Application of Rasch Model in validating the construct of measurement instrument

[1]Azrilah Abd Aziz, [2]Azlinah Mohamed, [3]NoorHabibah Arshad,
[4]Sohaimi Zakaria, [5]Azami Zaharim, [6]Hamza Ahmad Ghulman and [7]Mohd Saidfudin Masodi

Abstract**—** The development of Rasch Measurement Model in social science educational measurement has rapidly expanded to other areas of education including technical and engineering fields. Originally, there was substantial controversy between those who saw Rasch Model as a relevant method of measurement in technical fields and those who saw them as essentially different. This paper is an attempt of a paradigm shift in testing and validating a process towards bio-based Rasch Model. It is believed compatibility exist with the fundamental measurement currently used based on Kuhn's explanation on the role of measurement in physical science particularly in measuring competency which is categorised as latent trait. These cannot be gleaned from textbooks in computer engineering or statistics. Taking the paradigm shift, many technical faculties in Institutions of Higher Learning has embarked on the application of Rasch Model to measure the achievement of it's program Learning Outcomes (LO). Face validity tests were conducted subsequent to rigorous meta-analysis on the attributes identified from literature reviews. The major constraint in face validity test is the very small number of sample that is involved; hence reliability. Rasch Model tabulates these experts' opinion on a Person and Items Distribution Map (PIDM) which gives a summative over view on their Level of Agreement for the attributes duly identified. Comparative analysis against the traditional *t*-test to show the correlation between the experts and the attributes shows that Rasch measurement was found to give a better exploratory depth in understanding the experts level of agreement of an attribute. Despite the small sample size, the experts opinion were clearly defined as to their level of acceptance according to the respective dimension before an attribute can be considered for the development of the survey questionnaires as the research instrument; hence construct validity. This is of utmost importance as a bad construct is detrimental to a research finding.

Keywords**—** Rasch analysis, measurement, face validity, information professionals, competency.

Received April 31, 2008. Revised received August 30, 2008. Supported: Ministry of Science, Technology & Innovation, MALAYSIA under the National Science Fund, Grant No.: UiTM/ 100-PPSis(16/1/2)

A. Azrilah, a PhD (IT) candidate is the main recipient of the Scholarship Grant. Her research on latent trait measurement is extended Rasch Model; a paradigm in technical and engineering field which has lead to new approach in performance evaluation. She can be reached at Tel/SMS : +6019 3332661 Fax : +603 41072262 or e-mail: azrilah@gmail.com

A. Mohamed, Ph.D; Associate Professor in System Science at the Faculty of Information Technology and Quantitative Science Universiti Teknologi MARA, 40800 S.Alam, MALAYSIA and reachable at azlinah@tmsk.uitm.edu.my,

N. Arshad, Ph.D; Associate Professor in Intelligent System at the Faculty of Information Technology and Quantitative Science Universiti Teknologi MARA, 40800 S.Alam, MALAYSIA and reachable at habibah@tmsk.uitm.edu.my

Z. Sohaimi, Ph.D; Associate Professor in Information Retrieval, is the Director of Division of Academic Affairs, Universiti Teknologi MARA 40450 Shah Alam, Selangor, MALAYSIA can be reached at e-mail: sohaimiz@salam.uitm.edu.my

A. Zaharim; Ph.D; Associate Professor, is the HoD of Engineering Mathematics, Universiti Kebangsaan Malaysia 43250 Bangi, Selangor, MALAYSIA can be reached at e-mail: azaminelli@gmail.com

H.A. Ghulman; Ph.D; Vice Dean (Academic), Dept. of Mechanical Engineering, College of Engineering & Islamic Architecture, Umm al-Qura University, P.O. Box 9125 Makkah, Kingdom of Saudi Arabia, and reachable at: hghulman@uqu.edu.sa

M. Saidfudin, is the Quality Management Professional Certificate Program Director at the School of Professional Advancement & Continuing Education, University Technology Malaysia, 81300 Skudai MALAYSIA reachable at e-mail: saidfudin@gmail.com

I. INTRODUCTION

Information Professionals (IP) are professionals who managed and use information for critical decision making in an organization. The existence of these professionals has proven their significance and contributed to the success of the organization in achieving their strategic planning; hence, contributes to the economic growth of an organization. These people are computer engineers, communication engineers, system architects, knowledge navigators in information system, system analyst etc. Thus, it is essential that these IP's acquire the necessary essential knowledge and skills in order for them to be competent in performing their tasks during their study period in IHL.

Conventionally, knowledge and skills are predominantly about acquiring either of cognitive attributes or inclination towards technical perspectives [1-3]. However, now a days it requires more than educational qualities and intelligence to be successful. It also requires more than technical expertise in-order for the professionals to perform their tasks successfully and significantly contribute to the financial success of an organization [4-6].

This study endeavoured into the competency attributes required by Information Professionals in acquiring and sharing significant information in the digital landscape which then uses it for critical decision making. The collection of competency attributes will then be used as basis of constructing the assessment tool in measuring the acquired competencies among Information Professionals, specifically in Malaysian public organization. However, competency measurement is not quitely done accurately [7, 8] despite it's utmost importance and essential aspect in every organisation.

