Measuring the Content Validity of MEPI using Content Validity Ratio

Mengukur Kesahan Kandungan MEPI menggunakan Nisbah Kesahan Kandungan

Norashady Mohd Noor[1], Muhammad Aziz Shah Mohamed Arip[2], Eftah Moh @ Hj Abdullah[3], Haniza Yon[4]

[1]*Universiti Pendidikan Sultan Idris; norashady@gmail.com*
[2]*Universiti Pendidikan Sultan Idris; aziz.shah@fppm.upsi.edu.my*
[3]*Universiti Pendidikan Sultan Idris; eftah.a@fppm.upsi.edu.my*
[4]*Mimos Berhad; haniza.yon@mimos.my*

Abstract

The main intention of this research is to examine the content validity of Marine Engineer Personality Inventory or MEPI for student candidates of Diploma in Marine Engineering course at Malaysian Polytechnics by using Content Validity Ratio (CVR). The assessment was conducted through the evaluation among 14 subject matter experts (SME) selected via judgment sampling. Seven professional university experts involved with the expertise in psychology, psychometric, educational measurement, and linguistics. The field experts specializing in particular fields of study consisted of seven practitioners who worked in the polytechnics field and maritime industry. The instrument involved 288 items with six main constructs. The results of the research show that the instrument has a good content validity and proved that MEPI has great potential to be promoted as a good measurement instrument of personality screening for Marine Engineering student selection process. It is recommended to apply more sophisticated statistical analysis, such as the Item Response Theory (IRT) model for elaborating on quality items.

Keywords: Content validity, personality screening, content validity ratio, expert panel, marine engineering students.

Abstrak

Tujuan utama kajijan ini adalah untuk menguji kesahan kandungan bagi Inventori Personaliti Jurutera Marin atau MEPI untuk calon pelajar Diploma Kejuruteraan Marin di politeknik Malaysia dengan menggunakan Nisbah Kesahan Kandungan (CVR). Penilaian telah dilaksanakan melalui pengujian oleh 14 pakar rujuk kursus (SME) yang dipilih melalui persampelan penghakiman. Tujuh pakar profesional universiti terlibat merupakan pakar dalam psikologi, psikometrik, pengukuran pendidikan dan bahasa. Pakar lapangan khususnya dalam bidang pengajian terdiri dari tujuh pengamal yang bekerja dalam bidang politeknik dan industri marin. Instrumen melibatkan 288 item dengan enam konstruk utama. Hasil kajian menunjukkan instrument berkenaan mempunyai kesahan kandungan yang baik dan membuktikan MEPI mempunyai potensi besar sebagai instrument pengukuran yang baik untuk tapisan personaliti untuk proses pemilihan pelajar Kejuruteraan Marin. Ia adalah disyorkan untuk dilaksanakan analisis statistic yang lebih canggih seperti Model Teori Maklumbalas Item (IRT) untuk menjelaskan item yang lebih berkualiti.

Kata kunci: Kesahan kandungan, tapisan personaliti, nisbah kesahan kandungan, panel pakar, pelajar kejuruteraan marin.

INTRODUCTION

Measuring and reporting on the content validity of an instrument or personality inventory is the essence of this research. In the context of psychometrics, content validity refers to the extent to which a measure represents or cover all facets of a given social construct in MEPI. Content validity also functions by determining how well the dimensions and elements of a concept in personality screening can be successfully defined (Sekaran, 2011). The function is to validate every item in the instrument representing each measured construct (Miller, 2013).

The validity of an instrument defines the extent to which it reflects or able to measure the construct being examined (Grove, 2013). The more evidence of content validity such as the expert evaluation is obtained, the higher the confidence of the researcher in the validity of the instruments being constructed (Johnson, 2012). For this research context, expert's evaluation is very important to be part of this research flow. In summary, consensus of the experts is the key factor for this study.

