Students Performance Evaluation for Learning Outcomes Measurement: SPELOM Model

ZAMALIA MAHMUD [1], MOHD SAIDFUDIN MASODI [2] AND AZRILAH ABD. AZIZ[3]
[1] Associate Professor, Ph.D. Faculty of Computer and Mathematical Sciences,
Universiti Teknologi MARA, 48000 Shah Alam, Selangor MALAYSIA
Tel/SMS: +6012 2197985 Fax: +603-55435501 e-mail: zamal669@salam.uitm.edu.my
[2] Program Coordinator, Exec.Dip in QMS / QMS ISO9000 Lead Assessor,
School of Professional Advancement & Continuing Education,
University Teknologi Malaysia, 81300 UTM Skudai, MALAYSIA.
Tel/SMS: +6012 2402821 Fax : +603-41072262 e-mail:saidfudin@gmail.com
[3] Ph.D., Program Coordinator, Exec.Dip in Protective Security Science,
School of Professional Advancement & Continuing Education,
University Teknologi Malaysia, 81300 UTM Skudai, MALAYSIA.
Tel/SMS: +6019 3332661 Fax : +603-41072262 e-mail:azrilah@gmail.com

***Abstract**:* **-** The Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (FCMS) teaching and learning processes were certified to ISO 9001:2008 and now putting all the concerted effort to comply to Malaysia Quality Assurance (MQA) framework as part requirement of ISO9001:2008 compliance to statutory requirements. MQA adopts the American Accreditation Board of Engineering and Technology 2000 (ABET) principles which promote outcome based education (OBE) learning process. OBE calls for the evaluation of the course learning outcomes (CLO) as specified in each Course Outline. Performance Measurement has been largely dependent on students' performance in carrying out tasks such as tests, quizzes or submission of assignments. Evaluation on the performance outputs; categorized as mastery of knowledge and skill development, gives an indication on the achievement of the subject's expected CLO. This paper describes a computational model which can be used to measure a subject CLO in an undergraduate program. An overview of the measurement model and it's key concepts are presented. SPELOM Model is the acronym for Student Performance Evaluation on Learning Outcomes Measurement which is developed based on the Rasch Model. It can be used to improve the students' assessment method by CLO of each subject instead of the traditional raw score grading. Results obtained were assessed using Rasch Analysis where the strength of the measurement lies in its ability to precisely map out the CLO for evaluation of differences and correlation between the students, β_n performance and item difficulty, δ_i The study shows that this model of measurement adopting the Rasch Model can classify students learning ability more accurately based on Bloom's Taxonomy dimensions as compared to the traditional CGPA method. Hence, this model has the novelty to serve as a better ruler to more accurately measure students' knowledge mastery and skill development. The usefulness of this new measurement instrument is very significant especially in developing prudent continual quality improvement (CQI) measures of the teaching method effectiveness thus meeting the requirement of MQA holistically.

***Keywords**:-* Learning Outcomes, performance measurement, Quality education, Bloom's, Rasch Analysis

1.0 Introduction

The Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (FCMS) has taken the challenge to improve their undergraduate programs teaching and learning system by meeting ISO9001:2008 – Quality Management System requirements for the scope of service provision in teaching and learning of Computer and Mathematical Sciences. They are also doing their best endeavour to observe the Malaysia Quality Assurance (MQA) framework as stipulated in the Quality Assurance Code of Practice for Public Institutions of Higher Learning Education, Malaysia issued by the Quality Assurance Division, Ministry of Higher

Education; a Malaysian Government authority in quality education whose approval is of prime importance before any entity wish to offer any educational program to the Malaysian public. It is FCMS top management commitment to meet MQA program accreditation requirements where among others promote outcome based education (OBE) learning process. OBE calls for the evaluation of the course learning outcomes (CLO) as specified in each Course Outline hence program performance measurement. Practically, it has been largely dependent on students' performance in carrying out tasks such as a series of tests or quizzes, final examination and submission of assignments. Evaluation on the performance outputs; encompassing both categories, technical knowledge and generic skills gives an indication on the achievement of the subject's expected CLO. However, the current practice of Cumulative Grade Point Average (CGPA) is only a mean average of raw scores which lacks precision and linearity hence validity required to meet the fundamental criteria of measurement.

