Research Journal of Applied Sciences, Engineering and Technology 6(19): 3520-3526, 2013
DOI:10.19026/rjaset.6.3555
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: October 17, 2012 Accepted: January 11, 2013 Published: October 20, 2013

Research Article

Insights into Engineering Education Learning Outcome's Assessment with Rasch Model

[1]Azrilah Abdul Aziz, [1]Mohd Saidfudin Masodi, [2]Faridah Mohamed Ibrahim, [1]Mohd Zaidi Omar and [1]Azami Zaharim
[1]Centre for Engineering Education Research, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, 43600 Bangi, Malaysia
[2]Centre for Foundation Studies, University of Malaya, 50603 Kuala Lumpur, Malaysia

Abstract: Education is a process in which people learn. Knowledge, skill or ability is gained during the learning process to better equipped students on career opportunities. Learning process takes place in academic institutions in which programs are developed for the various fields. Accreditation to these programs requires the academic institutions to adopt qualification framework. The Engineering Accreditation Council of Malaysia (EAC) adopts the American Accreditation Board of Engineering and Technology 2000 (ABET) requirements which promote Outcome Based Education (OBE) learning process. OBE calls for the evaluation of the subjects Learning Outcomes (LO) as specified in the Program Specification. Evaluation method has been largely dependent on students' performance carrying out tasks in tests, quizzes or submission of assignments. The evaluation on the students' performance output gives an indication on the students' achievement of the subject's LO. However, measurements of the students' achievement from the observed outcomes remain vague when raw scores are the only means of assessment. Raw scores may not provide true conception on difficulty of tests and or the true ability of students. This article described the measurement and analysis of the Engineering Mathematics LO's achievement using Rasch Measurement Model. It gives a more meaningful analysis on the students' achievement.

Keywords: Accreditation assessment, engineering mathematics, learning outcome, Rasch measurement model, OBE measurement.

INTRODUCTION

Learning is a complex process where it involves cognitive, emotional and environmental influences and experiences in acquiring knowledge or skill (Illeris, 2004; Ormrod, 1995). Assessment should reflect these understandings by employing a diverse array of methods, including those that call for actual performance, using them over time so as to reveal change, growth, and increasing degrees of integration. Such an approach aims for a more complete and accurate picture of learning, and therefore firmer bases for improving our students' educational experience (Astin *et al*., 1991).

A good assessment recognizes the value of information for the process of improvement (Deming, 2000; Mok and Wright, 2004; Wright, 1997). Assessment approaches should produce evidence that relevant parties will find credible, suggestive, and applicable to decisions that need to be made. The point of assessment is not to gather data and return "results"; it is a process that starts with the questions that involve data gathering and subsequent analysis. It is of utmost importance on the onset that this fundamental of measurement must be clearly understood. Analysis must be based on valid data and duly interpreted to generate a reliable report with meaningful information for prudent decision making towards continuous improvement of teaching and learning. Rasch analysis provides more meaningful analysis enabling better inference be made. Rasch measurement model will be used in this study to assess students' achievement on basic mathematics for engineering. The findings will also lead to students' classification based on their ability-achievement hence better management in improving their achievement towards meeting the targeted learning outcomes (Azrilah *et al*., 2007).

METHOD OF ASSESSMENT

This paper explores the Learning Outcome (LO) on an engineering mathematical subject which exposes students to the basic concept of mathematical engineering. The subject is expected to be able to balance between practical applications of mathematical equations and in-depth understanding of real situation (Wright, 1989). The LO is the specification of what students should learn as a result, after a specified period

Corresponding Author: Azrilah Abdul Aziz, Centre for Engineering Education Research, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, 43600 Bangi, Malaysia

Table 1: Example of students' performance report

Students	Final test	Quiz 1	Quiz 2	Quiz 3	Quiz 4	Quiz 5	Laporan	Excel	Maths	Overall total
LC008	90	90	100	40	80	45	100	90	70	87
LP001	88	100	92	35	80	75	82	90	75	85
LC007	88	100	100	35	60	70	95	80	75	85
LC003	85	90	92	50	80	75	76	90	80	83
LC005	82	100	100	70	60	80	50	83	70	78
LC004	77	100	92	40	60	75	71	80	70	75
LM002	73	100	75	60	50	75	56	67	75	71
LC006	73	100	100	40	70	0	64	83	70	71

of time in the program. Apart from the technical knowledge, students are introduced on report writing and logical thinking. The data include several tests and quizzes taken during the first semester of their engineering program. Apart from the test results, the assessment also includes assessment on report writing, students' skills on Microsoft Excel and mathematical knowledge.