THEORITICAL FOUNDATION

The content validity of MEPI for polytechnic students' candidate who apply for Marine Engineering course is measured by quantitative measurement procedures by Lawshe (1975), which is the Content Validity Ratio or CVR. The CVR is an item statistic that is useful in the rejection or retention of specific items. After items have been identified for inclusion in the final form, the content validity index (CVI) is computed for the whole test. The CVI is simply the mean of the CVR values of the retained items.

CVR used for measuring the content validity items through empirical measurements. CVR is a method from the classical measurement literatures, which is more practical from the aspect of time and costs, besides being easy to administer and fast in implementing (Dewi Rooslani Tojib, 2006). These advantages have made CVR a choice among past researchers abroad (Allahyari, 2011; Baheiraei, 2013; Van Rensburg, 2011) and research in Malaysia (Mohd Arif Shuib, 2013).

CONCEPTUAL FRAMEWORK

MEPI been developed based on the combination of Big Five Personality Theory and Workplace Personality Theory. Therefore, it requires an expert in psychology to verify the items that are included in these instruments. Experts in the field of shipping and marine engineering lines are compulsory to include the elements of the working environment on board into every item. It includes also an expert in the field of industrial and organizational psychologist to make sure the item of personality and workplace match together.
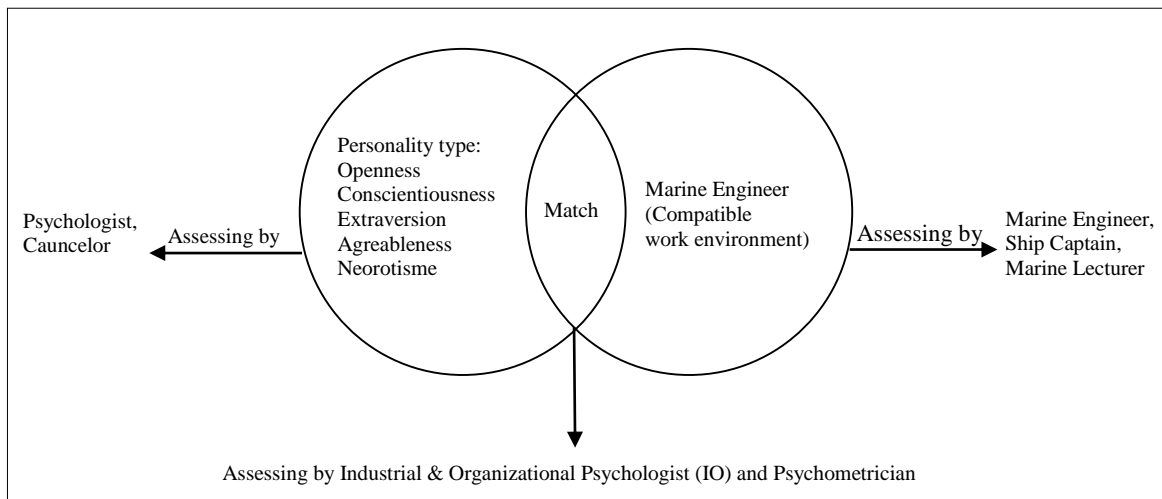
Figure 1: Conceptual framework of content validity assessment.

The Big Five Personality are broad categories of personality traits. While there is a significant body of literature supporting this five-factor model of personality, researchers don't always agree on the exact labels of each dimension. However, these five categories are usually described as follows:

i. *Extraversion*: This trait includes characteristics such as excitability, sociability, talkativeness, assertiveness, and high amount of emotional expressiveness.
ii. *Conscientiousness*: Common features of this dimension include high levels of thoughtfulness, with good impulse control and goal-directed behaviors. Those high in conscientiousness tend to be organized and mindful of details.
iii. *Agreeableness*: This personality dimension includes attributes such as trustworthiness, altruism, kindness, affection, and other pro-social behaviors.
iv. *Neuroticism*: Individuals high in this trait tend to experience emotional instability, anxiety, moodiness, irritability, and sadness.
v. *Openness*: This trait features characteristics such as imagination and insight, and those high in this trait also tend to have a broad range of interests.