2.0 Fundamentals of Measurement

Measurement is of utmost importance in our everyday life. Three(3) major use of measurement are; 1. Regulate trade, 2. Monitoring; and 3. Calibration. Academicians have great need for the development of valid measures, e.g., of the quantity and quality of education services and the outcomes of these services; be it teaching and learning as well as the conduct of researches. In FCMS, the theory and practice of classical test theory, the traditional approach of assessment and evaluation effectiveness by simple raw score is therefore thoroughly reviewed. It then provides an overview of "modern" measurement as practiced using item response theory with focus on Rasch Measurement Model [1].

This paper describes a computational model which has been used to measure a subject CLO in an undergraduate program in University Teknologi Malaysia (UTM) which is then further validated and confirmed in FCMS [2]. This hybrid model is developed largely based on Rasch Measurement Model can be used to improve the students' assessment method and verify each stage of the CLO for each course taught. Results obtained were evaluated against the CLO map; developed based on Bloom's Taxonomy and learning achievements described by SOLO Taxonomy for consistency [3, 4]. The information generated from this measurement are of meaningful use to guide us determine the appropriate quality improvement of the teaching method or style employed as well as in determining of the validity of the examination questions prepared thereafter. Questions were assessed on their Point Correlation Measure; whether it is measuring what it is supposedly to measure and subsequently scrutinised on its level of difficulty before it can be considered as a bankable item in FCMS question bank. Thus, the construct validity of a particular examination paper and the CLO measurement is therefore resolved simultaneously.

The data is then transformed into a linear interval scale using the *logit* ruler, primarily to obtain uni-dimensionality of measure with better precision to measure the ability of students in respect of their learning difficulty encountered. It can be shown by simple mathematical concept of indices that a series of probabilities of observed events described by log series maintained an equal separation; thus equal interval. This equal separation we termed it *logit* as unit of measurement for ability akin to ^{o}*Celcius* to measure temperature or alike in metrology [5].

This provides a sound platform of measurement equivalent to natural science which matches the following SI Unit criteria; there must be an instrument of measurement with a defined unit. It is quantifiable by mean of linearization with reasonable accuracy. The measurement shall be replicable and consistent and; is predictive to overcome missing data [6].

3.0 Measurement Methodology

Responses from the students examination results were analysed using 'park mark system' in which the students were rated according to their achievement by 'key words' of each topical area of study. Practically, this is only counting the responses of correct and wrong answers from the students responses who sat for the examination that gives a raw score for each course being evaluated. This serves as a guide to rank the students for grading. However, raw score can only give 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. It does not meet the fundamentals of sufficient statistics for evaluation. Alternatively, data set would normally be put on a scatter plot to establish the best regression. However, prediction from ordinal responses on the ability attributes are almost impossible due to absence of intervals in the scale. The normal solution is to apply the regression approach where a line which fits the points as best as possible; which is then use it to make the required predictions by interpolation or extrapolation as necessary as shown in Figure 1.

