# Application of Rasch-based ESPEGS Model in Measuring Generic Skills of Engineering Students: A New Paradigm

ROZEHA A RASHID [1], RAZIMAH ABDULLAH [2],
HAMZAH AHMAD GHULMAN [3] AND MOHD SAIDFUDIN MASODI [4]
[1] Senior Lecturer, Department of Telematics and Optics
Faculty of Electrical Engineering, Universiti Teknologi Malaysia
81310 UTM Skudai, Johor, MALAYSIA.
[2] Research Consultant, Center of Teaching and Learning,
Research Management Center, Universiti Teknologi Malaysia
81310 UTM Skudai, Johor, MALAYSIA.
[3] Vice Dean (Academic), Dept. of Mechanical Engineering
College of Engineering & Islamic Architecture,
Umm al-Qura University, PO Box 9125 Makkah, KINGDOM OF SAUDI ARABIA.
[4] QMS Professional Cert. Coordinator / QMS ISO9000 Lead Assessor,
School of Professional Advancement & Continuing Education,
University Teknologi Malaysia, 81300 UTM Skudai, MALAYSIA.
Tel/SMS: +60122402821 Fax : +603-41072262

e-mail: rozeha@fke.utm.my, arazimah@gmail.com, hghulman@uqu.edu.sa, saidfudin@gmail.com

*Abstract:* - The Faculty of Electrical Engineering, Universiti Teknologi Malaysia (FKE) teaching and learning processes was certified to ISO 9001:2000 and now seeks the Engineering Accreditation Council of Malaysia (EAC) approval. One of FKE's top management commitment is to ensure the programme accreditation requirements are met. EAC adopts 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; categorised as technical knowledge and generic skills, 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 electrical engineering program. An overview of the measurement model and its key concepts are presented. ESPEGS Model is the acronym for Engineering Student Performance Evaluation on Generic Skills. This model of measurement is developed based on students' marks entries and together with Rasch Measurement Model, it can be used to improve the students' assessment method on the CLO of each subject. Results obtained were assessed against the CLO maps for consistency and used as a guide for future improvement of the teaching method. The study shows that this model of measurement, which adopts Rasch Model based on Logistic Regression Model, can classify students learning ability more accurately with only very few randomly selected students and dimensions as compared to the traditional CGPA method.

***Keywords***:- Learning Outcomes, performance measurement, Quality, engineering education, Rasch Model

## 1 Introduction

Factors such as increased global competition, accelerated technological advancements and enhanced customer requirements have caused fundamental changes in the manner in which organisations compete. Organisations can no longer compete solely on the basis of cost; *value-for-money,* but must formulate competitive strategies defined by industrial market-driven requirements. Therefore, it has become increasingly important for organisations to develop strategic objectives which facilitate the development of a competitive advantage in specific markets or market segments. Strategic objectives are initiatives designed to have a significant and favourable effect on the long-term health of the organisations; in this case the Institutions of Higher Learning to remain relevant. The improvement of products or programmes offered together with the required processes in teaching and learning, and service quality have been adopted by many IHLs as the key strategic objectives for achieving world-class performance levels at optimum cost [1]. However, sustainable world-class performance will not occur if there is a misalignment between an Institution of Higher Learning (IHL) programmes objectives and actual market requirements. In addition, effective faculty wide coordination in relation to market driven initiatives is essential for ensuring the effective use of an institution's resources.

In order for an IHL to successfully compete on its strategic programmes objectives, relationships must exist between the IHL's strategies, faculty wide actions and performance measures. Not only are specific action programmes supporting strategic objectives required, an integrated performance measurement system (PMS) which facilitate consistent organizational actions toward objective achievement is also vital.

In an IHL, PMS provides a means for (a) maintaining an **alignment** between strategic programmes objectives and industry requirements; (b) coordinating the **effective** use of IHL resources; and (c) **monitoring** teaching and learning progress towards the achievement of pre-determined course learning outcomes (CLO); hence, the CLO Map. Consequently, a PMS is required for each CLO to serve as a mechanism for monitoring progress or the achievement of each CLO. It has found a place in the field of Information and Communication Technology education and on the same framework the principle applies in engineering education [2].

The purpose of this paper is to describe a measurement model to compute the development of students' learning in a quality focused engineering education system. The model is developed based on data from a longitudinal evaluation of selected electrical engineering students' performance since academic session 2004/05 in the Faculty of Electrical Engineering, Universiti Teknologi Malaysia, Skudai, Johor (FKE). The study focused on the relationship between teaching and learning (T&L) method and students' performance measurement systems designed to support FKE's program objectives so as to facilitate faculty wide coordination to continually improve the quality of engineering education.

FKE needs to comply with the Engineering Accreditation Council of Malaysia (EAC) programme accreditation requirements. EAC adopted 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 subjects LO's as specified in each Course Outlines. In FKE, Performance measurement has been largely dependent on students' performance in carrying out tasks such as tests, quizzes or project papers / assignments. Evaluation on the performance outputs, broadly categorised as technical knowledge and generic skills, gives an indication on the achievement of the subject expected LO's.

Monitoring and measurement is a vital process, meeting the PDCA approach – *Plan, Do, Check, Act*; the founding method of ISO9001:2000 to assure customer satisfaction where FKE's teaching and learning process is duly certified [3]. The resulting model provides a foundation for further development of quality-focused engineering education performance measurement system theory based on empirical findings. In addition, the model can be used as an instrument for the implementation of quality-focused performance measurement systems in IHL.

## 2. LITERATURE: An Overview of Performance Measurement

Performance measurement of LO in IHL is relatively new, undeveloped and yet to be studied systematically. Although some of the functionally-based performance measurement literature examined performance relationships across functional areas, little was written about the alignment of functional performance to overall programme outcomes. Moreover, many of these articles focused on the local optimization of each subject area with little regard on how other teaching and learning dimensions or attributes may be affected. This calls for the adoption of a more global engineering education performance measurement mechanism which would optimize the effectiveness of teaching and learning, thus, benefiting the entire programme.

The authors conducted a comprehensive review of literature pertaining to performance measurement system design and categorized it into three distinct areas [1, 4]:

1. individual performance measures;
2. performance measurement systems as an entity; and
3. relationships between performance measurement systems and the environment.

Early researchers who focused on individual performance measures examined various dimensions of quality education, cost, time, and flexibility from a strategic perspective. In an attempt to define the various attributes of a performance measurement system, researchers developed frameworks for relating functional or teaching and learning performance to overall programme performance.

Others have examined the interaction between a performance measurement system and its teaching and learning method. Literature pertaining specifically to quality-focused performance measurement systems can be classified into three(3) broad categories; a) quality measures, b) quality measurement; and c) frameworks for developing quality measurement systems.

