard deviation (in logits) for each construct was well appreciated, as shown in Figure 2. The variable map indicates that emotional eating construct was the most difficult item followed by uncontrolled eating and cognitive restraint eating. Rasch factor analyses for the constructs are shown in Table II.

Exploratory factor analysis

The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.86 and Bartlett's test of spherity was significant at $p<0.0001$, indicating the suitability of this data for factor analysis procedures. Initial factor analysis indicated a five-factor solution. All factors held Eigenvalues greater than 1 whilst this five-factor solution accounted for 56.8% of the total variance. The newfound factor came up from the separation of the cognitive restrain (i.e., CR4) and uncontrolled eating (i.e., UE8 and UE9). However, the correlation between the newfound factor and its expected domains was weak ($r = 0.17$ and $r = -0.09$ respectively). Cronbach's alpha was 0.04 for all the three problematic items; the value increased to 0.33 if the item CR4 was deleted. The number of factor recovery was close to the expected constructs if all three problematic items were removed, but the weakness of item UE3 was thus highlighted. Cronbach's alpha for cognitive restraint, uncontrolled eating and emotional eating were 0.66, 0.79 and 0.87 respectively. Emotional eating construct was found the most stable construct compared to the two constructs. Results are shown in Table III.

DISCUSSION

This study is one of its kinds to assess construct validity from two approaches using Winstep and IBM SPSS. The dual approaches were shown complement to each other. Rasch analysis supported the theoretical constructs of the TFEQ-R21 in surveyed population but exploratory factor analysis showed presence of newfound factors from cognitive restraint and uncontrolled eating constructs. However, the similar three problematic items were detected by both analyses, indicating that they may not well describe the cultural experiential concept of Malaysian male eating behaviour. The analyses revealed that 18-item questionnaire was better adapted to the population and had adequate psychometric properties in measuring three dimensions of eating behaviour among the surveyed male Malaysian workers. The rating scale quality was conformed to the standard criteria because respondents were forced to answer on four categories response of agreement or disagreement certainty. These types of categories are suitable for measuring eating behaviour to obtain more reliable data in the absence of "unsure or not applicable or do not know" response category in the middle response options of agreement or disagreement. This is because there are no logical reasons that the respondents could not answer it or the item is not applicable to them.[25]

Both analyses showed TFEQ-R21 consisted of three main different constructs. However, Rasch analysis offered more information in regard to appreciation of eating behaviour differences in the sampled respondents. It demonstrated that each construct was found to have two further sub-constructs in which those sub-constructs were highly interrelated to measure global construct of cognitive restraint, uncontrolled eating, and emotional eating, respectively. For cognitive restraint construct, three positive items (CR1, CR3, and CR2) appeared to have common meaning of control over food intake to influence body weight. In contrast, the remaining three negative items (CR6, CR5, and CR4) appeared to have

common meaning of control over food intake cognitively. For uncontrolled eating construct, four positive items (UE5, UE2, UE8 and UE9) appeared to have common meaning of tendency to lose control when exposed to stimuli. In contrast, the five negative items (UE7, UE4, UE6, UE3, and UE1) appeared to have common meaning of tendency to lose control when feeling hungry, other study found similar.[5] For emotional eating construct, three positive items (EE3, EE2, and EE1) appeared to have common meaning of propensity to overeat in relation to negative mood states due to social situation (performance situations and interpersonal interaction). In contrast, the three negative items (EE6, EE4, and EE5) appeared to have common meaning of propensity to overeat in relation to self-negative mood states. Interestingly, the emotional eating construct was the most stable in both exploratory factor analysis and Rasch factor analysis, yet was found the most difficult construct to be endorsed. Perhaps, male gender singularity may explain the findings but may not confirm it. Previous studies showed that emotional eating construct was significantly more pronounced among women compared to the other constructs[5,26] and had the highest factor loadings of all the three constructs.[9]

Item CR4 "How frequently do you avoid stocking up on tempting foods?" was also found to have weakness in United States and Canada population[26] and Greek population.[6] In Malaysia, the item was frequently commented by respondents for the keyword of "stocking up on tempting food". One frequent flying executive gave a remarkable comment. He said, "The item is not suitable for Malaysian multiracial population because presence of variety delicacies