Common practice of assessment would take percentages of each assessment components, including the final examination for the semester. The summation of the percentages will be the overall marks indicating students' performance. Students' with high marks are of good performance compared to those with low marks. However, it is always a difficult situation deciding who is better than another when any two or more of the students having similar total scores. Table 1 show example of students' performance result and decision have to be made on who has better ability when it involved similar total score.

Raw-score method of assessment provides unreliable facts about students' performances. The Overall Total is the count of all correct answers provided by the students when answering the tests and quizzes called raw scores. The scores do not take into consideration the students' ability on answering each task and do not consider the difficulty of each task (Azrilah *et al*., 2007; Wright, 1989). It would not be fair for those students with high marks but only able to answer easy tasks compared to those with high marks on difficult tasks. Surely the next question comes to mind is, how do we determine a task as easy or difficult? Rasch Model theorized (Rasch, 1960, 1961) that those tasks or what Rasch referred to as items, are of easy items when it can be answered by all students. The second theorem of Rasch Model is that, students with high ability should be able to answer all questions, while students with less ability may have some problems with difficult items (Azrilah, 2009; Bond and Fox, 2007). Rasch model will rearrange the pattern of responses according to Guttman pattern. In column-wise, data are sorted from most easy to difficult items and in row-wise, from most competent to less competent students. The probability of answering a yes or a no for each task (right/wrong responses) is modeled as a logistic function of the difference between the person and the item's parameter (Mok and Wright, 2004; Wright, 1989; Andrich, 2004; Zamalia *et al*., 2010). Refer to (1):

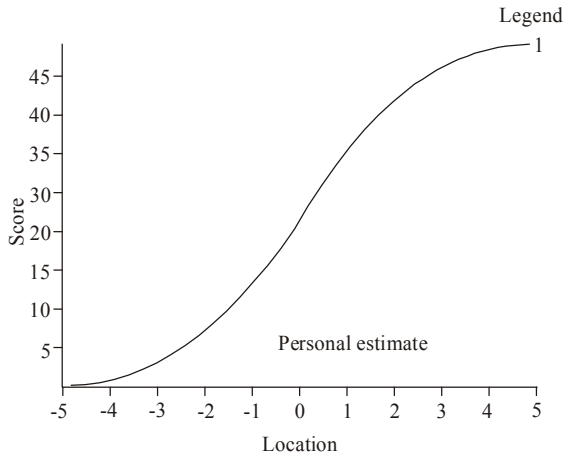

Fig. 1: Test characteristics curve showing the relationship between total score on a test and person location estimates

Probability of success = Person ability – Item difficulty

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

The relationship between total scores with ability estimates which is determined by the probability of success of that person to achieve an item, reveals a non-linear relationship. The shape of the Test Characteristics Curve (TCC) is ogival as depicted in Fig. 1.

Then Rasch model would scale the ability estimate according to the portion of correct responses. The smaller portion of the correct responses, the higher the difficulty of an item hence the higher the item's scale location. Both item and person are being scaled on the same or single scale. The higher a person's ability in relation to the difficulty of an item, the higher the probability of a correct response for that particular item. Person's location on the latent trait is equal to the difficulty of the item, when both the person and the item are located at the same location on the scale, there is a probability of 0.5 chance for a correct response according to Rasch model (Mok and Wright, 2004; Azrilah, 2009).