The use of statistically-based measures to monitor and control process and product quality was pioneered by Shewhart (1931), Juran (1951) and Deming (1975). In addition, Kane (1986) explored the use of capability indices as a measure of process quality. Initial attempt to measure the Learning Capability Index by the authors was presented at the International Forum on Engineering Education (IFEE 2006), Sharjah, UAE [5]. It provided some insight on the development for a more comprehensive evaluation system and assessment of the strategic impact of engineering education quality initiatives.

It is good to note other research done in PMS had addressed the problems associated with the use of quality measures in isolation and they highlighted the need for a holistic approach to quality performance measurement. In the advent of internet and globalisation, information system is also an essential dimension for an excellent framework of a quality management system.

Performance measurement should generate accurate, meaningful information i.e., be reliable and valid. Performance measurement represents a vision that can shape the future direction of classroom-based assessment; hence subjects LO, but it requires much additional scrutiny and development before it can fulfill its promise.

There is a need to articulate the necessity for IHLs to adopt a "customer-driven" approach to quality engineering education designed to avoid misalignments between an IHL products; the programmes offered, or service offerings; teaching and learning; and the requirements of the targeted industry market segment. A good PMS can enhance the understanding of such alignment, and assist academicians in developing and maintaining quality as well as relevant engineering programmes duly aligned between IHL and the industry.

Performance measurement can be summarily viewed in the correlational ABC Model on how cognitive skills and affective state is reflected in the behaviour of students during learning. Weybrew (1992) discussed at length on the repercussion of such development but believed that affective values are of significant importance in neuro-linguistic programming (NLP) [6].

## 3. Measurement Methodology

This study addresses the three(3) following questions:

1. What are the established LO's at the T&L levels by the IHL?
2. Do performance measurement systems used to evaluate progress on the subjects LO in an IHL applicable globally?
3. How are performance measurement system linkages accomplished on the LO's at T&L in an IHL ?

A method of defining the required metrics in Engineering Education Performance Measurement is setforth modelled on Razimah (*2006*) Plan-Execute-Report-Monitor (P-E-R-M) assessment method to measure the Value for Money (VFM) Audit performance [1]. This model is a variant of Shewhart's (*1939*) P-D-S-A Cycle which was subsequently developed into the infamous Deming's (*1957*) P-D-C-A Cycle by the Japanese industrial community. Then, in year 2000 in Geneva, this fundamental concept of P-D-C-A was adopted by the international community for ISO9000 certification. In IHL, performance measurement is given great importance by many IHL administrators, particularly with the world class ranking now being the ***in-thing*** viz; Times Higher Education Survey.

In measuring performance, the attempt is to use a simple statistical technique yet that can yield very accurate findings using data-driven approach to analyse the root causes of each learning problem encountered [3, 5]. It is a much more disciplined approach for assessing students' generic skills during a learning process.

Communication skill, teamwork, life long learning, etc., are generic skills which are termed as dimensions. Within these dimensions, relevant main areas related to the subject learning outcomes is then identified but not limited to, viz; vocabulary power, technical appreciation, software development and resourcefulness. Collectively, these main areas are known as attributes which are all measurable. This method of measurement using dimensions and attributes to measure performance has been applied in the evaluation of internal audits for constructing an Audit Performance Index [7]. Measurement is made possible by transforming the raw score ordinal data into quantifiable interval data [8] .

Data analysis utilised is within-case analysis technique. Summary tables of the CLO's were developed for assessment. A consolidated table is used to show how the performance measurement system linkages are accomplished on the identified CLO's. This tactic improved the probability of developing theoretical models which were a "close fit" with the data. A probabilistic algorithm of data fit is determined next.

The assessment form, as an instrument of measurement, is designed and developed for the attributes which is rated based on an even number scale of 1 – 6 dichotomously indicating **NO** – **YES** with 2,3 – 4,5 indicating their level of agreement to an attribute. This assessment form gathers empirical data as the main component of this study. Table 1 shows the conceptual format of the designed assessment form.

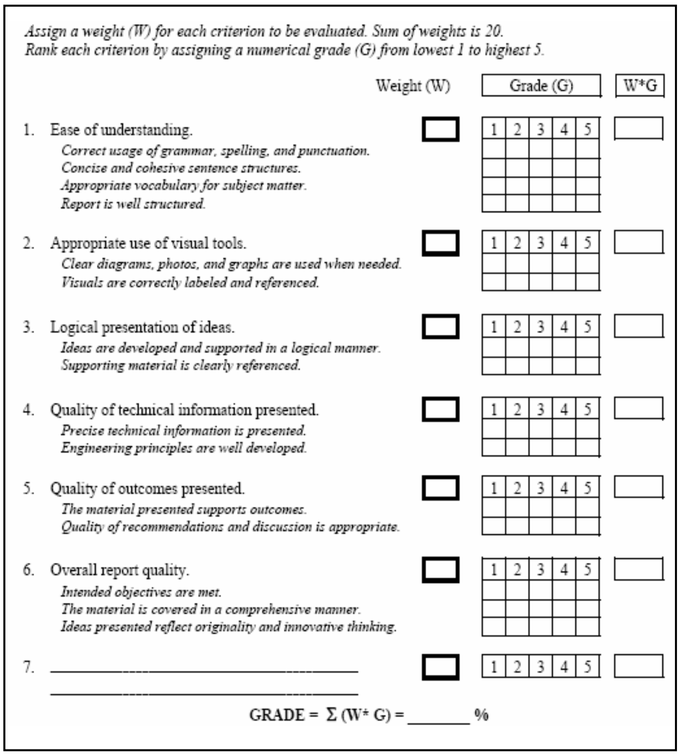
Table 1. Assessment Form Format

| PERFORMANCE SCORE FORM<br>Student: *XXX YYY* Date: ddmmyy | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| ***Ratings*** | ***W*** | RATING 1 | 2 | 3 | 4 | 5 | 6 | $G_n$*W |
| **Dimension A** | | | | | | | | |
| Attribute A1 | | | | | | | | |
| Attribute A2 | | | | | | | | |
| Attribute A$_n$ | | | | | | | | |
| **Dimension B** | | | | | | | | |
| Attribute B1 | | | | | | | | |
| Attribute B2 | | | | | | | | |
| Attribute B$_n$ | | | | | | | | |
| **Sum ($G_n$*W) x Total Allocated Marks** | | | | | | | | |

Dimension A, B…*n,* are the generic skills to be assessed; communication skills, resourcefulness, adaptability, etc. The attributes are finite skills within the dimensions. For example, in writing, it would be grammatical order, logic flow or reasoned arguments. A holistic discrete method of measurement is developed to enable the respective mean values, $\bar{x}$, for each generic skill to be established. These values serve as indicators; on the items easiness and give the locii on the quality level of the respective CLO [9].