These attributes need to go through a critical phase of face-validity from the local experts to confirm the significance of these supposedly deemed required competencies. This will eventually be used to establish a correlation between foreign and local experts; perceptions on the required competencies for local IP's; $\boldsymbol{\mu}_{IP}M$, againsts foreign IP's universally; $\boldsymbol{\mu}_{IP}F$, on the premise;

$$\mathbf{H_o : \boldsymbol{\mu}_{IP}F = \boldsymbol{\mu}_{IP}M}$$

II. OBJECTIVES

The face-validity test carried out is to confirm on the competency attributes derived subsequent to a rigorous meta-data analysis which has been summarized and consolidated from literature and researches done in the area of information handling and technology practitioners. The list of competencies is tabulated along with their expected outcomes using prescribed keywords that represent the competencies. The study uses the concept applied by the infamous American Board of Engineering and Technology (ABET) Table of Keyword and Expected Outcomes [9, 10]. This serves as the framework in developing the IP Competency Assessment Tool (IPCAT) which focuses on assessing observable behaviors or expected outcomes from the IP's. In psychometry, behaviour is the manifest of affective development and cognitive skills.

These attributes were put to professionals i.e Computer Engineers, System Analyst, Informations Managers etc.who are familiar with the working ambience of an IP. They are deemed experts in their respective fields but have sufficient working contact or experience with IP's or they are an IP themselves and knows well their role and responsibilities in a given scenario. These experts are required to state their agreement or disagreement on the listed competency attributes duly identified and then rank each of them according to their preference of priority; on a rating scale of 1 to 3, where 1 represents the lowest priority whilst 3 represents the highest priority, and a 2 as medium priority. Example of the face-validity test form is shown in Figure 1. The responses will then be tabulated and analysed using Rasch Unidimensional Measurement Model [11]; with the aid of Rasch analysis software [12]; **WinSteps®**.

III. CONCEPT TRANSFORMATION

Responses from the professionals on the face-validity are considered rating scale in which the professionals rated the competencies according to priority. In theory, at this stage the study is only counting the responses of priorities from the experts. The rating is only an order of preference; an ordinal scale which is continuum in nature, and do not have equal intervals which contradicts the nature of numbers for statistical analysis[13]. It does not meet the fundamentals of sufficient statistics for evaluation. In Traditional Test, these data set would normally be put on a scatter plot to establish the best regression. However, prediction from ordinal responses on the competency attributes are almost impossible due to absence of intervals in the scale. The normal solution in linear regression approach is to establish a line which fits the points as best as possible; which is then used to make the required predictions by inter-polation or extra-polation as necessary as shown in Figure 1.

$$y = \beta_0 + \beta_1 m \qquad (1)$$

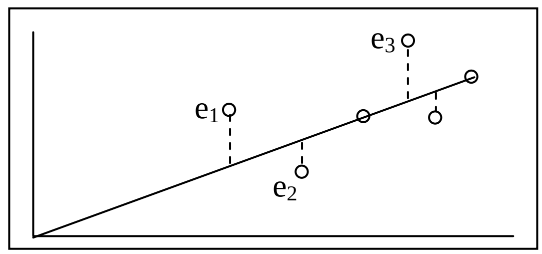


Figure 1 –Best fit line concept

In obtaining the best fit line; however, there exist differences between the actual point; y_i, and the predicted point; $\acute{y}_i$, that is on the best fit line,. The difference is referred here as error; *e*

$$y_i - \acute{y}_i = e_i \qquad (2)$$

By accepting the fact that there is always errors involve in the prediction model, the deterministic model of equation (1) renders itself less reliable. This can be resolved by transforming it into a probabilistic model by including the prediction error into the equation;

$$y = \beta_0 + \beta_1 m + e \qquad (3)$$

Rasch moves the concept of reliability from establishing "best fit line" of the data into producing a reliable repeatable measurement instrument [14] instead. Rasch focuses on constructing the measurement instrument with accuracy rather than fitting the data to suit a measurement model with of errors. By focusing on the reproducibility of the latent trait measurement instead of forcing the expected generation of the same raw score, i.e. the common expectation on repeatability of results being a reliable test, the concept of reliability takes its rightful place in supporting validity rather than being in contentions. Hence; measuring competency in an appropriate way is vital to ensure valid quality information can be generated for meaningful use; by absorbing the error and representing a more accurate prediction based on a probabilistic model.

In Rasch philosophy, the data have to comply with the principles, or in other words the data have to fit the model. In

Rasch point of view, there is no need to describe the data. What is required is to test whether the data allow for measurement on a linear interval scale specifically in a cumulative response process i.e. a positive response to an item stochastically implies a positive response to all items being easy or otherwise. This is dichotomous responses which can take only two values, 0 and 1 which is known as Bernoulli random variable; in our case a competent IP or otherwise.

Rasch Measurement Model is expressed as the ratio of an event being successful as;

$$P(\theta) = \frac{e^{(\beta n - \delta i)}}{1 + e^{(\beta n - \delta i)}} \qquad (4)$$

where;

e = base of natural logarithm or Euler's number; 2.7183

βn = person's ability

δi = item or task difficulty

Rasch exponential expression is a function of Logistic Regression which resulted in a ***S***igmoidal ogive and can be transformed into simpler operation by reducing the indices by logarithm :

$$\ln[P(\theta)] = \ln\left[\frac{e^{(\beta n - \delta i)}}{1 + e^{(\beta n - \delta i)}}\right] \qquad (5)$$

Now ln[P(θ)]; as the probability of a successful event; x=1 is reduced to the expression termed *logit* and can be construed simply as the difference of person ability; β_n and the item difficulty; δ_I, which can be represented as;

$$\ln[P(\theta)] = \beta_n - \delta_i ; \qquad (6)$$

The very reason why the need to transformed it to *logit* is primarily to obtain a linear interval scale. It can be readily shown mathematically that a series of numbers irrespective of based used is not equally spaced but distant apart exponentially as the number gets bigger while a log series maintain their equal separation; thus equal interval [15]. This equal separation is shown in Table 1 and we term it ***logit*** as unit of measurement of ability akin to *meter* to measure length or *kilogram* to weight. The difference between $\log_{10}5$ and $\log_{10}2$ is constant and remain of equal distant between $\log_{10}50$ and $\log_{10}20$. Similarly for $\log_e$; hence *logit*.