The purpose of applying Rasch model is to obtain measurement from categorical response data. In their raw form, categorical data cannot have mathematical

Table 2: Cronbach-A and Item Reliability

Valid responses: 99.9%
Person raw score-to-measure correlation = 0.99 (approximate due to missing data)
Cronbach alpha (KR-20) person raw score reliability = 0.33(approximate due to missing data)
Summary of 8 measured items

					Infit		Outfit	
	Raw score	Count	Measure	Model error	MNSQ	ZSTD	MNSQ	ZSTD
Mean	726.8	242.2	0.00	0.07	1.10	0.1	1.06	-0.2
S.D.	215.9	2.0	0.94	0.02	0.42	4.4	0.34	3.7
Max.	1105.0	243.0	1.82	0.10	1.91	5.9	1.06	4.3
Min.	329.0	237.0	-1.71	0.06	0.53	-7.3	0.85	-6.4

Real RMSE 0.08 ADJ.SD 0.94 Separation 11.67 Item reliability 0.99
Model RMSE 0.07 ADJ.SD 0.94 Separation 13.33 Item reliability 0.99
S.E. of item mean = 0.36
Minimum extreme score: 1 items
Umean = 0.000 Uscale = 1.000

Table 3: Person Reliability

Summary of 243 measured persons

					Infit		Outfit	
	Raw score	Count	Measure	Model error	MNSQ	ZSTD	MNSQ	ZSTD
Mean	28.9	9.0	-0.03	0.36	1.03	0.0	1.06	0.0
S.D.	3.8	0.2	0.48	0.02	0.58	1.2	0.84	1.1
Max.	40.0	9.0	1.62	0.52	3.13	3.1	7.00	4.7
Min.	16.0	8.0	-1.77	0.34	0.16	-2.9	0.18	-2.1

Real RMSE 0.40 ADJ.SD 0.26 Separation 0.66 Persons reliability 0.31
Model RMSE 0.36 ADJ.SD 0.32 Separation 0.89 Persons reliability 0.44
S.E. of person mean = 0.03

operations; addition, summation, and even mean. Therefore meaningful inference cannot be made upon categorical data. It only allows for data summarization using mode and median in reporting of the data (Leedy and Ormond, 2009). Rasch Model embodies certain criteria to be met in order to obtain measurement (Mok and Wright, 2004; Rasch, 1961; Bond and Fox, 2007; Andrich, 1988) and that the data should fit the model (Andrich, 2004) and not from the perspective of statistical modeling, model to fit the data.

ANALYSIS AND DISCUSSION

The Cronbach-α value is 0.33 reveals that the instrument might have some problems in assessing the students' performance. Rasch analysis allows further investigation, checking on the reliability of the person and item (Table 2). Item reliability is at 0.99, an excellent reliability value (Fisher, 2007). This indicates that the item size is sufficient to provide an excellent item spread with ranges of difficulty; there are items which are easy and some are difficult and some items in between.

It also reveals that there is one (1) item with minimum extreme score, pointing out that there is one (1) very easy item located at the end most of the measurement scale at -1.71 logit. The item is unable to demonstrate the variation of person ability; all those students who managed to answer the items correctly cannot be discriminated between a smart and a poor student.

The person reliability is rather low at 0.31, which indicate a poor spread of person; indicate a high level of inconsistencies in the students' responses (Table 3). This is further reveal by the Person Separation 0.66 which points out that the 243 person sample size has almost similar abilities but the instrument does not have enough good items to separate them into distinct group. More items is needed to be constructed to make the instrument better hence provide a laser cut to discriminate the students.

The poor value of person reliability however does not reveal that the instrument is not valid and analysis cannot be done. The Person Mean is located at -0.03 logit, indicating that the students have almost the same ability with the item difficulty (Item Mean = 0.00 logit) and the items in general are difficult to this group of students. Figure 2 depicts the location of items on the measurement scale, indicating their level of difficulties.

Observation made from the map in Fig. 2 showed that 4 of the items located above the mean item and the other 4 items located below the mean. Item KP which refers to chemical engineering appears to be the easiest item among those eight (8) items. KP is the minimum extreme item and is shown to be located at the lowest position on the logit measurement ruler in Fig. 2. ES-Electrical system, Excel-Microsoft Excel application, and *Laporan*-Report writing, are other items that are considered easy which is located below the item mean.