Table 3 shows the simple computation used for an assessment. The lecturer will give his evaluation on the student's performance using the pro-forma form. He will give his own weightage, ***W*** for each dimension which he opines best for the subject. This allows flexibility and freedom for each lecturer to make his own evaluation. This is vital because the lecturer is free to set his own criteria of assessment and lets the student know what is expected from the assignment [10]. Next, each attribute is given a grade rating; between **1** to **6**, being the lecturer's assessment; **1** being ***poor*** and **6** for ***excellent***. Each rating grade has its own criteria. An example of such grade rubric amended with a scale of 6 is shown in Figure 1 below;

Figure 1. Assessment Grade Rating


*Assign a weight (W) for each criterion to be evaluated. Sum of weights is 20.*
*Rank each criterion by assigning a numerical grade (G) from lowest 1 to highest 5.*

Weight (W) | Grade (G) 1 2 3 4 5 | W*G

1. Ease of understanding.
   *Correct usage of grammar, spelling, and punctuation.*
   *Concise and cohesive sentence structures.*
   *Appropriate vocabulary for subject matter.*
   *Report is well structured.*
2. Appropriate use of visual tools.
   *Clear diagrams, photos, and graphs are used when needed.*
   *Visuals are correctly labeled and referenced.*
3. Logical presentation of ideas.
   *Ideas are developed and supported in a logical manner.*
   *Supporting material is clearly referenced.*
4. Quality of technical information presented.
   *Precise technical information is presented.*
   *Engineering principles are well developed.*
5. Quality of outcomes presented.
   *The material presented supports outcomes.*
   *Quality of recommendations and discussion is appropriate.*
6. Overall report quality.
   *Intended objectives are met.*
   *The material is covered in a comprehensive manner.*
   *Ideas presented reflect originality and innovative thinking.*
7. ______________________

GRADE = Σ (W* G) = _______ %

In a hypothetical scenario as shown in Table 2, subsequent to a grading, the attributes score for Dimension A is totalled up:

Grade Total A$_n$= 4+5+6 =15; Eq.(1)
$n$=total attributes

The grade total is then divided by the expected full score; G*$n$ = 6 x 3 nos. attributes, to give

Dimension A factor, $G_A = {}^{15}\!/_{18} = 0.83$ Eq.(2)

The raw score for Dimension A is obtained by multiplying the factor with ***W***, the given weightage of a particular dimension to generate the percentage score for the said dimension;

Dimension $A_{RS}$, $G_A$****W*** = 40x0.83 = 33.33% Eq.(3)

Finally, each dimension raw score is then summed up to determine the actual score the student obtained for his assignment;

Sum ($G_n$*$W$) x Total Allocated Score   Eq.(4)
= ((0.83 x 40) +(0.67 x 60)) x 6
= (33.33 + 40.00) x 6
= 73.33% x 6
= 4.40

Table 2. Simulated Assessment

| PERFORMANCE SCORE FORM<br>Student: *XXX YYY*   Date: ddmmyy | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| ***Ratings*** | ***W*** | **GRADE** | | | | | | **$G_n$*W** |
| | | **1** | **2** | **3** | **4** | **5** | **6** | |
| **Dimension A** | 40 | | | | | | | %G= $\frac{4+5+6}{18}=0.83$<br>$0.83X40=33.33$ |
| Attribute A1 | | | | | ✓ | | | |
| Attribute A2 | | | | | | ✓ | | |
| Attribute $A_n$ | | | | | | | ✓ | |
| **Dimension B** | 60 | | | | | | | %G= $\frac{3+4+5}{18}=0.67$<br>$0.67X60=40$ |
| Attribute B1 | | | | ✓ | | | | |
| Attribute B2 | | | | | ✓ | | | |
| Attribute $B_n$ | | | | | | ✓ | | |
| **Assignment Score= Sum (G*W) x 6** | | | | | | | | 33.33+40<br>=73.33%x6<br>= 4.40 |

The assessment is well structured based on dimensions and attributes, lending it reliable. The actual score is now more reflective of the students' generic skill ability rather than arbitrarily assessed. FKE allows the liberty at the lecturers hand to decide the type of generic skills they want to assess. However, they have to submit their proposed course outline indicating such assessment. The Heads of Department and the Programme Coordinators shall ensure all CLO's as determined in the programmes CLO Map is duly assessed. Table 3 exhibits a *pro-forma* student's Generic Skill Score Card for a given semester showing the mapped generic skills assessed.

Table 3. Engineering Students Performance Evaluation of Generic Skills (ESPEGS) Model

| **STUDEN T'S GENERIC SKILL SCORECARD**<br>**Name : Muhammad Saifuddin   Program : SEE   Sem.1/2005-06   PA: Rozeha A Rashid** | | | | | | |
|---|---|---|---|---|---|---|
| | *Learning Ability,* | | | | | |
| **Kursus** | ***$LO_4$*** | ***$LO_5$*** | ***$LO_6$*** | ***$LO_7$*** | ***$LO_8$*** | $\bar{x}$ |
| 1. SEE1023- Elektromagnetic Theory | eg. 33/40 = **0.8300** | 40/60= **0.6667** | - | - | - | 0.7083 |
| 2. SEE1123 –Signal & Network | - | - | 0.8500 | 0.7430 | - | 0.7965 |
| 3. SEE1223 –Comm. Principle | - | 0.7186 | 0.8475 | - | - | 0.7830 |
| 4. SEE1712 –Electrical Lab | 0.6250 | 0.6413 | - | 0.8475 | - | 0.7046 |
| 5. SSE1793 – Digital Numbers | - | - | - | - | - | - |
| **Ability, Mean ($P_\theta$) x 4** | **0.7275** x 4 | **0.6755** x 4 | **0.8488** x 4 | **0.7953** x 4 | | **0.7618** x 4 |
| **GSSC$\theta_x$** | $\theta_c$ 2.91 | $\theta_t$ 2.70 | $\theta_l$ 3.39 | $\theta_a$ 3.18 | $\theta_e$ - | **3.05** |

The ratio of scores for each dimension; say, Dimension A=$LO_4$ and B=$LO_5$ is transferred to it's respective columns :

Dimension A score = 33.33   Eq.(5)
Score ratio factor = $^{33}/_{40}$ = 0.8300, *from* Eq.(*2*)

The whole process is repeated to complete the scorecard by entering all the score ratio for each subject. This will result in a complete matrix of generic skills assessment of an engineering programme for further evaluation. This method of evaluation on engineering students generic skills development yields two (2) important data;

1. Value $\bar{\bar{x}}$ for each subject LO of a student;
2. Value $\bar{P}_{\Theta n}$ for each dimension.

Value $\bar{\bar{x}}$ of all the individual students is then summed up to give the CLO index as shown in Table 4. Value $\bar{P}_{\Theta n}$ serves as an indicator of each student's ability; strength and weakness in a particular generic skill. Remedial action can be taken effectively on each specific generic skill without any hesitation. As shown in Table 5, students can now track their generic skills development on the same scale ranking like the typical technical knowledge report; the base 4 point Cumulative Grade Point Assessment (CGPA).