Table 1. Comparison of Numerical and Log intervals

Numerical series	**$\log_{10}$**	**$\log_e$**
1	0.000	0.000
2	0.301	0.694
5	0.699	1.609
10	1.000	2.303
20	1.302	2.997
50	1.699	3.912
100	2.000	4.606

IV. Discussion

The face-validity questionnaire was forwarded to experts that were identified from the Directory for Science & Technology Expert 2007/2008 produced by Malaysian Science & Technology Information Centre (MASTIC) under the jurisdiction of the Ministry of Science, Technology and Innovation (MOSTI). MASTIC has identified ten (10) knowledge domains in categorizing the science and technology information. Seven (7) out of sixteen (16) professionals have replied to the test and returned duly completed form, at the time this article was written. Those who have replied are from the following knowledge domains; Business & Economics, Environment & Biodiversity, Information & Communication Technology, Industry, Science & Technology Services, and Social Sciences.

V. Concept Transformation

The responses from the experts are tabulated and run in **WinSteps®** software to obtain the *logit* values. Figure 1 below shows the Person-Item Distribution Map (PIDM) where the person; Experts and the items; the competency attributes are plotted on

the same *logit* scale. By virtue of the same unit scale; then the basic rule of Additivity, the correlation of β_n and δ_i can be established.

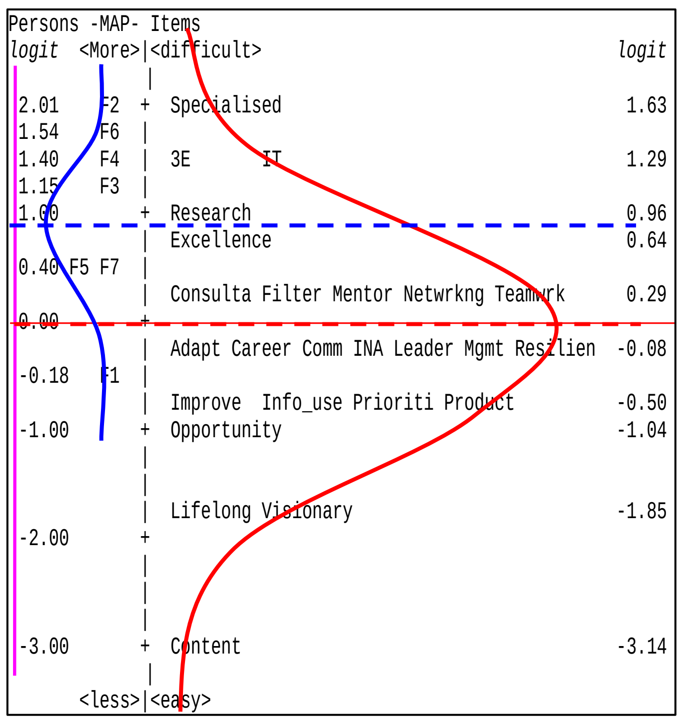


Figure 1. Person-Item Distribution Map

PIDM is similar to traditional histogram tabulation; however in PIDM it allows both the person and the item to be mapped together side-by-side to give a better picture how the person correlate to the respective items. This will give a clearer view of the persons' ability and the relevant item difficulty; From Figure 1, the Face Validity Level of Agreement given by the experts for the IP attributes identified can easily established. By Central Limit Theorem which states that the mean of a sample averages; $\boldsymbol{\mu}_A$, will approximate the mean of the population; $\boldsymbol{\mu}_P$, then the mean of the experts can be taken as representative measure of the experts. Since we are trying to establish the auditors' perception to the attributes; the item mean; $\boldsymbol{\mu}_i$ is pegged at *zero* to serve as the reference mark.

This is clearly shown here; where

$P(\theta) = \beta_v - \delta_i$ (f rom Equ.6)

$= \mathbf{0.96} - 0$ (= β_v values from Figure 1)

$$P(\theta) = \frac{e^{\beta v - \delta i}}{1+e^{\beta v - \delta i}}$$

$$= \frac{e^{0.96}}{1+e^{0.96}}$$

= 0.7230 **OK, p.s.i > 0.7**

Hence, Face Validity Level of Agreement = **72.30%,** is **acceptable.**

The experts indicated their Level of Agreement at 72.30% which is more than the 70% threshold limit of Cronbach Alpha of 0.6. Therefore, all the experts agree to the competency attributes. This can be determined conspicuously from the PIDM where the person mean $\boldsymbol{\mu}_{person}$=+0.96 *logit* is located higher than the item mean; $\boldsymbol{\mu}_{item}$ which is constrained to 0.00 *logit*. This indicates that all the experts involved in the Face-validity test have the tendency of agreeing to the entire competency attributes prescribed.

Expert F2 which has a *logit* of +2.01 being the highest on the PIDM, agrees that Information Professionals should have all the competency attributes in the Face-validity test. Whilst F1 located at -0.18 *logit*, being the lowest, only agrees to 26% (N=8) of the attributes out of the 30 competency attributes, and disputes the rest. It is interesting to note also that F5 and F7 being slightly below the $\boldsymbol{\mu}_{person}$ have reservation on Specialised, 3E, IT, Research and Excellence. On all the competency attributes prescribed in the Face-validity test, all the experts have no dispute and agrees to all the competency attributes that have *logit* of -0.18 below; similar to what experts F1 agrees.