Maths -Mathematical concepts are the easiest among those items that are difficult and all four (4) are located above mean item. Final-the final examination,

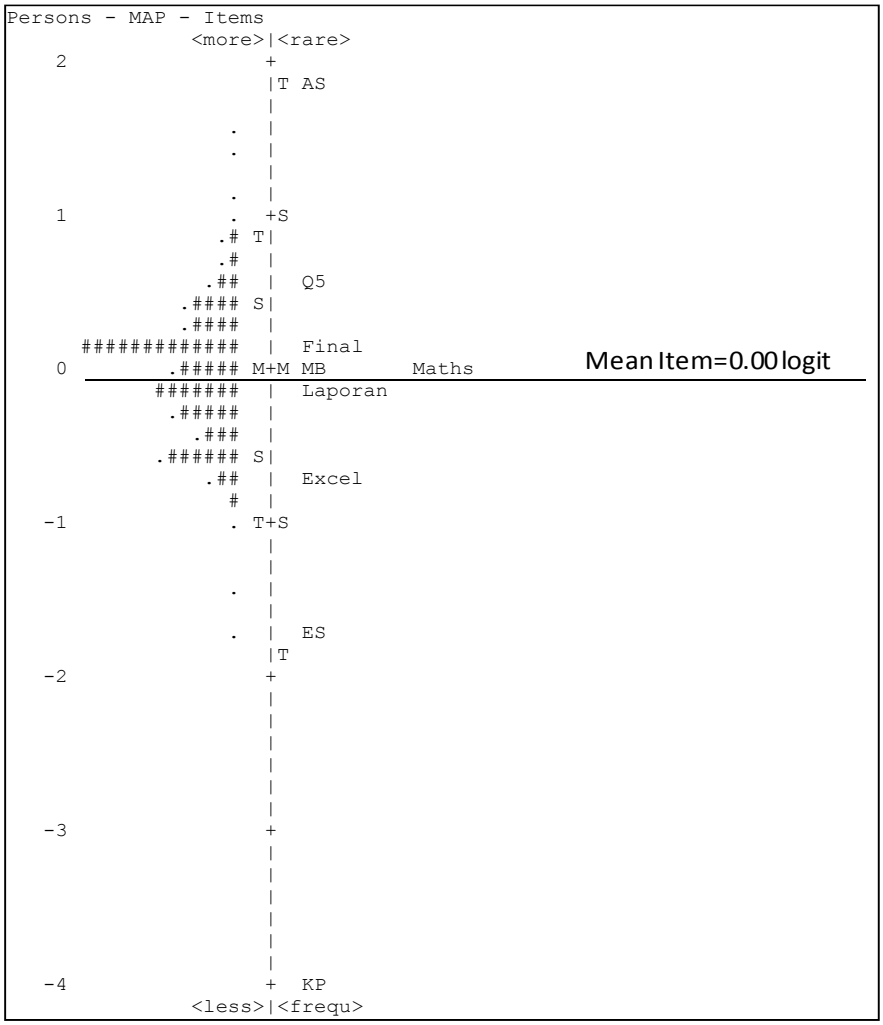

Fig. 2: Item map

Quiz 5 and AS -Structure and civil which is the most difficult item among the four (4).

Figure 3 is the person map where it depicts the location of each student according to their ability measure. In general, there is almost equal distribution of students above and below the item mean at 0.00 logit. The initial observation made to the summary statistics from Table 3, reveals that person separation is 0.66 near 1.00, indicating these group of students have almost the same ability thus a very strong homogeneity.

In common practice of assessment, those with high total marks are deemed have good command of the subject. The data used in this paper included the total marks of overall test percentages, and converted it into rating of 1<60, 2>60, 3>70, 4>80, 5>90. Logically, those with a '5' and '4' should be located on the upper location, above mean item = 0.00 logit, of the measurement scale on the map in Fig. 3. A detailed scrutiny of the Person-Item map is shown in Fig. 4 segregated between male and female student. Figure 4 showed female students' ability location. The achievement overall test rated from 1 to 5 is the last character (6th character) coded in the person demographic. There are some students with rating '2' and even '1' located on the upper portion of the map; above the mean item. Refer to person in the square in Fig. 4. Even though by summing up all the score of the test and quizzes gives those students low total score, to the contrary, Rasch finds them a more able student.