To compute the Learning Outcome index, each student's score ratio average, $\bar{x}$, is obtained from Table 3. This is done by summing up the student's individual score and averaging the output to arrive at Student ***A*** generic skill score***;***

$$\bar{\bar{x}} = \text{Students Averaged Attribute Score} \quad \text{Eq. (6)}$$

$$= \frac{0.7275 + 0.6755 + 0.8488 + 0.7953}{4}$$

$$= 0.7618$$

The result is tabulated as in Table 4 to establish the subject CLO index. Once each student's generic skills score ratio is completed the LO index can now be established; hence,

$$\text{Subject CLO} = \sum_n \bar{x} \quad \text{Eq.(7)}$$

$$= \frac{0.7618 + 0.7806 + 0.7957 + 0.6926 + 0.7256}{5}$$

$$= 0.7513$$



ISSN: 1790-1979   595   Issue 8, Volume 5, August 2008

The result above is then multiplied with 100 to obtain the CLO index in percentage form. This number is easily interpreted and understood by many;

$$CLO_{index} = 0.7513 \times 100 \qquad \text{Eq.(8)}$$
$$CLO_i = 75.13\ \%$$

The subject learning outcome of 75.13% in Table 4 indicates that the achievement is commendable. A qualitative scale may be developed with descriptives like; exemplary, commendable, mediocre, poor, etc., to give direction whether necessary actions need to be taken to improve teaching instructions.

Table 4. Course Learning Outcome Index

| LEARNING OUTCOMES PERFORMANCE MEASUREMENT<br>Subject : SEE3512 EMT-1 Section: 12 Sem.1/2005-06 Lecturer: Rozeha A Rashid | | | | | | |
|---|---|---|---|---|---|---|
| | *Learning Ability* | | | | | |
| Student's Name | $LO_4$ | $LO_5$ | $LO_6$ | $LO_7$ | $LO_8$ | $\bar{\bar{x}}$ |
| 1.Muhd. Saidfudin | 0.7275 | 0.6755 | 0.8488 | 0.7953 | - | 0.7618 |
| 2. Student B | 0.7638 | 0.8236 | 0.7813 | 0.7536 | - | 0.7806 |
| 3. Student C | 0.8360 | 0.7468 | 0.7863 | 0.8136 | - | 0.7957 |
| 4. Student D | 0.6638 | 0.6918 | 0.7136 | 0.7014 | - | 0.6926 |
| 5. Student E | 0.6415 | 0.6839 | 0.8139 | 0.7632 | - | 0.7256 |
| OVERALL | 0.7265 | 0.7244 | 0.7888 | 0.7654 | - | 0.7513 |
| Learning Outcome Index; $Al_i$ = | | | | | *x 100* | 75.13% |

Given the set of raw score, it is now possible to establish a link between the students generic skills development and the CLO's. Table 5 shows the student's generic skills development over a study period. This overall longitudinal score, termed GSSC, shows some relationship between the outcome variable, $LO_i$ , and the student's ability, ( $\beta$ ). The students level of generic skills development is pretty cohesive through out the years. The raw score in Table 5 can be scrutinized further using probabilistic Rasch Unidimensional Measurement Model (Rasch Model) to measure their true development more accurately [11-13].

Table 5. Students Generic Skills Scorecard

| GENERIC SKILLS SCORECARD TRANSCRIPT<br>Name : Muhammad Saifuddin Program : SEE Sem.1/2005-06 PA: Rozeha A Rashid | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *Learning Ability* | | | | | | | | | | |
| Yr | Session | $LO_4$ | | $LO_5$ | | $LO_6$ | | $LO_7$ | | $LO_8$ | | GSSC |
| 1 | Sem 1-2004/05 | 2.91 | | 2.68 | | 3.39 | | 3.18 | | - | | 3.05 |
| | Sem 2-2004/05 | 2.96 | | 2.68 | | | | 3.39 | | 3.26 | | 3.05 |
| 2 | Sem 1-2005/06 | 3.05 | | 2.68 | | 3.39 | | 3.18 | | - | | 3.08 |
| | Sem 2-2005/06 | 2.75 | | 2.98 | | 2.68 | | 3.39 | | 3.18 | | 3.37 |
| 3 | Sem 1-2006/07 | 2.85 | | 3.06 | | 3.39 | | 3.18 | | | | 3.16 |
| | Sem 2-2006/07 | 2.85 | | 2.68 | | 2.75 | | 2.68 | | 3.39 | | 3.23 |
| | SUMMARY | $\theta_c$ | 2.87 | $\theta_t$ | 2.79 | $\theta_l$ | 3.12 | $\theta_a$ | 3.16 | $\theta_e$ | 3.28 | 3.15 |

## 4. Validation of Construct

A sample of comprehensive students assessment results, in this case EMT2523-Electro-Magnetic Theory, based on Table 4 is compiled and tabulated as shown in Table 6 for further analysis using Rasch software, ***Winsteps***. Students were coded as ST*nn*X. The study also delve into Differential Item Functioning (DIF) between genders; X coded 1 for Male and, 2 for Female.

This will enable the establishment of the discrimination index; the construct validity of the instrument in separating the students of different ability irrespective of gender or socio-economic background.

Table 6. Students Assessment Result: CLO Marks Tabulation

| Student | RAW SCORE BY LEARNING OUTCOMES | | | | | |
|---|---|---|---|---|---|---|
| | DIF | CLO1 | CLO2 | CLO3 | CLO4 | CLO5 |
| ST01 | 1 | 75.5 | 76.1 | 76.1 | 84.1 | 84.1 |
| ST02 | 2 | 70.7 | 71.4 | 64.0 | 69.1 | 69.1 |
| ST03 | 1 | 73.1 | 72.7 | 63.4 | 62.2 | 62.2 |
| ST04 | 1 | 58.8 | 60.4 | 55.9 | 57.8 | 57.8 |
| ST05 | 2 | 77.9 | 77.7 | 80.5 | 86.9 | 86.9 |
| ST06 | 2 | 77.5 | 79.0 | 76.9 | 85.9 | 85.9 |
| ST07 | 1 | 68.2 | 68.9 | 51.8 | 38.1 | 38.1 |
| ST08 | 2 | 83.2 | 82.8 | 66.5 | 60.6 | 60.6 |
| ST09 | 2 | 72.4 | 74.3 | 72.4 | 84.1 | 84.1 |
| ST10 | 1 | 65.8 | 65.4 | 62.7 | 68.1 | 68.1 |
| ST11 | 1 | 62.2 | 62.0 | 65.4 | 73.8 | 73.8 |
| ST12 | 2 | 86.3 | 87.1 | 85.5 | 90.6 | 90.6 |
| ST13 | 2 | 70.6 | 70.6 | 64.0 | 67.2 | 67.2 |
| ST14 | 1 | 78.8 | 80.3 | 84.9 | 94.7 | 94.7 |
| ST15 | 1 | 62.5 | 62.2 | 54.7 | 58.1 | 58.1 |

These raw score results are then transformed to numeric Grade Rating by cluster similar to the typical order rank A-E; in this case the following rating is used;

Grade Rating based on marks cluster:
>80 =**5** >70 =**4** >60 =**3** >50 =**2** >40 =**1** <40 =**0**

The grade rating is then tabulated in Excel **.prn* format as shown in Table 7. This numeric coding is necessary for further evaluation of the CLO achievement using Rasch Measurement Model.