There is another construct in this instrument which is 'Survival' construct. Researchers has developed their own items containing six sub construct. All sub-constructs were obtained from a survey conducted on 80 marine engineers all over Malaysian shipping company and interviews among number of engineers and highly experienced sailor. These six sub-constructs are defined as follow:

i. *Adaptability* / Flexibility: Job requires being open to change (positive or negative) and to considerable variety in the workplace.
ii. *Initiative*: Job requires a willingness to take on responsibilities and challenges.

iii.   *Independence*: Job requires developing one's own ways of doing things, guiding oneself with little or no supervision, and depending on oneself to get things done.
iv.   *Integrity*: Job requires being honest and ethical.
v.   *Persistence*: Job requires persistence in the face of obstacles.
vi.   *Leadership*: Job requires a willingness to lead, take charge, and offer opinions and direction.

Research Methodology

This research has referred to the definition of the word 'expert' as a panel of experts being made up of two categories: professional experts and field experts (Rubio, 2003). Professional experts help determine whether the measurements are well constructed for the psychometric testing (Davis, 1992). The second type is the field experts. The total number of experts comprising in this research is 14 including seven professional experts and seven of field experts covered polytechnic lecturers from marine engineering department and marine engineer who got more than ten years of experience. The sampling technique used was one form purposive sampling, which is judgment sampling.

The criteria for selecting the panel of professional experts and field experts are based on academic qualification and experiences. The researcher also contacts the experts via telephone, letter and emails to explain the purpose of the study, the procedures and seek their approval to participate. Although Lawshe's method only requires at least four members for the panel, the researcher has decided to involve as many experts as possible to increase the value of the model (Lawshe, 1975). The total of 14 experts in this research is exceeding the recommendations from past researchers (Baheiraei, 2013; Delgado-Rico, 2012).

Research Design

To validate the content validity of the constructs, the quantitative approach (Lawshe, 1975; Lewis et al, 2005) was undertaken in the following manner:

i.   First, relevant items from the existing literature on human and workplace personality were identified. This led to the construction of the questions and statement.
ii.   Second, a content evaluation panel, consisting of experts from academia and/or industry who were related to the desired research area, was selected.
iii.   Third, each member of the panel was then provided with the questionnaire formed in step 1. The panel members were requested to respond independently to each item in relation to a particular construct on a three-point scale as mention before.
iv.   Fourth, the responses from the overall panelists were then pooled. This step also included counting responses that indicated 'essential' for each item.
v.   Fifth, the content validity ratio (CVR) for each item was estimated utilising the formula $CVR=(n-N/2)/(N/2)$ (Lawshe 1975), where N is the total number of respondents and n is the frequency count of the number of panelists rating the item as "3=essential".

vi.   Finally, the CVR values obtained for each item were examined for their significance employing the standard table provided by Lawshe (1975). If the estimated CVR value was equal to or above the standard value, then the item was accepted; other-wise it was eliminated. The significance level (standard value) depended upon the number of experts rating the item. The minimum number of experts required to rate each item should be five. The value of CVR ranged from 0 to I (Lawshe, 1975; Lewis et al, 1995).

RESULTS AND DISCUSSION

The demographic profile of the profesional experts (N = 7) shows female (5, 71.4%) dominated male experts (2, 28.5%). The area of expertise covers psychology (2, 28.5%), industrial-organizational psychology (2, 28.5) psychometric (1, 14.8%) and linguistic (1, 14.8%). All of them are from various universities such as University Kebangsaan Malaysia, University Pendidikan Sultan Idris and University Malaysia Terengganu.

For field experts (N=7) distribution shows Male (5, 71.4%) dominated female experts (2, 28.5%). It has three area of field expertise such as marine engineering education (4, 57.1%), industrial-organizational psychology (1, 14.8%) and psychology and counselling (1, 14.8%). The experts including the senior lecturer in the Department of Marine Engineering in Polytechnic of Ungku Omar, psychology and counselling officer from Polytechnic of Sultan Azlan Shah and Industrial Organization officer from Mimos Berhad.

Table 1: Examples of items need purification based on type of experts (N=14).