In this article, the data are of the scores for each tests and quizzes. It will be a more meaningful analysis if further investigation can be conducted on the items sub-construct of each of the tests and quizzes. Each tests or quizzes would have several sub-items (questions) within them, and usually it is based on the learning outcome of the Table of Test Specification, normally defined according to Bloom's level of knowledge. Then further findings can be explained according to the respective Learning Outcomes (LO); either the hierarchical achievement of the Bloom is according to the taxonomy or the reverse and confirms whether the students met the specified LO or not.

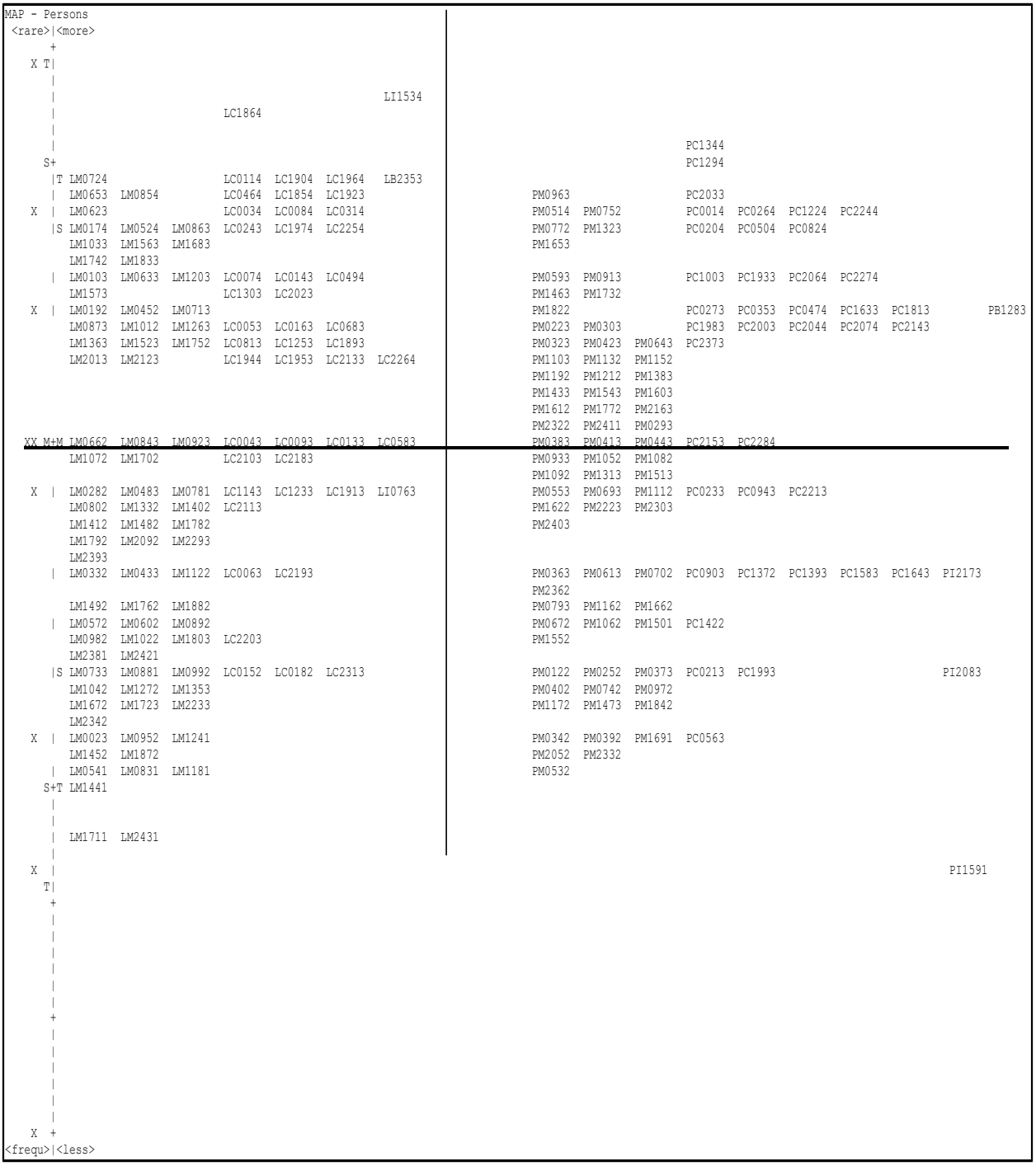

Fig. 3: Person - item map

CONCLUSION

The basic engineering mathematics subject introduces the students to the basic mathematics in engineering for all engineering students. Engineering students are exposed to the initial and necessary concepts for their preparation towards focused engineering field later in their engineering program. Besides assessing the students' knowledge on the subject, the assessment includes report writing, their knowledge on the Microsoft excel application and also on their mathematical skill.

The assessment reveals that the students find it easy in the area of chemical engineering and electrical engineering. The students also find it easy to handle report writing and have no problem with Microsoft