The spread of logit scale from Table 2 of Item Measure, shows that maximum value is at +1.63 *logit* and the minimum value is at -3.14 *logit*. The study refers to the common *logit* scale, since this is the same scale that is used in measuring both the person ability and the item difficulty; comparing both variables on the same interval scale. The difference between *logit*$_{max}$ where *Specialised* is, and the min *logit*$_{min}$ where *Content* located, is δ = 4.77. This indicates that the Item difficulty of the items spread over 4.77 *logit* unit.

Table 2. Item Measure

TABLE 13.1 FACE VALIDITY RESPONSE_LOGIT ZOU158WS.TXT Apr 13 11:04 2008
INPUT: 7 Persons 25 Items MEASURED: 7 Persons 25 Items 3 CATS MINISTEP 3.64.2

Person: REAL SEP.: 1.72 REL.: .75 ... Item: REAL SEP.: .59 REL.: .26

Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	RAW SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEA CORR.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
3	12	7	1.63	.59	1.03	.2	.97	.1	.41	57.1	55.7	Specialised
4	13	7	1.29	.58	.66	-.8	.69	-.6	.53	42.9	49.6	3E
6	13	7	1.29	.58	.74	-.5	.76	-.4	.44	42.9	49.6	IT
9	14	7	.96	.57	1.04	.2	1.03	.2	.45	57.1	48.2	Research
16	15	7	.64	.58	.23	-2.5	.25	-2.4	.60	100.0	48.8	Excellence
2	16	7	.29	.59	.66	-.7	.63	-.7	.75	71.4	51.7	Filter
14	16	7	.29	.59	.93	.0	.87	-.1	.90	42.9	51.7	Consultative
19	16	7	.29	.59	.81	-.3	.77	-.4	.59	42.9	51.7	Teamwrk
20	16	7	.29	.59	1.26	.7	1.24	.6	.13	14.3	51.7	Netwrkng
24	16	7	.29	.59	.73	-.5	.70	-.6	.68	71.4	51.7	Mentor
5	17	7	-.08	.63	1.95	1.7	2.04	1.7	-.24	42.9	59.9	INA
7	17	7	-.08	.63	.94	.0	.88	-.1	.62	57.1	59.9	Mgmt
12	17	7	-.08	.63	1.35	.8	1.17	.5	.34	57.1	59.9	Resilient
18	17	7	-.08	.63	.68	-.6	.73	-.4	.28	57.1	59.9	Career
22	17	7	-.08	.63	1.03	.2	1.11	.4	.49	57.1	59.9	Adaptability
23	17	7	-.08	.63	1.38	.9	1.16	.5	.75	57.1	59.9	Leadershp
25	17	7	-.08	.63	1.38	.9	1.16	.5	.75	57.1	59.9	Communication
8	18	7	-.50	.68	.53	-.9	.53	-.7	.61	71.4	63.0	Info_use
10	18	7	-.50	.68	1.18	.5	1.43	.8	-.28	42.9	63.0	Product
11	18	7	-.50	.68	1.16	.5	1.21	.5	-.18	42.9	63.0	Improve
21	18	7	-.50	.68	.83	-.2	.88	.0	.21	42.9	63.0	Prioritized
15	19	7	-1.04	.79	.75	-.2	.56	-.4	.49	57.1	73.2	Opportunity
13	20	7	-1.85	1.06	1.20	.5	2.05	1.1	-.33	85.7	86.4	Visionary
17	20	7	-1.85	1.06	1.20	.5	2.05	1.1	-.33	85.7	86.4	Lifelong
1	21	7	-3.14	1.85	MINIMUM ESTIMATED MEASURE							Content
MEAN	16.7	7.0	-.13	.71	.99	.0	1.04	.0		56.5	59.5	
S.D.	2.1	.0	1.02	.27	.35	.8	.46	.8		17.7	10.1	

Attributes were check on the Point Measure Correlation with acceptable parameters; PMC = *x*, 0.4 < *x* <0.8. Table 2 shows the following items; i.e. Consultative, Networking, Info. Needs Analysis, Resilient, Career, Product, Improve, Prioritized, Visionary and Lifelong fall outside the range. Rasch requires further verification by looking at the OUTFIT column for Mean Square value; MNSQ = *y*, 0.5 < *y*< 1.5. InfoNeeds Analysis, Visionary, and Lifelong were found beyond this parameter. Further checks on the Z-Std value, where Z-Std = *z*, -2 < *z* < +2; shows none were beyond the set value. Hence, all items are acceptable for further analysis.

The mean raw score obtained is 16.7; giving the average score of ;

$$\text{Mean Average} = \frac{16.7}{\text{7 Professionals x 3 (ideal high score)}}$$

$$= \frac{16.7}{21}$$

$$= 0.79 \approx 80\%$$

Let us look at two of the experts from Table 3 as a focus of discussion in this article. Let us look at F2=2.01 and F1=-0.18, having the highest and lowest score respectively:

For the case of F2; his levels of agreement for the highest item and the lowest item, can be derived as follows:

Table 3. Correlation of experts' Level of Agreement to Information Professionals attributes matrix.

