

Rasch Model Analysis on the Effectiveness of Early Evaluation Questions as a Benchmark for New Students Ability

Norhana Arsad^{1,2}, Noorfazila Kamal², Afida Ayob^{1,2}, Nizaroyani Sarbani^{1,3}, Chong Sheau Tsuey⁴, Norbahiah Misran^{1,2} & Hafizah Husain^{1,2}

¹ Centre for Engineering Education Research, Faculty of Engineering and Built Environment, Universiti Kabangsaan Malaysia, Selangor, Malaysia

² Department of Electrical, Electronic and Systems Engineering, Faculty of Engineering and Built Environment, Universiti Kabangsaan Malaysia, Selangor, Malaysia

³ Department of Mechanical and Materials Engineering, Faculty of Engineering and Built Environment, Universiti Kabangsaan Malaysia, Selangor, Malaysia

⁴ School of Psychology and Human Development, Faculty of Social Sciences and Humanities, Universiti Kebangsaan Malaysia, Selangor, Malaysia

Correspondence: Norhana Arsad, Department of Electrical, Electronic and Systems Engineering, Faculty of Engineering and Built Environment, Universiti Kabangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia. Tel: 60-389-118-363. E-mail: norhana@eng.ukm.my

Received: February 7, 2013 Accepted: March 25, 2013 Online Published: May 30, 2013

doi:10.5539/ies.v6n6p185

URL: <http://dx.doi.org/10.5539/ies.v6n6p185>

Abstract

This paper discusses the effectiveness of the early evaluation questions conducted to determine the academic ability of the new students in the Department of Electrical, Electronics and Systems Engineering. Questions designed are knowledge based - on what the students have learned during their pre-university level. The results show students have weak basic knowledge and this is in contrast to the results obtained during the application for admission to Year 1 of university. Thus, early evaluation questions were implemented to see the relevance in assessing the student's ability, obtained by the use of Rasch analysis, WinSteps. The findings show that the initial assessment is an effective and appropriate method to assess the ability of students, where the Cronbach- α is 0.69 and achieve the acceptable ranges of PT-Measure, Mean Square Outfit or Outfit Mean Square (MNSQ) and z-standardized values (ZSTD) Outfit. This shows that Rasch analysis can be used to classify the questions and the students according to their performance level accurately and thus, reveal the true level of the students' ability, despite the small number of samples.

Keywords: early evaluation questions, bloom taxonomy, rasch analysis, new students

1. Introduction

Education is the main driving force for the development of the nation. The growth in higher education institution is important not only to produce a person who is knowledgeable in any particular field, but most importantly, to produce one who has excellent soft skills such as thinking skills, communication skills, team work, problem solving skills and other skills that are essential to meet up to the challenge of the 21st century (Lee & Tan, 2003). At school, the education system employed is more text-book oriented. However, the education system at university seems to be more tailored to students who have the skills to search for, understand and analyze information critically. In addition, over the last decade, the university has shifted to more outcome-based education system, where problem-based learning has been introduced.

Another factor that contributes to the declining performance of the students is the learning environment. Back when they were at school, students were supervised and controlled by the parents or caretakers and there were also extra classes available for those could afford them. Meanwhile, for students who stayed at the hostel, they had been allocated some time to study at some specific time. This is different with the university environment, where students have more freedom in determining the kinds of activities that they want to concentrate in or do.

Therefore, with a different environment and study methods, very few students were able to adapt themselves to the university environment. Meanwhile, the rest of them faced the problem of adjusting to the new environment

so much so that their academic performances were affected.

Another most important skill that has to be acquired by UKM students is the learning skills (Mohamed Amin, 2010). It is also one of the contributing factors to academic performance. It has been found that students with poor study habits are more likely to withdraw from university or to have academic performance problems during the transition from secondary school to university (Pantages & Creedon, 1975; Abbott-Chapman, Hughes & Wyld, 1992).

Pre-university student results showed marked improvement from year to year, but some of these outstanding students result plunge at the university level. Why are these students not able to continue their prior excellence? Is the scale used in assessing the results of the students different at pre-university and university? Are the methods of teaching and learning (T&L) for the two institutions vastly different? A study was undertaken to answer all these questions.