In Rasch, the theoretical mean is deemed to be the probability of success. It is readily shown that a way to calculate the mean, $\bar{x}_i$ , is simply to take the proportion of cases with each score, multiply by the value of the score, and add them up; expressed in equation form as:

$$\bar{x}_i = \sum_{x=0}^{k} p_{xi} x_i \qquad \text{Eq.(9)}$$

*where*, k = maximum grade rating
$P_{xi}$ = proportion of event for each Grade Rating
$x_i$ = ascertained Grade Rating; n=1,2,$n_i$… $n_k$

The frequency proportion of events where the student obtained a certain Grade Rating is then established to compute the probability of achievement for each given CLO.

Table 7. Students Assessment Result Tabulation *.prn

| Student | RATED LEARNING OUTCOMES | | | | | |
|---|---|---|---|---|---|---|
| | DIF | CLO1 | CLO2 | CLO3 | CLO4 | CLO5 |
| ST01 | 1 | 4 | 4 | 4 | 5 | 5 |
| ST02 | 2 | 4 | 4 | 3 | 3 | 3 |
| ST03 | 1 | 4 | 4 | 3 | 3 | 3 |
| ST04 | 1 | 2 | 3 | 2 | 2 | 2 |
| ST05 | 2 | 4 | 5 | 5 | 5 | 5 |
| ST06 | 2 | 4 | 4 | 5 | 5 | 5 |
| ST07 | 1 | 3 | 2 | 0 | 0 | 1 |
| ST08 | 2 | 5 | 3 | 3 | 3 | 5 |
| ST09 | 2 | 4 | 4 | 5 | 5 | 5 |
| ST10 | 1 | 3 | 3 | 3 | 3 | 5 |
| ST11 | 1 | 3 | 3 | 4 | 4 | 5 |
| ST12 | 2 | 5 | 5 | 5 | 5 | 5 |
| ST13 | 2 | 4 | 3 | 3 | 3 | 3 |
| ST14 | 1 | 5 | 5 | 5 | 5 | 5 |
| ST15 | 1 | 3 | 2 | 2 | 2 | 5 |

**Grade Rating**: Marks >80=**5** >70=**4** >60=**3** >50=**2** >40=**1** <40=**0**

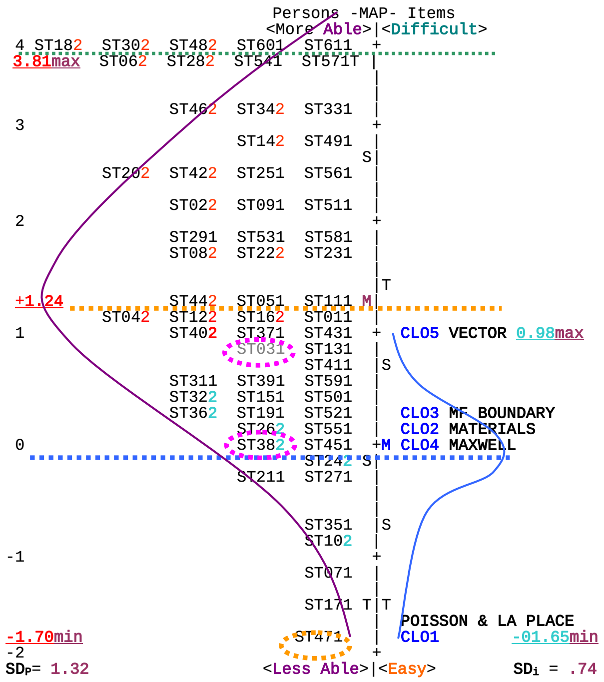
## 5. Findings

The data is processed using ***Winsteps***, a Rasch analysis software to conduct the necessary computation. First is to establish the Person-Item Distribution Map (PIDM). Figure 2 shows the Person=ST*nn* GenderX; i.e. student's n[th] series geographical position in relation to the test Item distribution; CLO's.

Note the similarity to the Traditional Histogram tabulation; but now both Person and Item distribution are put on the same *logit* scale in line with the Latent Trait Theory. PIDM is concerned with how likely a person $v$ of an ability $\beta$ on the latent trait is to respond to an item $i$ of difficulty $\delta_i$. This change of paradigm for measurement in education offers many value-adding features for potential development [14].

The instrument gives, the parameter $\delta$ the location of the item on the same trait: if $\beta_n$ is greater than $\delta_i$ then the person is likely to be able to respond to the item correctly [15, 16]. The degree of a person's ability is indicated by the separation of the item against the person's location on the map; the further the separation, the more able a person is likely to respond correctly to the said item. Similarly, the extent of an item difficulty is reflected by the spread of the item over the scale; akin to the high jump bar; the higher the location from the item mean, $Mean_{item}$, then the item is perceived to be more difficult as compared to an item on a lower location [17]. Thus, the $Mean_{item}$ serves as the threshold where it is set to zero on the *logit* scale.

Figure 2. Person-Item Distribution Map




It is observed in the PIDM that the cohort $Mean_{person}$=+1.24 which is higher than the threshold value, $Mean_{item}$=0 indicating these students cohort are of high ability. Further inspection of their Cumulative Point Average shows that they have an average of CPA above 3.30 which confirms the reliability of Rasch Measurement. The positive skew of the binomial graph further supports this observation. Only 8 students (13.11%) were found to be below $Mean_{item}$. These poor students generally have difficulty in achieving all the CLO's except CLO6-*Poisson and LaPlace*. By inference, the Lecturer can now take specific instructional measures of corrective action on the respective students. For example, a bridging tutorial can be arranged for this purpose.

The PIDM reveals that CLO7-*Vector* is the most difficult item encountered whilst CLO6-*Poisson and LaPlace* is the easiest item understood by the students. 37 students (60.66%) were found to have a good command over all the expected CLO performance. This map details out the relative position of each student ST*nn*X in relation to the respective CLO's; where ST*nn*X is Student n[th] serie of Gender X: Male is coded as 1 and Female, 2.