$$y = \beta_0 + \beta_1 m \qquad \text{Equ.(1)}$$

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

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

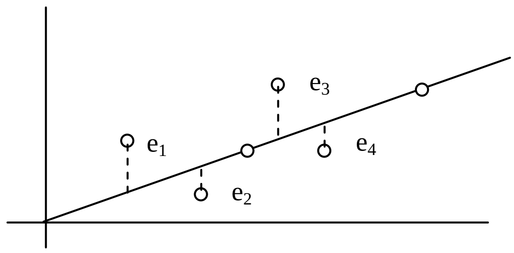


Figure 1. Best fit line: Linear Regression Model

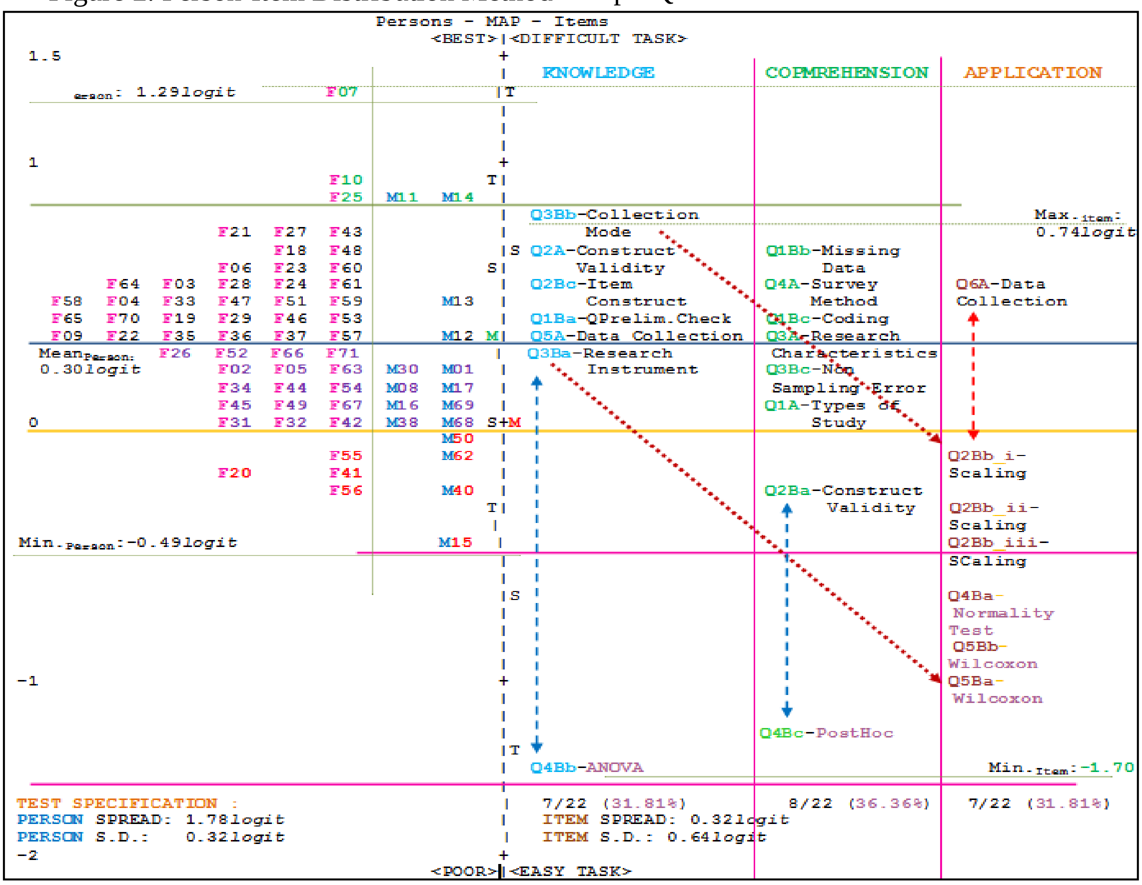
By accepting the fact that there is always error involved in the prediction model, the deterministic model of equation (1) can be transformed into probabilistic model by including the prediction error into the equation;

$$y = \beta_0 + \beta_1 m + e \qquad \text{Equ. (3)}$$

Rasch moves the concept of reliability from establishing "best fit line" of the data into producing reliable repeatable measurement instrument. Rasch focuses on constructing the measurement instrument rather than fitting the data to suit the measurement model. By focusing on the reproducibility of the latent trait 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 ability in an appropriate way is vital to ensure valid quality information can be generated for meaningful use.

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 [7].