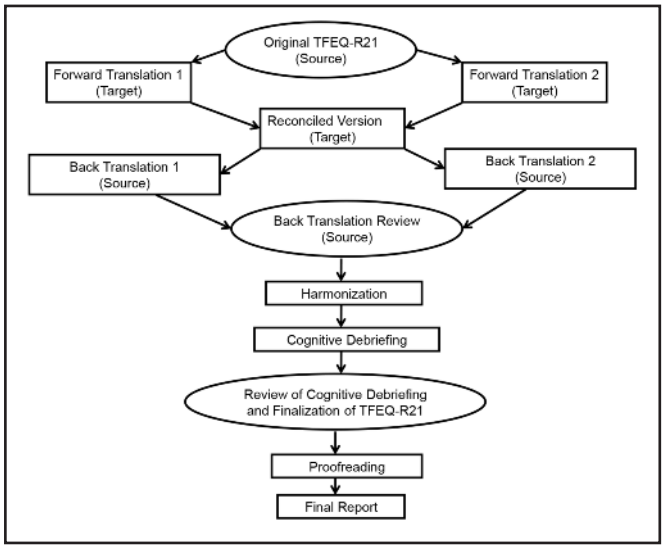

Fig. 1: Translation process.

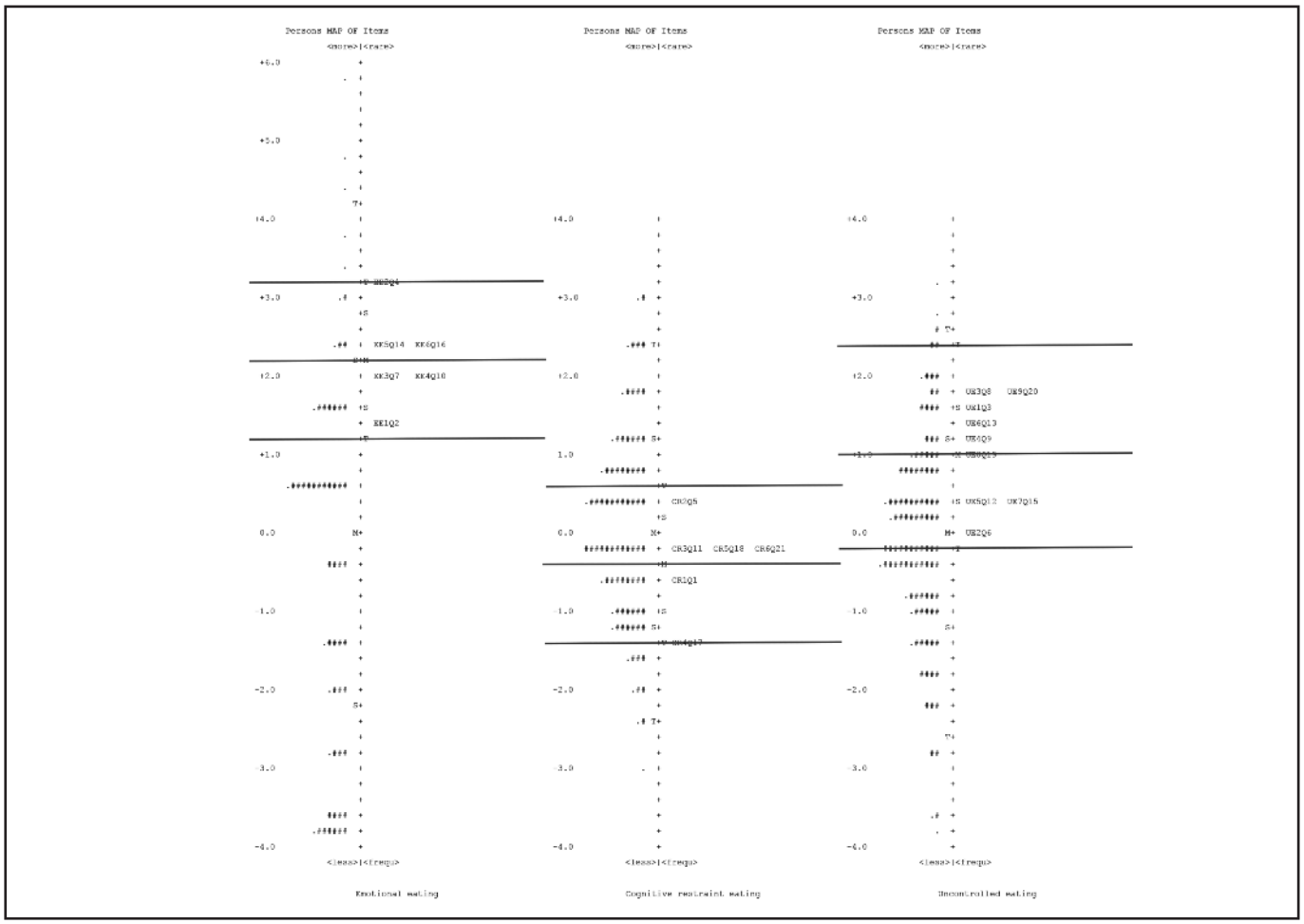

Fig. 2: Three-Factor Eating Questionnaire (TFEQ-R21) variable maps comparison.