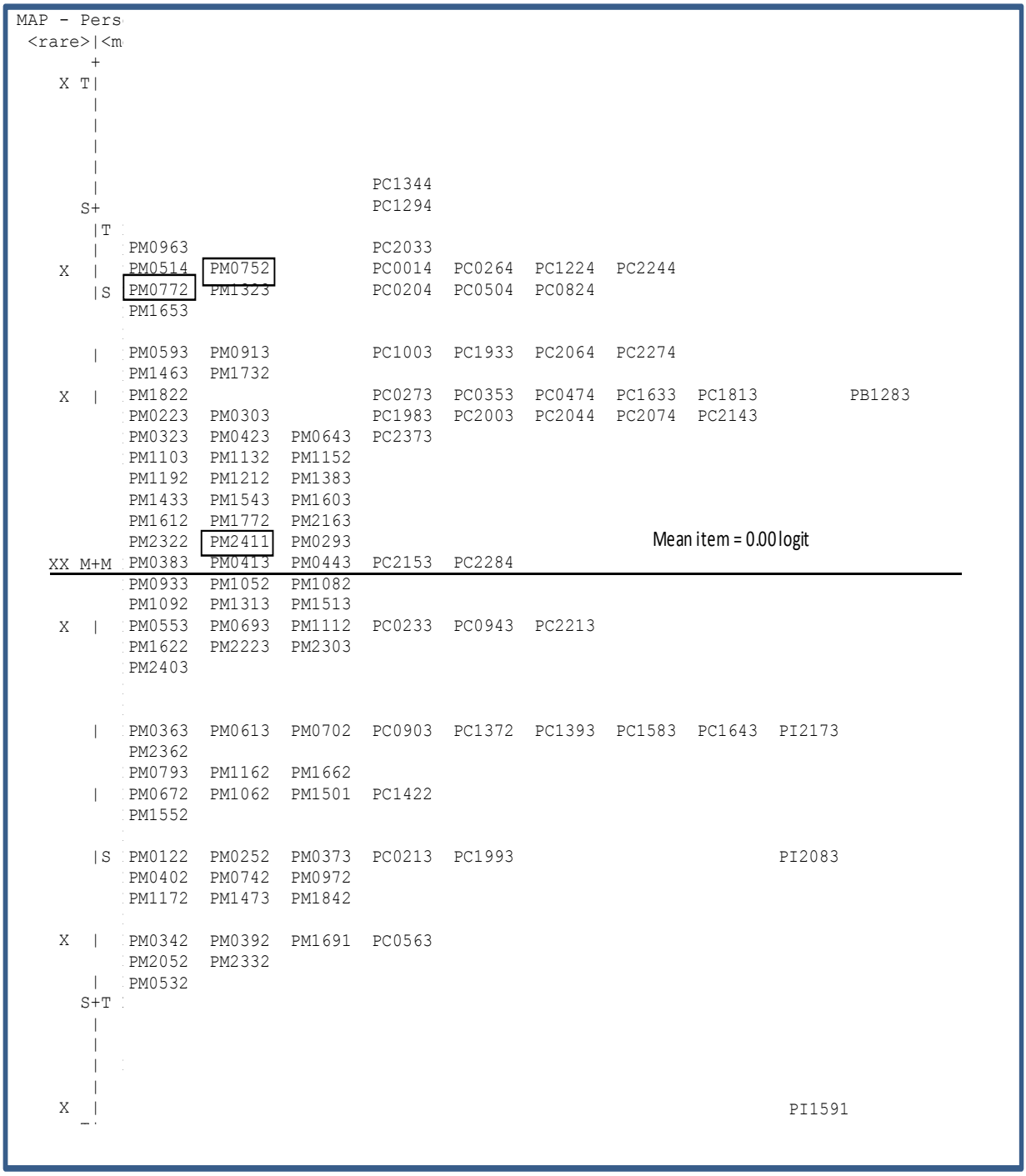

Fig. 4: Female distribution on person map

excel application. However, the students have some problem in basic mathematical knowledge, material science, Quiz 5, and find the final examination to be difficult. The most difficult item is on the area of civil and structural engineering, where all students are unable to solve the tasks given under this topic. There is definitely a need to identify the reason on why such preference is shown by this group of students.

ACKNOWLEDGMENT

The authors wish to acknowledge the financial support received from the Centre for Academic Advancement, Universiti Kebangsaan Malaysia as research grant; PTS-2012-088 and the Centre for Engineering Education Research, Universiti Kebangsaan in the effort of improving the quality of teaching and learning in engineering.

REFERENCES

Andrich, D., 1988. Rasch Models for Measurement. Sage Publications, Beverly Hills.

Andrich, D., 2004. Controversy and the Rasch model: A characteristic of incompatible paradigms. Medical Care, 42(1): 1-16.

Astin, A.W., T.W. Banta, K.P. Cross, E. El_Kawas, P.T. Ewell, P. Hutchings and B.D. Wright, 1991. Principles of Good Practice for Assessing Student Learning. The American Association for Higher Education.

Azrilah, A.A., 2009. Rasch Model Fundamentals: Scale Construct and Measurement Structure. ISBN: 978-967-5772-00-9 CIP: 519.54: Azrilah.

Azrilah, A.A., M. Saidfudin, M. Azlinah, A. Nor Habibah and Z. Sohaimi, 2007. Appraisal of course learning outcomes using rasch measurement: A case study in information technology education. Int. J. Syst. Appl. Eng. Dev., 1(4): 164-172.

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