Item Number	Item	The CVR Category Expert Panel			Item Status
		Profesional (N=7) CVRcrit = 0.741	Field (N=7) CVRcrit = 0.741	Total (N=14) CVRcrit = 0.524	
7.	As a child, I rarely enjoyed games of make believe.	0.429	-0.142	0.143	Purification
12.	I am intrigued by the patterns I find in art and nature	0.429	-0.142	0.143	Purification
28	I often try new and foreign foods	0.429	0.429	0.429	Purification
50	I don't take civic duties like voting very seriously	-0.142	0.429	0.143	Purification

The overall findings showed that only 37 items that are just below the critical value of 0.524. Almost all 37 items that need to be modified because of the word which is not in compliance with Malaysian culture. Further analysis should be carried out to test the content validity via statistical methods such as IRT model. Table 1 shows the examples of four items from 37 items that needs purification based on comparison among experts.

Based on Table 1, examples of four items had purified started with item (Q7, "When I was a child, I rarely enjoyed games of 'pondok-pondok' or 'masak-masak'."), (Q12, "Sometimes I get excited by certain artistic patterns that I found), (Q28, "I like to try new foods or recipes that I had never tried to taste") and (Q50, I don't take civic duties like keeping the environment clean very seriously").Most of the items in the MEPI was adapted from NEO-PI-R, International Personality Item Pool (IPIP Scale) and Workplace Personality Inventory. Therefore, there are some item that do not fit with the socio-cultural in Malaysia and needs to be purified. Q7 were under Fantasy construct which shows clearly that the game of make believe is not quite familiar among student in Malaysia.

The SME have recommended to change the word of 'make believe' to 'pondok-pondok' or 'masak-masak'. Q28 was under Action construct which changed to items that explained more common. This was due not everyone had the opportunity to enjoy a meal from abroad. The changes to items were not limited to only these eighty items, but also other items deemed inappropriate by matching indicator. Thus, after this, these items will go through the pilot study process among the Marine Engineering students at Polytechnic of Ungku Omar, Ipoh, Perak. Items that do not achieve minimum agreement by the expert panel must be either eliminated from the instrumen or revised (DeVon, 2007). In this context, items will be revised by the research as preparation for pilot testing.

Table 2: Item distribution after verifications by expert.

Construct	Subcostruct	Number of items	Number of revised / eleminated item
Openness vs Clossedness to experience	Ideas (curious)	8	1 item revised
	Fantasy (imaginative)	8	2 item revised
	Aestatics (artistic)	8	2 item revised
	Action (wide interest)	8	3 item revised
	Feeling (excitable)	8	1 item revised
	Values (unconventional)	8	
Conscientious vs Lack of direction	Competence (efficient)	8	
	Order (organized)	8	
	Dutifulness (not careless)	8	1 item revised
	Achievement striving (thorough)	8	
	Self-discipline (not lazy)	8	
	Deliberation (not impulsive)	8	
Extraversion vs Introversion	Gregariousness (sociable)	8	1 item revised
	Assertiveness (forceful) Activity	8	1 item revised
	Activity (energetic)	8	1 item revised
	Excitement-seeking (adventurous)	8	1 item revised
	Positive emotions (enthusiastic)	8	1 item revised
	Warmth (outgoing)	8	
Agreeableness vs Antagonism	Trust (forgiving)	8	
	Straightforwardness (not demanding)	8	
	Altruism (warm)	8	1 item revised
	Compliance (not stubborn)	8	
	Modesty (not show-off)	8	
	Tender-mindedness (sympathetic)	8	1 item revised
Neuroticism vs Emotional stability	Anxiety (tense)	8	
	Angry hostility (irritable)	8	1 item revised
	Depression (not contented)	8	
	Self-consciousness (shy)	8	1 item revised

Construct	Subcostruct	Number of items	Number of revised / eleminated item
	Impulsiveness (moody)	8	1 item revised
	Vulnerability (not self-confident)	8	
Survival vs Give up	Adaptability (Flexibility)	8	1 item revised
	Initiative	8	3 item revised
	Independence	8	3 item revised
	Integrity	8	3 item revised
	Persistence	8	5 item revised
	Leadership	8	3 item revised
6 main construct	36 sub construct	288 item	37 item revised