which are easily accessible and readily available at any time of the day." He later added, "Malaysians rarely stock up those foods." The comment reflects the true situation of urbanisation-induced lifestyle changes in big cities of Peninsular Malaysia that result in mushrooming variety of food premises and cooked food services nearby staying area, working places, and shopping units/complexes, leading to changes of eating-out behavioural pattern among urbanites.[27] The keyword "tempting food" also had aroused different perception and understanding among the respondents. Some respondents perceived the tempting food was the basic food like staple food rice. Rice is normally stocked up for a month supply for majority households. Others stock up foods like chicken, meat, and fish for two weeks supply. These feedbacks were concordance with previous research.[28] They found rice is the first in home list for food expenditure and the trend of food expenditure changes from rice to meat, fish, vegetables, oils and fats and, fruits with increase of income per capita. The feedbacks were further supported by Malaysian Adult Nutrition Survey that revealed the top-three daily consumptions were 2.5 plates of rice (97.2%), 1.5 medium-size marine fish (40.8%), and 1 cup of green leafy vegetables (39.9%), while bean vegetables, chicken egg, and chicken meat were the top-three consumptions weekly.[29]

Mushrooming variety of food premises and cooked food services nearby staying area, working places, and shopping units/complexes may have encouraged the respondents to enjoy their food to fulfil their biological need with circle of friends leisurely during rest hour or after working hour. Recent study concludes that eating behaviour is no longer associated with three main meal times, i.e., breakfast, lunch, and dinner.[27] Morning tea break, afternoon tea break, and supper have been seen as additional meal times at individual, family, and work unit levels. Such situation may promote eating binges even though not hungry. Both items of UE8 and UE9 may need to be improved in future to capture such eating habit in Malaysian context. Item UE9 "How often do you feel hungry" was also found to have weakness in Greek population.[6]

The number of respondents from various organisations was the strength in this study. The sample size of 313 was sufficient for this validation analysis as a minimum of 234 respondents is recommended for exploratory factor analysis with expected 3-factor recovery and factor loading of 0.4 and above.[16] The high response rate of 99% may minimise the risk of non-response bias. The respondents were obtained from various organisations, which gave enough spread of eating behaviour variability as expected. Researcher was able to capture any remarkable comments from the respondents and to read their body languages while they were answering the TFEQ-R21 questionnaire. These inputs were very helpful to explain the pattern of data. Furthermore, the TFEQ-R21, which was arranged as second set of questionnaires, gave the golden opportunity to the respondents to answer with fresh mind. TFEQ-R21 was also not biased to waist circumference as an indicator for obesity ($r=0.93$, $p<0.01$), other study found similar in which the mean domain scores were comparable and not significantly different between obese and non-obese male population.[26]

The purposive convenient sampling of the organisations to capture male respondents may be regarded as a weakness in many instances. However, eating behaviour is a latent trait that is influenced by multiple factors such as internal factors (e.g., genetic) and external factors (including environmental factors such as availability and accessibility of food products).[27] Absence of heterogeneous gender may limit understanding on emotional eating construct, but it provides an opportunity to understand more pertaining to eating behaviour pattern among males that may relate to the development of chronic diseases.

CONCLUSION

A Malay version of TFEQ-R21 measuring the latent variable of human eating behaviour is now available. Our findings showed that the TFEQ-R21 has promising psychometric properties and valid measure of the tendencies of cognitive restraint, uncontrolled eating, and emotional eating dimensions in Malaysia, at least among male workers of between 20 and 60 years old. However, further development should focus on the items for Malaysian cultural adaptation before its use in future setting. The TFEQ-R21 is a handy instrument to be used in similar study setting or validated in other desired groups of population especially in larger-scale self-administered study, in which respondent's burden and cost of administration are considered.

ACKNOWLEDGEMENT

This study is a part of doctorate research which is supported by Peruntukan Dana Fundamental PPUKM (3200070001) Project Code FF-288-2012. Special thanks to those involved directly and indirectly in the research.