Deming, W.E., 2000. Out of Crisis. MIT Press, Cambridge.

Fisher, W.P., 2007. Rating scale instrument quality criteria. Rasch Measurement Trans., 21(1): 1095.

Illeris, K., 2004. Three Dimensions of Learning Malabar. Krieger Publishing, Florida.

Leedy, P.D. and J.E. Ormond, 2009. Practical Research: Planning and Design. 9th Edn., Pearson, ISBN-13: 9780137152421.

Mok, M. and B. Wright, 2004. Overview of Rasch Model Families: In Intoduction to Rasch Measurement: Theory, Models and Applications. Jam Press, Minnesota.

Ormrod, J.E., 1995.Human Learning, Englewood Cliffs. Prentice Hall, NJ.

Rasch, G., 1960. Probabilistic Models for Some Intelligence and Attainment Tests. (Copenhagen, Danish Institute for Educational Research), Expanded Edition (1980) with Foreword and Afterword by B.D. Wright, the University of Chicago Press, Chicago.

Rasch, G., 1961. On general laws and the meaning of measurement in psychology. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, California, 4: 321-334.

Wright, B.D., 1989. Rasch model from counting right answers: Raw scores as sufficient statistics. Rasch Measurement Transactions, 3(2): pp. 62.

Wright, B.D., 1997. Fundamental measurement for outcome evaluation. Physical medicine and rehabilitation: State of the Art Reviews, 11(2): 261-288.

Zamalia, M., M. Saidfudin and A.A. Azrilah, 2010. Students Performance Evaluation for Learning Outcomes Measurement: SPELOM Model. In: Mastorakis, N., V. Mladenov, A. Zaharim and C.A. Bulucea (Eds.), Recent Advances in Management, Marketing, Finances: Recent Advances in Electrical Engineering, a Series of Reference Books and Textbooks. WSEAS Press, March 2010, pp: 172-178.