Figure 2. Person-Item Distribution Method Map - QMT455:Research



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

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

where;

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

β_n = person's ability

δ_i = item or task difficulty

4.0 Data Analysis and Discussions

The test QMT455-Research Method was administered on 3rd year undergraduate students in Statistics from the Department of Statistical Study, FCMS. The result from the test were tabulated and run in *WinSteps 3.68.2,* a Rasch Analysis software; to obtain the *logit* values. Figure 2 shows the Person-Item Distribution Map (PIDM) where the *person*; i.e. the Students and the *item*; the learned topics were plotted on the same *logit* scale. By virtue of the same scale; then the basic rule of additivity, the correlation of the *person*, β_n and *item*, δ_i can now be established as in equation (6).

Summary statistics of Person and Items measures were next captured. It is then used to complete the PIDM indicating both the Person and Item maximum and minimum to give an indication of the person and item spread hence Standard Deviation (SD). The respective summary measurements is shown in Figure 3 – Persons Measure and Figure 4 for Items Measure.

SUMMARY OF 71 Persons MEASURED

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	66.5	19.1	.30	.18	.99	.0	1.09	.2
S.D.	10.4	0.6	.32	.02	.32	1.0	.90	1.1
MAX.	90.0	22.0	1.29	.33	1.90	2.4	6.68	5.8
MIN.	43.0	17.0	-.49	.16	.37	-2.4	.26	-1.5

REAL RMSE .19 ADJ.SD .26 SEPARATN 1.34 RELIABILITY .64
MODEL RMSE .18 ADJ.SD .26 SEPARATN 1.46 RELIABILITY .68
S.E. OF Person MEAN = .04

LACKING RESPONSES: 1Person VALID RESPONSES: 86.9%
Person RAW SCORE-TO-MEASURE CORRELATION = 0.97
CRONBACH ALPHA (KR-20) Person RAW SCORE RELIABILITY = 0.61

Figure 3 –Summary Statistics: Person Measure

Figure 3-Person Summary reveals a good person spread of 1.78*logit* where Max_{Person}=1.29*logit* and Min_{Person}= -0.49*logit* with person SD_β=0.32 and Separation, G_β=1.34 but rather low reliability of Cronbach α = 0.61. The major finding is the Person Mean, μ_{Person}= **0.30***logit* (P(θ)=0.5744) where the Students were found to be merely above the expected performance with poor Person Reliability=0.64. In SOLO Taxonomy terms; students are at unistructural level of learning where simple and obvious connections are made but their significance is not grasped. From Figure 2, only 7.04%(N=5) of the students measured were found to be exemplars having acquired the expected Learning Outcomes whilst 11.26 % (N=8) students were discovered to have difficulties in grasping the subject matter proper. Further scrutiny is done on items by topic and Bloom's Taxonomy cognitive learning curve categorised in six (6) domain from the simplest to complex;*knowledge, understanding, application, analysis, evaluation and synthesis.*

SUMMARY OF 22 Items MEASURED

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	211.6	60.8	.00	.13	1.11	.0	1.06	.0
S.D.	86.5	16.2	.64	.07	.40	1.2	.32	1.0
MAX.	345.0	70.0	.74	.38	2.44	1.7	1.80	1.8
MIN.	46.0	21.0	-1.70	.08	.51	-2.6	.50	-2.1

REAL RMSE .18 ADJ.SD .61 SEPARATN 3.34 RELIABILITY .92
MODEL RMSE .14 ADJ.SD .62 SEPARATN 4.34 RELIABILITY .95
S.E. OF Item MEAN = .14

Figure 4 –Summary Statistics: Item Measure

Generally, the students find difficulties with 45.46% (N=10) of the questions asked; where item *logit* > person mean, μ_{person}= 0.30*logit*. The most difficult item is the item at the top, like the high jump bar analogy. Being high then it is a difficult item to attempt. Figure 4- Item Summary gives a good summary with Item Separation, G=3.34 and a very high reliability=0.92*logit* reflecting the true measurement of the instrument. However, it has poor item spread of 0.32*logit* with SD*i*=0.64 and the findings require a serious review as knowledge items were found to be more difficult than the application items. This negative trend is just totally opposite the norm when learning is an upward trend. One possible reason is due to the emphasis of mathematics rather than making sense of an observation during the conduct of this course. Nevertheless, students F55, F41, F20, F56, M50, M62, M40 and M15 is definitely in trouble as they have serious difficulty understanding this course where they are located well below all items. From Figure 2, it is interesting to note the correlation of the difficult items from each domain; with Q3Bb - Data Collection Mode (*Knowledge*), Q1Bb - Missing Data (*Understanding*) and Q6A - Data Collection (*Application*) are of common nature. Q4A-Survey Method (*Comprehension*) high location further reinforced the students unistructural mind development (*SOLO Level 2*). Items concerning mathematics; Q4Bb-ANOVA (*Knowledge*), Q4Bc-Post-Hoc (*Comprehension*) and Q5ba-Wilcoxon (*Application*) are easy items but seems detached from the statistical fundamentals.