Person	Logit Person Measure	Item	Logit Item measure	P(F2)	P(F6)	P(F4)	P(F3)	P(F5)	P(F6)	P(F1)	Avg item
F2	2.01	Specialised	1.63	0.59	0.48	0.44	0.38	0.23	0.23	0.14	0.36
F6	1.54	3E	1.29	0.67	0.56	0.53	0.47	0.29	0.29	0.19	0.43
F4	1.4	IT	1.29	0.67	0.56	0.53	0.47	0.29	0.29	0.19	0.43
F3	1.15	Research	0.96	0.74	0.64	0.61	0.55	0.36	0.36	0.24	0.50
F5	0.4	Excellence	0.64	0.80	0.71	0.68	0.62	0.44	0.44	0.31	0.57
F6	0.4	Filter	0.29	0.85	0.78	0.75	0.70	0.53	0.53	0.38	0.65
F1	-0.18	Consultative	0.29	0.85	0.78	0.75	0.70	0.53	0.53	0.38	0.65
		Teamwrk	0.29	0.85	0.78	0.75	0.70	0.53	0.53	0.38	0.65
		Netwrkng	0.29	0.85	0.78	0.75	0.70	0.53	0.53	0.38	0.65
		Mentor	0.29	0.85	0.78	0.75	0.70	0.53	0.53	0.38	0.65
		INA	-0.08	0.89	0.83	0.81	0.77	0.62	0.62	0.48	0.72
		Mgmt	-0.08	0.89	0.83	0.81	0.77	0.62	0.62	0.48	0.72
		Resilient	-0.08	0.89	0.83	0.81	0.77	0.62	0.62	0.48	0.72
		Career	-0.08	0.89	0.83	0.81	0.77	0.62	0.62	0.48	0.72
		Adaptability	-0.08	0.89	0.83	0.81	0.77	0.62	0.62	0.48	0.72
		Leadershp	-0.08	0.89	0.83	0.81	0.77	0.62	0.62	0.48	0.72
		Communication	-0.08	0.89	0.83	0.81	0.77	0.62	0.62	0.48	0.72
		Info_use	-0.50	0.92	0.88	0.87	0.84	0.71	0.71	0.58	0.79
		Product	-0.50	0.92	0.88	0.87	0.84	0.71	0.71	0.58	0.79
		Improve	-0.50	0.92	0.88	0.87	0.84	0.71	0.71	0.58	0.79
		Prioritized	-0.50	0.92	0.88	0.87	0.84	0.71	0.71	0.58	0.79
		Opportunity	-1.04	0.95	0.93	0.92	0.90	0.81	0.81	0.70	0.86
		Visionary	-1.85	0.98	0.97	0.96	0.95	0.90	0.90	0.84	0.93
		Lifelong	-1.85	0.98	0.97	0.96	0.95	0.90	0.90	0.84	0.93
		Content	-3.14	0.99	0.99	0.99	0.99	0.97	0.97	0.95	0.98

$P(\theta) = \beta_v(F2) - \delta_i max(\text{Specialised})$ (from Equ.6)
$= 2.01 - 1.63$
$= 0.38$

$$P(\theta) = \frac{e^{\beta v - \delta i}}{1+e^{\beta v - \delta i}}$$

$$= \frac{e^{0.38}}{1+e^{0.38}}$$
$= 0.5934$

$F2_{min}$ % Level of Acceptance = **59.34%**

$P(\theta) = \beta_v(F2) - \delta_i min(\text{Content})$ (from Equ.6)
$= 2.01 - (-3.14)$
$= 5.15$

$$P(\theta) = \frac{e^{\beta v - \delta i}}{1+e^{\beta v - \delta i}}$$

$$= \frac{e^{5.15}}{1+e^{5.15}}$$
$= 0.9942$

$F2_{max}$ % Level of Acceptance = **99.42%**

For the case of F1; his levels of agreement for the highest item and the lowest item, can be derived as follows:

$P(\theta) = \beta_v(F1) - \delta_i max(\text{Specialised})$ (from Equ.6)
$= (-0.18) - 1.63$
$= -1.81$

$$P(\theta) = \frac{e^{\beta v - \delta i}}{1+e^{\beta v - \delta i}}$$

$$= \frac{e^{-1.81}}{1+e^{-1.81}}$$
$= 0.1401$

$F1_{min}$ % Level of Acceptance = **14.01%**

$P(\theta) = \beta_v(F1) - \delta_i min(\text{Content})$ (from Equ.6)
$= (-0.18) - (-3.14)$
$= 2.96$

$$P(\theta) = \frac{e^{\beta v - \delta i}}{1+e^{\beta v - \delta i}}$$

$$= \frac{e^{2.96}}{1+e^{2.96}}$$
$= 0.9510$

$F1_{max}$ % Level of Acceptance = **95.10%**

Table 4 list the summarized findings for comparison between the two expert extremes perception; F2 and F1 on IP attributes.

Professional F2	Professional F1
From the overall probability of agreement at 72%, F2 has the probability of agreement to most of the competency attributes except three (3) items; IT, 3E, Specialised.	Compared to F1 who only agrees to items; Content, Lifelong learning, and Visionary. Whilst disagree with the rest of the competency attributes.
The average probability of agreement is at 72%.	The average probability of agreement is only at 40%. Indicating that the difference of opinion is quite far apart. If trimming of the extreme scores is applied, the change will not be significant.
F2 minimum acceptance is at 59.34% for Specialised.	F1 minimum acceptance is at 14.01%, which gives a larger disparity of 45.33% on diagreement
F2 maximum acceptance is at 99.42% for Content.	Whilst F1 maximum acceptance is at 95.10%, which gives a disparity of minute 4.32% for attributes agreed.
Both the experts agree to attributes; Content, Lifelong learning, and Visionary.	