Take ST031-Male student; his PIDM shows that he has fulfilled all the CLO requirements except for CLO7-*Vector*; while ST382-Female student is having problem with CLO1-*Charge and Current*, CLO2-*Materials*, CLO3-*Magnetic Field Boundary* and CLO7-*Vector* respectively. The poorest student ST471-Male *logit* -1.70 is certainly a problematic Outfit whose score is even lower than item CLO6= *logit*-1.65: *Poisson and LaPlace*. On the other extreme, ST182, ST302, ST482-Female and ST601, ST611-Male are excellent students beyond *logit* 3.81.

The PIDM also allows us to scrutinise the gender performance; where 75.0% Females students (N=18) posess high ability above the $Mean_{person}$= +1.24 as compared to only 51.35% Male students (N=19). This is called Differential Item Function (DIF) showing the variance of learning ability between the genders.

Such detailed information is not available in typical Histogram reporting based on Traditional Test Theory [18]. Rasch Measurement is far superior and rich with specific information that enables the Lecturer to pinpoint the exact nature of the instruction problems and how the students progress over each course towards meeting the expected Programme Objectives.

Evaluation of Person-Item Correlation Order is done next. This is to establish the construct validity of the instrument of assessment. Table 5 shows interestingly a very close value between Cronbach-α and it's equivalent, Rasch Reliability, for items = 0.89 and 0.92 respectively on a scale of 1.0. Since the acceptable threshold value for Cronbach-α =0.6; these values validate the instrument construct validity as excellent. This is supported by Rasch separation index; equivalent to factor analysis, of an acceptable 3.46 on the scale of 6.0.

The mean in Rasch Model is the probability of success. As such, the level of CLO achievement can be taken as;

Raw Score mean =192.3

Expected full score = $N_{students}$ x max. rating; 56 x 5 = 280

$$Pr\ (\ Success\ ) = \frac{192.3}{280} \times 100$$, from Fig.5 Raw Score

Based on for Cronbach-α > 0.6 the overall CLO achievement = **68.68%** is therefore acceptable.

Subsequently, a check of Point Measure Correlation (PMC) gives the content validity of the items. The working parameter for an acceptable PMC value shall be between: $0.4 < x < 0.8$ [19]. CLO6- *Poisson and La Place* has the lowest PMC of 0.25 whilst CLO3- *Magnetic Field Boundary*, CLO4-*Maxwell's Law* and CLO5-*Stored Charges* is slightly above 0.8 which needs further evaluation (see Table 8).

Next is to check the corresponding Oufit values. The choice is obvious because it is easier to explain an outfit as compared to infit problems. The acceptable Root Mean Square (MNSQ) value is between; $0.5 < x < 1.5$. Any values beyond this range has implication for measurement where it distorts or degrades the measurement system and is unproductive for the construction of accurate measurement [10a]. Analysis shows CLO6 has a very high unexpected MNSQ=4.68 and a narrow margin MNSQ=0.45 for CLO3- *Magnetic Field Boundary*. This is further confirmed by the respective Z-standard score; the equivalent *t*-test, of a very high 5.1 and -3.4 respectively. This figures are well beyond the acceptable range of Z-STD value = $\pm 2$ [10b].

Table 8. Consolidated FKE SEE2523 EMT Learning Outcomes Student's Assessment: Persons Item Statistics:Correlation Order

INPUT: 61 Persons Items MEASURED: 61 Persons Items 6 CATS
Person:REAL SEP.:2.10 REL:.81 Item:REAL SEP.:3.46 REL.:.92
CRONBACH ALPHA (KR-20) Person RELIABILITY = 0.89 SD = 1.32

| ENTRY NUMBER | RAW SCORE | COUNT | MEASURE | MODEL S.E. | INFIT MNSQ | INFIT ZSTD | OUTFIT MNSQ | OUTFIT ZSTD | PTMEA CORR. | Item |
|---|---|---|---|---|---|---|---|---|---|---|
| 6 | 251 | 56 | -1.65 | .21 | 3.49 | 6.1 | 4.68 | 5.1 | .25 | L06 |
| 7 | 149 | 56 | .98 | .14 | 1.52 | 2.5 | 1.48 | 2.2 | .7[illegible] | L07 |
| 1 | 185 | 56 | .23 | .15 | .67 | -1.9 | .64 | -1.9 | .80 | L01 |
| 2 | 190 | 56 | .12 | .15 | .65 | -2.0 | .60 | -2.1 | .80 | L02 |
| 4 | 196 | 56 | -.02 | .15 | .59 | -2.4 | .55 | -2.4 | .82 | L04 |
| 5 | 196 | 56 | -.02 | .15 | .59 | -2.4 | .55 | -2.4 | .82 | L05 |
| 3 | 179 | 56 | .36 | .15 | .46 | -3.6 | .45 | -3.4 | .86 | L03 |
| MEAN | 192.3 | 56.0 | .00 | .16 | 1.14 | -.5 | 1.28 | -.7 | | |
| S.D. | 28.2 | .0 | .74 | .02 | 1.01 | 3.2 | 1.43 | 2.9 | | |

The results from Table 8 warrants an in-depth review of the items construct to ascertain the instrument validity in measuring what is it supposedly to measure; i.e. the Bloom's cognitive skills development as stipulated in the Table of Test Specifications. CLO6 -*Poisson and LaPlace* is further evaluated by Item Characteristic Curve technique to check the misfit data.

Figure 3 shows the Item Characteristic Curve indicating the spread of respondents against item difficulty.

It is found that CLO6 –*Poisson and LaPlace* is over-discriminating. The less able respondent were unexpectedly found to have scored very high beyond the 95% confidence interval whilst the more able respondent did not score as expected, hence, showing misfit data. Several possible reasons could have contributed to such anomaly or perhaps the assessment form itself need to be discarded.

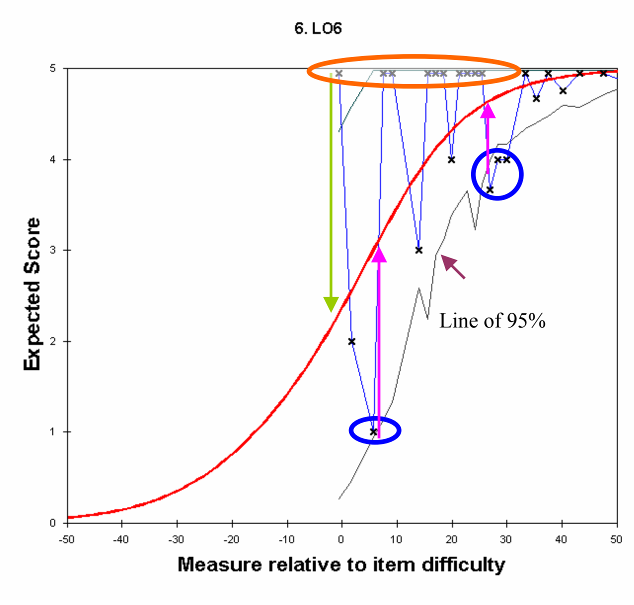



Figure 3 – Item Characteristic Curve of CLO-6: Poisson's and Laplace's equations

## 6. Discussion

There is a major mathematical flaw in the present method in computing students' performance using the class interval graded point. A conspicuous gap exists in the given grading which resulted in grotesque error in measurement. The present structure is shown below;

| Grade | Class Interval | CPA |
|---|---|---|
| A | ≥ 80 | 4.00 |
| B | ≥ 70 < 80 | 3.50 |
| C | ≥ 60 < 70 | 3.00 |
| D | ≥ 50 < 60 | 2.50 … |