It was also noted there is a huge gap indicating very easy questions denoted by (← →), between Q4Bb-ANOVA (*Knowledge*) and Q4Bc-Post-Hoc (*Comprehension*) and Q3Ba-Research Instrument (*Knowledge*), Q2Ba-Construct Validity (*Comprehension*). A difficult item; Q6A-Data Collection (*Application*) is noted (← →) by the gap against Q2Bb (*Application*). Take note of the other items in this cluster; Q4Ba-Normality, Q5Ba and b-Wilcoxon does not correspond to any Person at all. These are too easy items which need review to make the task a little bit more difficult or even possible discard. On the opposite end, we have Persons but without any items against it. Hence, students F07, F10, F25, M11 and M14 are exceptional students in this cohort who does not have much difficulty in attempting any given task.

In summary, the PIDM analysis clearly identify that there are four(4) groups of students profile from the poor to excellent as demarcated in the person, β_n column. Similarly, the items i.e. Questions is basically of four(4) types too.

Inspite the high item reliability, the construct validity of the of the items is further verified by analysis of the Point Measure Correlation as shown in Figure.5.

Controls applied was to checked the item as acceptable when the Point Measure=x; $0.4< x <0.8$. Next is to verify the suspect by looking at the Outfit Mean Square (MNSQ)= y-value to be in the range of $0.5 < y<1.5$. The final check would be on the Outfit z-standard (ZSTD)= z-value if it is within the range of; $-2< z<2$. Q10- MNSQ=2.47>1.5 and ZSTD>2.2; thus it confirms an item misfit.

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	Item
20	339	70	-1.18	.24	1.52	1.0	1.44	.9	.12	Q4Bc
19	345	70	-1.70	.38	2.44	1.7	1.74	1.1	.12	Q4Bb
10	285	68	-.28	.10	1.14	.7	1.29	1.1	.12	Q2Ba
21	281	59	-.93	.20	1.84	1.6	1.80	1.4	.16	Q5Ba
14	162	68	.58	.08	1.21	1.5	1.33	1.8	.24	Q2Bc
22	266	59	-.54	.13	1.01	.1	1.08	.3	.26	Q5Bb
1	226	70	.21	.08	1.02	.2	1.00	.0	.29	Q1A
13	290	68	-.33	.10	1.35	1.6	1.44	1.4	.30	Q2Bb
18	314	70	-.52	.12	1.17	.7	.96	.0	.31	Q4Ba
5	193	70	.41	.08	1.03	.3	1.06	.4	.36	Q5A
2	167	70	.57	.08	1.05	.4	1.04	.3	.38	Q2A
3	212	70	.29	.08	.86	-1.2	.90	-.7	.38	Q3A
16	130	63	.74	.09	.75	-1.6	.84	-.6	.43	Q3Bb
7	53	21	.49	.15	.90	-.3	.85	-.4	.44	Q1Ba
12	287	68	-.30	.10	1.21	1.1	1.09	.4	.45	Q2Bb
17	192	63	.29	.08	.92	-.6	.87	-.9	.48	Q3Bc
15	190	63	.30	.08	.87	-1.1	.88	-.8	.49	Q3Ba
4	179	70	.50	.08	.71	-2.6	.74	-1.8	.52	Q4A
11	270	68	-.15	.09	1.12	.8	.97	-.1	.53	Q2Bb
6	172	68	.51	.08	.94	-.4	.86	-.9	.56	Q6A
8	46	21	.66	.16	.83	-.5	.70	-.8	.64	Q1Bb
9	57	21	.40	.15	.51	-2.3	.50	-2.1	.78	Q1Bc
MEAN	211.6	60.8	.00	.13	1.11	.0	1.06	.0		
S.D.	86.5	16.2	.64	.07	.40	1.2	.32	1.0		