REFERENCES

1. IBM. Item Response Theory/Rasch Models in SPSS Statistics, 2013. http://www-01.ibm.com/support/docview.wss?uid=swg21488442 [25 March 2013]
2. Bond TG, Fox CM. Applying the Rasch Model: Fundamental measurement in the human sciences (2nd ed.). New York: Routledge, 2012.
3. Sick J. Rasch measurement in language education part 3: the family of Rasch models. JALT Testing & Evaluation SIG Newsletter 2009; 13(1): 4-10.
4. Stunkard AJ, Messick S. The three eating factor questionnaire to measure dietary restraint, disinhibition and hunger. J Psychometric Res 1985; 29(1): 71-83.
5. Karlsson J, Persson LO, Sjostrom L, Sullivan M. Psychometric properties and factor structure of the Three-Factor Eating Questionnaire (TFEQ) in obese men and women. Results from the study Swedish Obese Subjects (SOS) study. Int J Obes 2000; 24: 1715-25.
6. Kavazidou E, Proios M, Liolios I, Doganis G, Petrou K, Tsatsoulis A, Tsiligiroglou-Fachantidou A. Structure validity of the Three Factor Eating Questionnaire-R18 in Greek population. J Hum Sport Exerc 2012; 7(1): 218-26.
7. De Lauzon B, Romon M, Deschamps V, *et al.* The Three-Factor Eating Questionnaire-R18 is able to distinguish among different eating patterns in a general population. J Nutr 2004; 134: 2372-80.
8. Tholin S, Rasmussen F, Tynelius P, Karlsson J. Genetic and environmental influences on eating behavior: the Swedish Young Male Twins Study. Am J Clin Nutr 2005; 81(3): 564-9.
9. Angel S, Engblom J, Erikson T, *et al.* Three factor eating questionnaire-R18 as a measure of cognitive restraint, uncontrolled eating and emotional eating in a sample of young Finnish females. Int J Behav Nutr Phys Act 2009; 6: 41.
10. Chearskul S, Pummoung S, Vongsaiyat S, Janyachailert P, Phattharayuttawat S. Thai version of Three-Factor Eating Questionnaire. Appetite 2010; 54(2): 410-3.

11. Natacci LC, Ferreira MJ. The Three Factor Eating Questionnaire-R21: translation and administration to Brazillian women. Rev Nutr 2011; 24(3): 383-94.
12. Rosnah I, Noor Hassim I, Shafizah AS. Translation and Cultural Adaptation Process for Three-Factor Eating Questionnaire (TFEQ-R21). Med J Malaysia 2013; 68(5): 424-34
13. Thabane L, Jinhui Ma, Rong Chu, *et al.* A tutorial on pilot studies: the what, why and how. Med Res Met 2010; 10:1
14. Naing L. Reliability analysis. First UKMMC Intermediate to advance biostatistics in Medical Sciences Workshop. Anjuran Jabatan Kesihatan Masyarakat, Fakulti Pendidikan, Universiti Kebangsaan Malaysia. Cheras, Kuala Lumpur, 2010
15. Azrilah Abdul Aziz. Rasch model fundamentals: scale construct and measurement structure. Malaysia: Integrated Advance Planning Sdn Bhd, 2011.
16. De Winter JCF, Dodou D, Wieringa PA. Exploratory factor analysis with small sample sizes. Multivariate Behavioural Research 2009; 44(2): 147-81
17. Brown JD. What is construct validity? JALT testing & Evaluation SIG Newsletter 2000; 4(2): 8-12
18. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Edu 2011; 2: 53-5.
19. Friis RH, Sellers TA. Chapter 11: Screening for disease in the community. Epidemiology for Public Health Practice. United States of America: Jones and Bartlett Publishers, 2009: 409-36.
20. Aron A, Aron EN, Coups EJ. Statistics for psychology (5th ed.). USA: Pearson Prentice Hall, 2009.
21. Hyot WT, McCullough ME. Issues in the multimodial measurement of forgiveness. In: Worthington EL (eds). Handbook of Forgiveness. USA: Taylor & Francis Group, 2005: 109-23.
22. Wild D, Grove A, Martin M, *et al.* Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. Value in Health 2005; 8(2): 94-104.
23. Karlsson J. Personal communication via email dated 13 June 2012.
24. Linacre JM. Correlation coefficients: describing relationships. Rasch Measurement 2005; 19(3): 1028-129.
25. Dolnicar S. Asking Good Survey Questions. J Travel Res 2013;52(5): 551-74.
26. Cappelleri JC, Bushmakin AG, Gerber RA, *et al.* Psychometric analysis of the Three-Factor Eating Questionnaire-R21: results from a large diverse sample of obese and non-obese participants. Int J Obes 2009; 33: 611-20
27. Noraziah Ali & Mohd Azlan Abdullah. The food consumption and eating behavior of Malaysia urbanites: Issues and concerns. Mal J Soc Space 2012; 8(6): 157-65.
28. Yeong-Sheng Tey, Mad Nasir Shamsudin, Zainalabidin Mohamed, Amin Mahir Abdullah, Alias Radam. Emerging food demand behaviors in Malaysia: Incorporating quality effects in demand analyses, 2008. http://mpra.ub.unimuenchen.de/14832/1/MPRA_paper_14832.pdf [25 July 2014]
29. Norimah AK, Safiah M, Jamal K, *et al.* Food consumption survey patterns: Findings from the Malaysia Adult Nutrition Survey. Mal J Nutr 2008; 14(1): 25-39.