CONCLUSION

In conclusion, a total of only 37 items required refinement thus showing that the items were built with a good operationalization and conceptualization. The strength of CVR was prominent in this study when the differences in expert opinions could be seen clearly and easily. The researcher suggested that all 288 items that were refined would undergo a pilot study by polytechnic students using the IRT model. Through the IRT model, the items were selected after some due consideration such as the appropriateness statistics such as unidimensionality, local independence, item fit, item polarity and differential item functioning in order to fulfil Item Response Theory assumptions.

ACKNOWLEDGEMENTS

RUJUKAN

Allahyari, T., Rangi, N.H., Khosravi, Y., & Zayeri, F. (2011). Development and evaluation of a new questionnaire for rating of cognitive failures at work. *International Journal Occupation. Hygiene* 3(1), pp. 6–11.

Baheiraei, A., Khoori, E., Ahmadi, F., Foroushani, A.R., Ghofranipour, F., & Weiler, R.M. (2013). Psychometric properties of the adolescent health concern inventory: the Persian version. *Journal of Psychiatry* 8(1), pp. 28–36.

Davis, L.L. (1992). Instrument review: Getting the most from a panel of experts. *Application Nursery Research.* pp. 194–197.

Delgado-Rico, E., Carretero-Dios, H. & Ruch, W. (2012). Content-related validity evidence in test development. *International Journal of Clinical and Health Psychology*. 12(3), pp. 449-460.

DeVon, H.A., Block, M.E., Moyle-Wright, P., Ernst, D.M., Hayden, S.J., Lazzara, D.J., Savoy, S.M. & Kostas-Polston, E. (2007). A psychometric toolbox for testing validity and reliability. *Journal of Nursing Scholarship.* 39(2), pp. 155-164.

Dewi Rooslani Tojib & Ly-Fie Sugianto (2006). Content validity of instruments in IS research. *Journal of Information Technology Theory Applications.* 8(3), pp. 31-56.

Grove, S.K., Burns, N., & Gray, J.R. (2013). *The practice of nursing research: Appraisal, synthesis, and generation of evidence*. St Louis: Elsevier Health Sciences.

Johnson, B. & Christensen, L. (2012). *Educational research quantitative, qualitative, and mixed approaches*. Alabama: SAGE Publications Inc.

Lawshe, C.H. (1975). A quantitative approach to content validity. *Journal of Personality and Social Psychology*. 28(4), pp. 563-575.

Lewis, B.R., Templeton, G.F. & Byrd, T.A. (2005). A methodology for construct development in MIS research. *European Journal of Information System*. 14(4), pp. 388-400.

Miller, L.A., Lovler, R.L. & McIntire, S.A. (2013). *Foundations of psychological testing: a practical approach*. California: SAGE Publications Inc.

Mohd Arif Shuib, Shukran Abdul Rahman and Nor Diana Mohd, Mahudin (2013). Assessing psychosocial elements of crowds during Hajj: Scale construction and content validation. Published in *Proceedings of the 5th National Seminar on Hajj Best Practice*.

Rubio, D.M., Berg-Weger, M., Tebb, S.S., Lee, E.S., & Rauch, S. (2003). Objectifying content validity: Conducting a content validity study in social work research. *Social Work Research: Oxford Journals.* 27(2), pp. 94-105.

Sekaran, U. & Bougie, R. (2011). *Research methods for business: A skill building approach*. New Delhi: John Wiley & Sons.

Van Rensburg, H., Basson, J.S. & Carrim, N.M.H. (2011). Human resource management as a profession in South Africa. *Journal of Human Resource Management.* 9(1), pp. 1-16.

Wilson, F.R., Pan, W. & Schumsky, D.A. (2012). Recalculation of the critical values for Lawshe's content validity ratio. *Journal of Measurement and Evaluation in Counseling and Development*. 45(3), pp. 197-210.