Figure 5 - Point Measure Correlation: Item validity

It is considered as an misfit only when all the three(3) controls are violated. This is a more detailed controlled as compared to the traditional Classical Test Theory (CTT) where it only applies simple discrimination index to make an item bankable or not.

CATEGORY LABEL	CATEGORY SCORE	OBSERVED COUNT	OBSERVED %	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
1	1	348	26	-.22	-.25	1.08	1.34	NONE	(-1.08)	1
2	2	86	6	-.04	-.10	1.28	2.03	1.22	-.40	2
3	3	126	9	-.02	.10	.85	.56	-.38	-.01	3
4	4	132	10	.23	.38	1.06	.67	.18	.39	4
5	5	646	48	.81	.78	.90	.93	-1.02	(1.10)	5
MISSING		202	13	-.04						

Figure 6 -Summary of Category Structure

The structure calibration; ‘*s*’ is assessed to confirm the rating classification used is applicable where *s*-value being the difference between each structure;

e.g; s_{3-2}= 1.22-(-0.38) =1.60; > 1.4, OK.

S_{5-4}= 1.02-0.18 =0.84; < 1.4, ***Not*** OK

The result shall be in the range where s; $1.4< s<5$. It is noted that the difference for each category are irregular where the difference between category 1, 2, 3, 4 and 5 are all less than 1.4. Therefore, the classification A,5>90; B, 4>80; C,3>70; D,2>60 and Fail,1<60 is not reflective of this cohort person separation. In Rasch, this is termed as collapsing.

In summary, Figure 7 shows there are only two groups of students; between who knows and knows not. It reveals that the responses pattern is conspicuously dichotomous of 1 and 5 only. The rest of the other ratings were practically submerged. This call for Rasch Analysis by dichotomous approach. If the SD is found to be larger, then the dichotomous results shall be used or *vice-versa*.

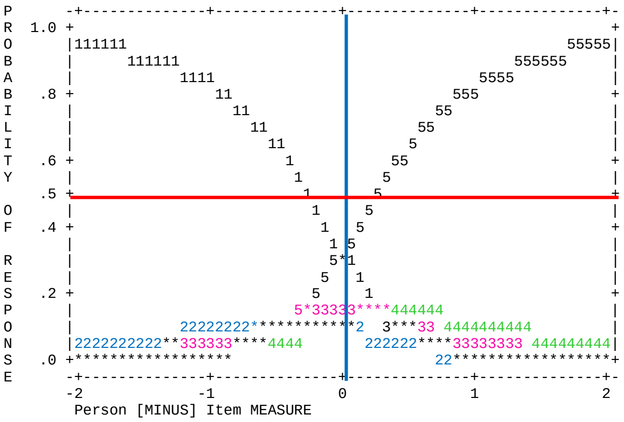


Figure 7 -Category Probability Curve

Next the students learning ability for each CLO identified can be derived from the Person Measure Table as follows, e.g. μ_{Person}:

$$P(\theta) = \beta_v - \delta_i \quad \text{(f rom Equ.6)}$$
$$= -0.30 - 0$$

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

$$P(\theta) = \frac{e^{0.30}}{1+e^{-0.30}}$$

$$P(\theta) = 0.5744$$

Thus, Students Learning Ability = 0.5744

Generally the students fared poorly below the expected performance achieving a poor mean; μ_{person} of only **57.44%** which is below the 68% threshold limit; being 2-δ Level of Learning Competence set by the Academic Council of the University Senate . It is very interesting to note that all (*Knowledge*) questions; viz; Q3Bb, Q2A, Q2Bc, Q1Ba, Q5A and Q3Ba are are located towards higher order of difficulty where students find hard to resolve. Instead students find otherwise in attempting question on (*Application*) viz; Q2Bb, Q4Ba, Q5Ba and Q5Bb were found to be easier task kept at bay below the person mean; μ_{person}= -0.49.