This are meaningful information which will be developed into what we termed as a 'Descriptive Scale'; which is actually profiling the IP according to their classification as shown in Figure 5.

The overall findings indicate that all the Professionals involved in the Face-validity test agrees to the competency attributes prescribed. Therefore, Information Professionals should possess all the required competency attributes for them to successfully perform their duties.

Now we can establish the corelation between foreign and local experts perceptions on the required competencies for local IP's; μ_{IP}M, againsts foreign IP's universally; μ_{IP}F, is valid and true;

$$\mathbf{H_o : \mu_{IP}F = \mu_{IP}M}$$

There is no difference in the expected competency attributes of an IP in Malaysia as compared to an IP at international level. This shows that Malaysia IP expertise is of international standard.

The tabulation of Level of Agreement by competency attributes reveals that disparity on agreed attributes irrespective of experts, is small amongst the experts. Meanwhile, for disagreed items, the experts shows a large contrast in percentage of Level of Agreement. Even if all the probability of disagreements is summed up, the mean average probability of disagreement stands at 36%. Table 3 points that only expert F5, F6 and F1 contributed to some disagreement.

Further scrutiny reveals that, three of them disagree on the following attributes; Specialised, 3-E and IT.

This reservations can probably be explained that most of us learned about usage of computer only after graduation and it's not to be something fundamental. As an IP, these expert expectation has some validity. There need some flexibility of some sort in terms of their ability and not being focussed only in a particular field only.

A professional needs to be quite encompassing knowledge wise. The 3-E issue; effective, efficient and economic could be due to dependancy on technology and the outside world beyond an IP control. However, it only involves three (3) experts affecting 5.24% (N=11 out of 210 points); then all the competency attributes will be used for the construction of the research instrument; the assessment tool in this research, with confidence interval of 90%. These items were included in the development of the questionaire due to it's importance which is profound in IT ambience.

Rasch enable in-depth analysis of each expert for every competency attributes identified. It allows one to delve further on the level of acceptance for each attributes. Since the PIDM can provide indication on an Information Professionals profiling on the probability of agreement towards the competency attributes; the first group consists of professionals which achieve probability average; $\bar{x}$>0.96 *logit*, while the second group ranges between 0.96 < $\bar{x}$ < -0.18 *logit*, and the third group are those who falls below $\bar{x} \leq$ -0.18.

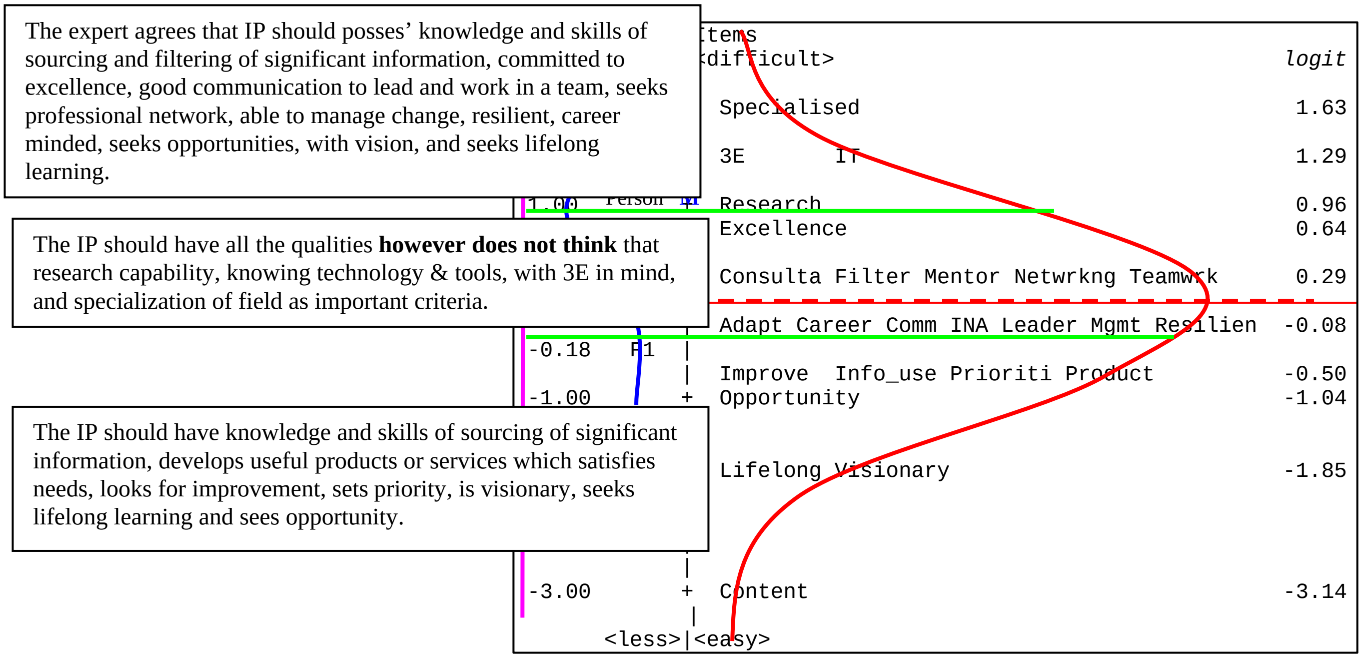
It was found that Rasch Measurement Model is an effective tool in assessing face-validity accurately and fast despite the small sample N. It is a prudent model of measurement which is reliable that generates quality information for meaningful use to construct a valid and reliable instrument. IHL can now conduct a more accurate measurement of the learning ability likelihood achieved by the computer engineering student during their tenure of study and how it contributes to their well being in their career.