CPA; Cumulative Point Average is a continuum scale with members set **P** = ( 0, .., 1, .., 1.98, 2, .., 3, .., 3.5, .., 3.75, .., 4.00). The gaps, which are described as fuzzy, materially affect the resulting CGPA, the Cumulative Grade Point Average, over the years. It is further aggravated by the fact that the final CGPA is highly dependant on the result of the first year CGPA. This is due to the fundamental mathematical flaw in the formula for calculating the CGPA itself where the denominator is smallest in the first year and it naturally becomes bigger towards the end of a student's study period. Consequently, if a student excels at the end of a program but is poor at the beginning, he will never obtain a good final CGPA. This flaw can be readily shown by applying a simple hypothetical student achievement data. This is certainly ironic because the significant content of any major in a particular course of study only appears towards the end of the study period. As such, the evaluation process is not measuring what it should be measuring. This indicates that the programme evaluation construct validity is questionable.

Matters get worst when the required measurement involves abstract and intangible items. The stochastic value is not readily available which renders the data unreliable.

Going back to the raw score as shown in Table 5, does not help either as this would revert to pure ordinal data. The percentage obtained as in Table 4, is only good to give a rank order. It gives an order of marks obtained based on count of correct answers hence representing ordinal data. Measurement does not occur here but rather it is merely a statistical representation. Raw score is not measurement. At best, the data summary can only be expressed by the median. Using just raw score fractions or percentages tends to group students around the middle scores and does not adequately contrast the results of the students with low and high abilities [20].

On the other hand, ESPEG Model is a comprehensive pro-forma evaluation for the required quality criteria grouped by dimensions and attributes which meets ABET evaluation requirement. ESPEGS is based on marks entries given during an assessment done by a lecturer for a given task. Thus, it provides an indicator on the item difficulty for each defined attribute; the generic skills to be accomplished. It allows subsequent further exploration of this evaluation method using Rasch probabilistic model.

Though closely related to Item Response Theory (IRT), Rasch Model was derived from a distinct set of fundamental postulates, and the most important concept is being specific objectivity [18]. It is the consequence of fundamental measurement principles deemed important and indispensable. The model is the one that meets Thurstone's requirement for useful measures and scale validity [21]. In the Rasch philosophy, the data have to comply with these principles, or in other words the data have to fit the model. 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. Whether the data will fit depends on many factors e.g., How good is your substantive theory? Does the latent variable actually exist? Are the items uni-dimensional?

If the data fail to fit the model, there is something wrong with the data. Resorting to a more general model and still claiming to measure something is obviously a self deception. Rather, in Rasch, exploration is made on the possibilities in what way could the data be wrong; possibly the items may not be good enough, the setting of the data collection may be inappropriate, or there may be more dimensions, etc. One should try to find out what causes the misfits by reasoned arguments rather than accounting for part of the misfit by extending the model; doing so by incorporating discrimination parameters which can lead to grossly skewed statistical outcomes. As a result, the fundamental principles are no longer applicable.

Rasch Model offers an excellent and comprehensive CLO assessment which can enhance the understanding of education policy alignment. In that regard, it assists educators in developing and maintaining quality education in Malaysian IHL that is aligned with the national interest; quality engineering education. Psychometrically, Rasch Model in its dichotomous case is expressed as follows:

$$\Pr\{ x_i = 1 \} = \frac{e^{\beta v - \delta i}}{1 + e^{\beta v - \delta i}} \qquad \text{Eq. (10)}$$

*where* $\Pr\{ x_i = 0,1 \}$, is the probability of a turn of event upon the interaction between the relevant person and assessment item;

e = Euler's number, 2.71828

$\beta_v$ = the ability of person $v$

$\delta_i$ = the difficulty of assessment item $i$

The expression yields a *S*igmoidal-curve, the Rasch Model, with locii indicating the person's ability for a given task. When responses of a person are listed according to item difficulty, from lowest to highest, it generates a likely pattern known as Guttman pattern or vector; i.e. {1,1,...,1,0,0,0,...,0}.

In this case, the probability of success can be simplified and re-written in *logit*, which is a Logistic Regression Linear Hierarchical Model [7]. The log odds, or *logit,* of correct response by a person to an item based on the Rasch Model can be simplified as:

$$\textbf{Logit}\ (P/1\text{-}P) = \beta v - \delta i \qquad \text{Eq.(11)}$$

From equation (11), it can be construed that the probability of a CLO is achieved is determined by the student's ability and the difficulty of the task given as shown in Figure 4.

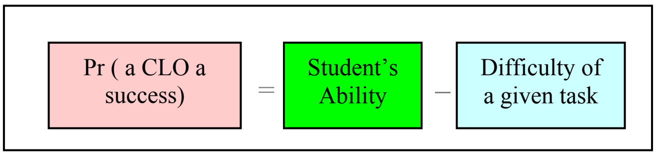



Figure 4. CLO Success Model

Therefore, the student's ability development may be tracked over their study period and the teaching method and style may be improved to facilitate such latent development. Symptoms of student's weaknesses in certain generic skill trait can be traced more effectively and easily. Similarly, students who excel can be equally tracked. The assessment form though dichotomous, used ordinal variables and they are not linear measures. By applying Rasch Model, this non-linearity problem is solved and an algorithm to link the correlation can therefore be established. This algorithm will help lecturers to give students the best experiential learning as well as provide a scale for the lecturers to improve their teaching competencies as well.

Rasch Measurement Model uses *logit* as the unit scale of measurement. A mathematical transformation is done to convert any numbers on a continuum scale to its natural logarithm which is an interval scale, ensuring the measurability of generated data. First, the raw score percentage is converted to its success-to-failure ratio or odds. Next, the scores are converted to their natural log odds or logits [20]. Simple index and logarithm can easily show that integers of raw score grow exponentially whilst $log_n$ increment is on a straight line.