This peculiar observation is also noted in a similar research experience in the middle east [8]. It can be deduced that technical students does not go for reading and understand fundamental concepts. Their preference is towards mechanical questions involving applied mathematics that belongs more to *application* and *analysis* in Bloom's Taxonomy learning domain. In fact it can be seen that these questions; viz. Q4Bb-ANOVA, Q4Bc-Post-Hoc and Q5Ba-Wilcoxon are located very much towards the easy task end.

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	
7	90	19	1.29	.33	.74	-.1	.51	-.2	F07
10	87	20	.92	.20	.57	-1.1	.37	-.9	F10
25	84	19	.89	.22	.53	-1.2	.45	-.5	F25
11	86	20	.88	.19	.80	-.4	.52	-.6	M11
14	84	19	.88	.22	.56	-1.1	.41	-.6	M14
43	81	20	.71	.17	.67	-1.2	.77	-.2	F43
21	80	19	.71	.20	1.03	.2	.67	-.2	F21
35	67	19	.31	.17	1.07	.3	.87	.0	F35
57	67	19	.31	.17	1.23	.9	3.45	2.9	F57
66	67	19	.29	.17	.91	-.2	.91	.0	F66
71	67	19	.29	.17	1.04	.2	.98	.2	F71
26	66	19	.26	.17	1.08	.4	.97	.1	F26
30	56	18	.01	.18	1.62	1.7	1.48	1.0	M30
31	57	19	-.01	.18	1.45	1.3	1.34	.8	F31
32	57	19	-.01	.18	.75	-.7	1.01	.2	F32
68	57	19	-.01	.18	.62	-1.2	.58	-.8	M68
38	53	20	-.03	.17	1.14	.5	1.08	.3	M38
42	53	20	-.03	.17	1.90	2.4	6.68	5.8	F42
50	51	20	-.09	.17	1.12	.5	1.39	.9	M50
55	43	17	-.13	.21	.37	-1.8	.31	-1.5	F55
62	57	22	-.16	.17	1.58	1.6	1.60	1.3	M62
41	51	19	-.20	.18	1.01	.1	.89	-.1	F41
20	47	20	-.21	.18	1.36	1.1	1.50	1.0	F20
40	50	19	-.24	.18	1.14	.5	1.07	.3	M40
56	49	19	-.27	.18	1.12	.5	.93	.0	F56
15	43	19	-.49	.20	.63	-1.1	.54	-.7	M15

Figure 8 –Person Measure Table

Figure 8 also exposed that student F42 who appeared to have passed the examination is found to be a misfit where he failed the 3-criteria. Scrutiny of the person key performance discloses the student response pattern does not meet Rasch Model expected outcome. Based on the pattern of her given answers, in Figure 9, Rasch Model expected him to give a response of 4 and 5 for Q4Ba- Normality Test (*Application*) and Q4Bb- ANOVA (*Knowledge*) respectively instead of 1. Rasch has this special predictive feature embedded in the model to make it very reliable.

This call for reasoned argument on their level of true ability. Student F42 has given responses in the opposite direction where she is able to answer the difficult questions but somehow failed to answer the very easy questions.