VI. VALIDATING THE CONSTRUCT

Subsequent to the findings from the face validity, the study conducted a pilot study to validate the construct of the competency instrument. Pilot study were conducted on a Malaysian government agency and the respondents involved is 23; N_{IP}=23. The respondents are from the officers which is of officers with salary grade of 41 and above, which is equivalent to executive and above in corporate agency.

The responses were tabulated and sorted accordingly to enable analysis using Rasch measurement model for further scrutiny.

Figure 5. Development of Descriptive scale on IP profiles



The ability and the task difficulty measurement is then being plot onto the Person-Item Map for a clearer view of the the correlation between the person and the attributes on the same linear scale called ***logit*** plotted on the PIDM as shown in Figure 6.

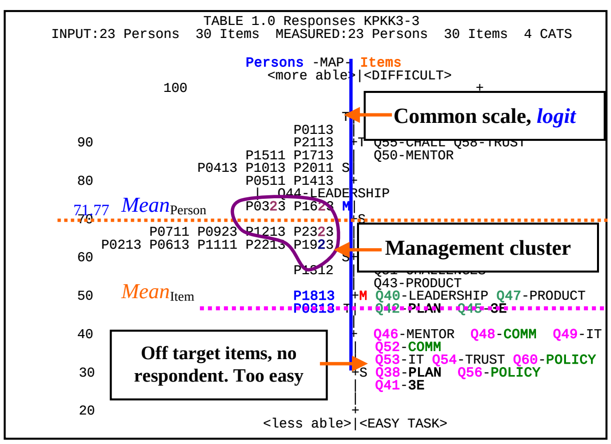


Figure 6. Sample of Person-Item Distribution Map -Wright Map (PIDM)

From Table 5, the reliability provided by the instrument yield a Cronbach-alpha value of 0.72 with high valid responses of 99.8%. This serves as a basis on the construct validity of the instrument which exceed the minimum acceptance Cronbach value of 0.6 at 95% confidence interval; p=0.05. The instrument is fit for purpose; where from Figure 6, it is measuring the ability of the IP. It gives the IP person mean;

TABLE 3.1 Responses KPKK3-3 ZOU814WS.TXT
INPUT: 23 Persons 20 Items MEASURED: 23 Persons 20 Items 4 CATS 1.0.0

SUMMARY OF 23 MEASURED Persons

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	58.0	20.0	71.77	6.09	.94	-.1	.93	-.2
S.D.	3.3	.2	12.64	.66	.61	1.6	.68	1.5
MAX.	65.0	20.0	94.26	6.80	2.21	3.0	2.17	2.5
MIN.	51.0	19.0	45.57	5.04	.10	-2.6	.06	-2.5

REAL RMSE 6.69 ADJ.SD 10.72 SEPARATION 1.60 Person RELIABILITY .72
MODEL RMSE 6.13 ADJ.SD 11.06 SEPARATION 1.81 Person RELIABILITY .77
S.E. OF Person MEAN = 2.70

VALID RESPONSES: 99.8%
Person RAW SCORE-TO-MEASURE CORRELATION = .98
CRONBACH ALPHA (KR-20) Person RAW SCORE RELIABILITY = .72

Table 5. Summary Statistics

$Mean_{Person}$= 71.77 logit, slightly above the 70% mark which is considered the international benchmark as meeting expectation [9, 16]. It also indicates that generally all the 23 respondents meet the 70% competent benchmark. The PIDM revealed that all the respondents from the **Management** group; marked by freeform line, are seen to be locked in between respondents from sub-ordinate officers. The person's description which differentiated the two groups is at the 4th character of the person description field; a "**2**" represent *management* and a "**1**" represent *sub-ordinate officers*. This is certainly a going concern when they are expected to be leading the organisation and therefore exhibit higher competency.

Rasch analysis is able to give a better assessment of a person's ability. It has value adding features which is not available in traditional statistics. The PIDM gives a quick overview but nevertheless accurate estimate of a Person ability; i.e. IP need attention; in this case P1813 and P0813 which obtained a measure of 48.38 and 45.57 logit respectively as shown in Figure 6. Person Measure Order gives the *logit* value of each IP and the corresponding item which is shown in Figure 7. However, the corresponding Point Measure correlation found x is; $0.4 > x > 0.85$; hence, the IP need not be subjected to further scrutiny.

Even if x exceeds the specified value, Rasch require verification by looking at the OUFIT for Mean-Square value MNSQ-y, is in the range of $0.5 > y > 1.5$ whilst for Z-Std value; $-2 > z > +2$. No such finding is observed in this data tabulation for the respective IP's. Figure 6 PIDM also gives an indication that some attributes i.e; Q38, Q41, Q48, Q49, Q52, Q53, Q54, Q56, Q60 are off target. Being in the lower rung, these items are therefore readily achievable by the said IP.