Let's look at a number series; *n* ( 1,2,5.. 100) as shown in Table 9. The series is a function of a curve expressed as follows:

$$y = ax^n \qquad \text{Eq.(12)}$$

Simple log rule converts high order arithmetic into simple arithmetical functions of lower order using no more than the four(4) fundamental arithmetical operations, +, -, X, ÷ ;

$$\ln(y) = \log(nx) + \log a \qquad \text{Eq.(13)}$$

This bears fundamentally the same expression as a simple linear regression model;

$$y = mx + c \qquad \text{Eq.(14)}$$

The fuzzy values along the continuum scale, ln (y), are now a straight line with a definite interval.

The very reason why there is a need to transform observed data to *logit* is primarily to obtain an interval scale. The need for equal-interval scale was forwarded by Thorndike where "0 will represent just not any of the ability in question, and, 1, 2, 3, 4 and so on will represent amounts increasing by a constant difference" [22]. Mathematically, a series of numbers irrespective of the base used is not equally spaced but distant apart exponentially as the number gets bigger while a log series maintain their equal separation, hence, interval. This equal separation is shown in Table 9 and is termed as ***logit,*** being the unit of measurement for ability; in this case '**learning ability**'. The difference between $\log_{10}5$ and $\log_{10}2$ is constant and remain of equal distance between $\log_{10}50$ and $\log_{10}20$. Similarly, for $\log_e$ or *logit*.

Table 9. Comparison of Original numbers and Log intervals

| Original *n*-series | $\mathbf{Log_{10}}$ | $\mathbf{log_e}$ |
|---|---|---|
| 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 |

# 7. Conclusion and Recommendations

This simple but prudent conceptual theoretical framework of measurement is capable of examining an engineering CLO in great breadth and width. It is capable of providing multi-faceted views with specific and objective measurement on the CLO's established by an IHL.

This measurement model uses empirical data directly from the lecturers' assessment on a student for a given task. ESPEGS offers a more reliable measurement model where the results can distinguish the examinees more accurately. The statistical technique employed very fundamental statistical approach; mean and mean average, and simple linear interval measurement, thus, globally applicable.

ESPEG Model enables each CLO to be evaluated discretely. CLO index generated gives a fairer view on the program CLO achieved. Coupled with Rasch Model, the proposed measurement model is able to show reliably accurate result with small number N. However, the dimensions affecting the performance of a teaching method shall be subjected to further study.

***References***:

[1]. Razimah, A., Rozeha, AR, and Saidfudin MS., "Continuous Improvement in Engineering Education through Effective Value for Money Audits", *Proceeding, 2nd International Conference on Engineering Education and Training,* Kuwait; April 9 - 11th, 2007.

[2]. Azrilah A.A, Azlinah M., NoorHabibah A, Sohaimi Z and Saidfudin M.S., "*Appraisal of Course Learning Outcomes Using Rasch Measurement: A Case Study In Information Technology Education*", International Journal of Systems Applications, Engineering & Development, Issue 4, Volume 1, pp.164-172. 2007

[3]. "ISO 9001:2000, Third edition, Quality Management Systems - Requirements," International Organization for Standardization, Geneve, Switzerland. 15 Dec 2000

[4]. Crawford, K.M., Cox, J.F. and Blackstone, J.H., *Performance Measurement Systems and the JIT Philosophy,* APICS, Fall Church, VA., 1988.

[5]. Rozeha AR and Saidfudin, MS.; "Establishing a method of measurement towards developing prerequisite engineering students' generic skills"; *Proceeding of the 4th International Forum on Engineering Education(IFEE),* Sharjah,UAE; 2006.

[6]. Weybrew, Benjamin B.; *The ABC's of Stress: A Submarine Psychologist's Perspective,* Praeger Publishers, 1992

[7]. Razimah, A., Rosmawati, A.H., Rozeha, AR., Azami, Z. and Saidfudin MS, "Application of Logistic Regression Model in Rasch Measurement to establish a Performance Index: A case study in Audits on Malaysian Institutions of Higher Learning", *Proceeding of the 12th WSEAS International Conference on APPLIED MATHEMATICS*, Cairo, Egypt; Dec, 2007.

[8]. Razimah, A., Rozeha, AR., and Saidfudin MS, "Establishment of a Corporate Governance Effectiveness Performance Index using Rasch Measurement: Internal audit experience in Malaysian Institutions of Higher Learning", *Proceeding of IIUM International Accounting Conference IV*, Putrajaya, Malaysia; June, 2008.

[9]. Rozeha AR., Halim, A and Saidfudin, MS.; "The construct validity of Electrical

Engineering Students Assessment Method based on Bloom's Taxonomy: A Case Study", *Proceeding, 2nd International Conference on Engineering Education and Training,* Kuwait; April 9 - 11th, 2007

[10]. Lawrence M. Rudner, "Measurement Decision Theory", *Tutorial,* University of Maryland, College Park, USA, 2001.

[11]. Kaz Uekawa, Rasch model: The reason why it is impossible to understand and how it still is possible to understand. Downloaded from http://www.estat.us/id115.html as at 16 April 2007.

[12]. Zoubir, AM and Iskander, DR.; "*Bootstrap methods and Applications in Signal Processing*"; pp. 10-19. IEEE Signal Processing Magazine, July 2007

[13]. Thompson, D., DR.; "*Jacknifing Multitaper Spectrum Estimates*"; pp. 20-13. IEEE Signal Processing Magazine, July 2007

[14]. Kuhn, T., *The Structure of Scientific Revolutions,* Harvard University Press. 1970

[15]. Rozeha, AR and Saidfudin, MS. ; *"An Instrument for Examining the LO's of Electrical Engineering subjects: A Performance Measurement"*; *Proceeding of Pacific Rim Objective Measurement Symposium*, Taipei, Taiwan. 2007

[16]. Rozeha, AR, Ibrahim, M., Azami, Z. and Saidfudin MS, "Outcome based Education Performance Assessment: A Computational Model to measure Electrical Engineering subjects Learning Outcomes", *Proceeding of WSEAS 4th International Forum on Engineering Education,* Crete Island, Greece; July, 2007

[17]. Mok, M. and Wright, B; "*An overview to Rasch Model and Measurement*", University of Chicago, 2004

[18]. Marcoulides, G. A. (1999). *Generalizability theory: Picking up where the Rasch IRT model leaves off?* Mahwah, NJ, USA: Lawrence Erlbaum

[19]. Linacre, M.*"What do Infit and Outfit Mean-Square and Standardized mean?"*, Rasch Measurement Transactions 16:2 Autumn 2002

[20]. Bond, T. & Fox, C. (2007). (2nd. Ed.) *Applying the Rasch Model: Fundamental Measurement in Human Sciences.* New Jersey:Lawrence Album Associates.

[21]. Wright B.D. (1988). The model necessary for a Thurston scale. *Rasch Measurement Transactions*. 2(1), 2-3.

[22]. Thorndike, E. L. (1926). *Educational Psychology*. New York: Columbia University, Teachers College.