In Rasch, we are interested to find out why such behavioural change occurs. Two important things is happening here; first is whether the learning process is at stake or secondly the teaching process need to be reviewed whereby reasoned arguments is given emphasis rather than pure technical and mechanical. Such skewed perception requires correction in our effort to produce not only quality graduates but also quality human capital who are equally qualified and competent.

```
NUMBER - NAME ------------------ MEASURE - INFIT (MNSQ) OUTFIT - S.E.
     42  F42                        -.03     1.9    A    6.7      .17

-3         -2          -1           0           1           2
|-----------+-----------+-----------+-----------+-----------|  NUM   Item
                                    2          .4.              14  Q2Bc
                                    2     .3.                    2  Q2A
                            .1.     2                            7  Q1Ba
                           .1.      2                            5  Q5A
                                    2               .5.         15  Q3Ba
                          .1.       2                            3  Q3A
                                    2      .4.                  17  Q3Bc
                                    2              .5.           1  Q1A
                    .1.             3                           11  Q2Bb_i
                                    4        .5.                12  Q2Bb_ii
                                    4       .5.                 13
Q2Bb_iii
                (1)                 4                           18  Q4Ba
  (1)                               5                           19  Q4Bb
|-----------+-----------+-----------+-----------+-----------|  NUM   Item
-3         -2          -1           0           1           2
```

Figure 9 -Person Key Performance

5.0 Conclusion

SPELOM model developed based on Rasch Model, Bloom's and SOLO Taxonomy learning domain provides a sound platform of measurement equivalent to natural science which matches the SI Unit measurement criteria; where it behaves as an instrument of measurement with a defined unit and therefore replicable. It is also quantifiable since it's linear. Rasch has made it very useful with its predictive feature to overcome missing data. SPELOM give a new measurement paradigm which is easier to read and better analysis using Rasch-based approach

The measurement conducted using SPELOM reveals the true degree of learning abilities of the undergraduates. Previously, lack of such measurement in Malaysia as well as in FCMS has made the necessary corrective actions in the form of skills development, education and competency training difficult to formulate. This major problem faced by Technical Education Administrators in an IHL to design the necessary curriculum to mitigate the going concern is therefore resolved. A Computer Aided Test software is currently being rigorously tested for validation before used [9].

References

[1] B. D. Wright and M. M. C. Mok, "An overview of the family of rasch measurement models," in *Introduction to Rasch Measurement: Theory, Models, and Applications*, J. Everett V.Smith and R. M.Smith, Eds., 2004, p. 979

[2] Saidfudin, M., Zaharim, A, Rozeha, A.R., Razimah, A. & Ghulman, H.A., "Engineering Students Performance Evaluation of Generic Skills Measurement: ESPEGS Model" *in* Mathematics and Computer Science Engineering: New Aspects of Engineering Education, Mastorakis et al., a series of reference books and textbooks published by WSEAS Press, July 2008, pp. 377-383. indexed in ISI, ACM

[3] Saidfudin, M, Azlinah M , Azrilah AA, Nor Habibah, A. & Sohaimi Z, "Appraisal of Course Learning Outcomes using Rasch measurement: A case study in Information Technology Education", *International Journal of Systems International Journal of Systems Applications, Engineering & Development*; Issue 4, vol.1, University Press, UK. pp.164-172, July 2007

[4] Atherton, J.S (2009) *Learning and Teaching; SOLO taxonomy* [On-line] UK: Available at:http://www.learningandteaching.info/learning/solo.htm accessed: 18 December 2009

[5] Saidfudin, M., and Azrilah, AA; "*The Rasch Model: A simple Scale Construct and Measurement Structure"*; UPENA, S.Alam, 2010

[6] Saidfudin, M, Azlinah M , Azrilah AA, NorHabibah, A; Hamza A Ghulman & Sohaimi Z, "Application of Rasch Model in validating the construct of measurement instrument", in *International Journal of Education and Information Technologies*, Issue 2, Volume 2,. pp. 105-112; May 2008

[7] T.G. Bond and C. M. Fox, *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*, 2nd ed. Mahwah, New Jersey: Lawrence Erlbaum, 2007

[8] Saidfudin, M & Hamza A Ghulman, "*Modern measurement paradigm in Engineering Education: Easier to read and better analysis using Rasch-based approach*", in proceeding of 2009 IEEE International Conference on Enginering Education (ICEED 2009), 7-8th Dec. 2009, Kuala Lumpur, Malaysia

[9] Azrilah, A.A., "Information Professional Competency Model in Public Institutions", unpublished Ph.D thesis, UiTM 2009.