TABLE 17.1 Responses KPKK3-3 **Person STATISTICS MEASURE ORDER**
INPUT:23 Persons 30 Items MEASURED:23 Persons 30 Items 4 CATS

Person: REAL SEP.:1.60 REL.:0.72 Item: REAL SEP.:3.07 REL.:0.90

RAW SCORE	COUNT	LOGIT MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEA CORR.	Prson
65	20	94.26	5.04	1.01	.1	1.28	.9	0.19	P0113
63	20	89.06	5.20	1.78	2.6	2.17	2.5	0.75	P2113
61	20	83.33	5.55	2.21	3.0	2.07	1.9	0.85	P0413
61	20	83.33	5.55	0.60	-1.4	.45	-1.3	0.19	P1013
61	20	83.33	5.55	1.38	1.2	1.99	1.8	0.34	P2011
60	20	80.11	5.81	0.50	-1.6	.33	-1.6	.00	P0511
60	20	80.11	5.81	1.71	1.7	1.19	.5	0.82	P1413
57	20	68.17	6.77	.20	-2.1	.11	-2.1	0.84	P0923
57	20	68.17	6.77	1.68	1.2	1.52	.9	0.27	P1213
57	20	68.17	6.77	.53	-.8	.51	-.7	0.87	P2323
56	20	63.54	6.80	.10	-2.6	.06	-2.5	0.92	P0213
56	20	63.54	6.80	.74	-.3	.90	.1	0.65	P0613
56	20	63.54	6.80	.10	-2.6	.06	-2.5	0.92	P1111
56	20	63.54	6.80	1.04	.3	1.08	.3	0.85	P1923
56	20	63.54	6.80	.10	-2.6	.06	-2.5	0.92	P2213
54	20	55.03	6.15	1.06	.3	.76	-.2	0.65	P1312
52	20	48.38	5.43	1.72	1.9	1.96	1.8	0.43	P1813
51	20	45.57	5.17	.94	-.1	.94	.0	0.60	P0813
MEAN:58.0	20.0	71.77	6.09	.94	-.1	.93	-.2		
S.D. 3.3	0.2	12.64	0.66	.61	1.6	.68	1.5		

Figure 7 Person Measure Order

Generally, the IP in this organisation are very diligent professionals. They are good officers who would complete their tasks no matter how they do it, but still act on basis of directions and regulations. This can be seen from the PIDM where most of the items placements are of positively inclined, no matter how the pattern placement of each of the items.

On the good perspective, they are obedient officers who will complete their task according to the requirements and abiding to regulations and directives from their superiors. However, on the negative perspective, they are not willing or not comfortable to react on their own credibility. This is revealed clearly from the positioning of the management team among other officers on the Person-Item map.

On the contradicting findings; it may be due to their fundamental duty which call them to strictly follow orders from their superiors. They got confused while serving the other Government agencies in their course of duty where the IP has the duty to explore in sufficient depth without being asked to do so. Instead, they have responded otherwise.

This further validate the content of the instrument which reflect the organisation's IP ability in sufficient depth.

VII. Conclusion

The observations explained above indicated that the IP Competency Model developed thus far is capable of "reading" or assessing the competency of the IP and able to profile the professionals according to their behaviour and ability towards the prescribed tasks. A descriptive scale with calculated values hence can therefore be established.

Controls can be more effectively applied with regards to training as well as promotion. The PIDM will conspicously indicate each staff geographical location and help the Management to make prompt decision as to the direction and career development of each staff pretty accurately.

This is only possible by using Rasch Model of Measurement. A computer adaptive test using this algorithm is currently being developed at quite an advanced stage, termed *IPCAT*- acronym for Information Professional Computer Adaptive Test to give *Azrilah Index* on IP competency which is soon available for pilot testing in the participating institutions.

Acknowledgment

The authors wish to express gratitude to the Ministry of Science, Technology and Innovation for their support in providing scholarship: Grant No.: UiTM/100-PPSis(16/1/2) to enable this subject matter to be pursued as a research subject towards the award of Doctor of Philosophy (IT) by University Teknologi MARA, Shah Alam, MALAYSIA.

References

[1] A. Bonder, "What it takes to do the job: the world before and after competencies," 2004.

[2] Bernama, "PSD aims to "go back to basics" move set to transform performance," in *The Star* Kuala Lumpur, 2008, p. N12.

[3] R. Boyatzis, *The competent manager: a model for effective performance*. Chichester: John Wiley, 1982.

[4] C. Hughes and C. Thompson, "The international IT professional practice programme," 2007.

[5] M. Moullin, "Eight essentials of performance measurement," *International Journal of Health Care Quality Assurance*, vol. 17, pp. 110-112, 2004.

[6] R. Rada, "IT skills standard," *Communications of the ACM* vol. 42, pp. 21-26, April 1999.

[7] B. Wright and M. Stone, *Measurement essentials*, 2nd ed. Delaware, 1999.

[8] D. C. McClelland, "Testing for competence rather than intelligence," *American Psychologist*, vol. 28, pp. 1-14, 1973.

[9] R. M. Felder, "ABET criteria 2000: an exercise in engineering problem solving," *Chemical Engineering Education*, vol. 32, 1998.

[10] R. Adams, "Clarification_OBE Attributes vs Blooms'," 2000.

[11] E. Smith and R. Smith, *Introduction to Rasch Measurement: Theory, Models and Applications*. Minnesota: JAM Press, 2004.

[12] Tevor G. Bond and C. M. Fox, *Applying the Rasch Model*, 2nd ed. New Jersey: Lawrence Erlbaum 2007.

[13] B. D. Wright, "Fundamental measurement for outcome evaluation," *Physical medicine and rehabilitation : State of the Art Reviews*, vol. 11, pp. 261-288, 1997.

[14] M. Mok and B. Wright, "Chap 1: Overview of Rasch Model Families," in *Intoduction to Rasch Measurment: Theory, Models, and Applications* Minnesota: Jam Press, 2004, pp. 1-24.

[15] D. Andrich, *E444-Advance Social and Educational Measurement, Semester 2 Unit Materials*. Perth, W.A: School of Education, Murdoch University, 1999.

[16] "ISO 9001:2000 Quality Management Systems - Requirements," International Organization for Standardization, Geneva.