For this preliminary study, the initial evaluation questions were formulated and used to measure the level of basic knowledge and the ability of new students enrolled in the Department of Electrical, Electronics and Systems Engineering (JKEES). This is to see the actual performance of these students before going further with their university studies. This study is important to look at other factors that can contribute to the deterioration of the performance of students while studying at university. In addition, the results also give an early warning so monitoring of weak students can begin as early as possible to ensure that students are continuously motivated and do not drift in activities that could distract them from their studies. As study conducted by Hafizah, Norbahiah, Norhana, Wan Mimi Diyana and Sarifah Nurhanum (2011) in Seminar Pendidikan Kejuruteraan & Alam Bina 2011 (PeKA '11) has found that the achievements of students from the Department of Electrical, Electronic and Systems Engineering (JKEES) is on the decline and some students have been disqualified from their degree program as a result to such performance.

However, are the questions set for the evaluation enough and suit with the objective of the conducted study? Are the questions fulfilled all required analysis factors? The appropriateness and effectiveness of the initial evaluation questions can be tested using Rasch analysis, using WinSteps software. This is important to make sure the questions are ready to get useful information in analyzing the performance of the undergraduate students who are excellent in their academic before. A good assessment recognizes the value of information for the process of improvement. Assessment approaches should produce evidence that relevant parties will find credible, suggestive, and applicable to decisions that need to be made, it is a process that starts with the questions of decision-maker that involves them in the data gathering and subsequent analysis (Saidfudin et al 2007).

Rasch analysis is a 'modern' alternative method of measurement that creates a measurement platform that matches the criteria of an SI unit where it acts as an instrument with a clear unit of measurement and can serve as a good model (Saidfudin & Ghulman 2009). Analysis of these measurements is done using empirical data obtained directly from lecturers' assessment for an assignment given to students and then converted to a logit scale having the same interval. Rasch uses logit as a measurement unit. Then the results are approximated to linear correlation. Rasch enables the movement from defining the concept of reliability from the 'most compatible data line' to produce repeatable measurement instruments that can be trusted. It is more focused on building a measurement instrument than loading data to adjust to the measurement model (Osman et al. 2011). Therefore, the results produced can give an accurate picture of the designed questions as a benchmark for new students which consequently would allow for follow-up actions to be taken by department.

2. Method

Initial evaluation questions were formulated as a result of discussions between the members of the Student Development Committee (JPPEL) of the department. Questions varied according to the Bloom taxonomy levels set for new students. Questions will be distributed to the new students during orientation week held a week before lectures begin. Students are required to answer all questions within one hour. A total of 34 new student admissions for Semester 1 2012 to 2013 in the Department of Electrical, Electronics and Systems Engineering, Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia took part in this initial assessment.

To evaluate the student performance, the question were arranged according to Bloom taxonomy level suitable for first year students as determined in the CO-PO list of the department. There are six domains from the simplest to the most complex; Tier-1: Knowledge, 2: Comprehension, 3: Application, 4: Analysis, 5: Evaluation, and 6: Synthesis: thus, everything can be measured.

Table 1 shows the learning topics and Blooms taxonomy domains assessed for each question. Part A:

Self-assessment consists of Question 1 only which reflect student’s ability in English language. Part B: *Basic Knowledge* contains 7 questions which tested the students’ basic understanding of electrical and electronic engineering, whilst Part C: *General Knowledge* examines the students’ general knowledge with 2 questions. All these questions can be found in (Norhana et al 2012).

Table 1. Topics and domains of Bloom Taxonomy assessed for each question

Part	Question	Learning topics	Bloom taxonomy
A	1	Explanation	Knowledge (K)
	2	General solution	Application (A)
	3	Specific solution	Comprehension (C)
	4	Specific solution	Comprehension (C)
	5	Difference	Application (A)
	6	Difference	Comprehension (C)
	7	Specific solution	Evaluation (E)
B	8	General solution	Knowledge (K)
	9	General knowledge	Knowledge (K)
C	10	General knowledge	Evaluation (E)

3. Results

Through Rasch analysis, via the WinStep software, the responses of the Person-Item Distribution Map (PIDM) which is obtained as shown in Figure 1, the reference to 'person' as the student and 'item' refers to the examination topics/ questions that are plotted on the left on the same logit scale in line with the Latent Traits Theory. PIDM considers how someone, x , with the capability of β as their latent trait, responds to those items, i , which have the difficulty level of δ (Saifudin et al. 2008). From the figure, we can get an overview of the overall or actual achievement for a student in their final exam later on in their study.

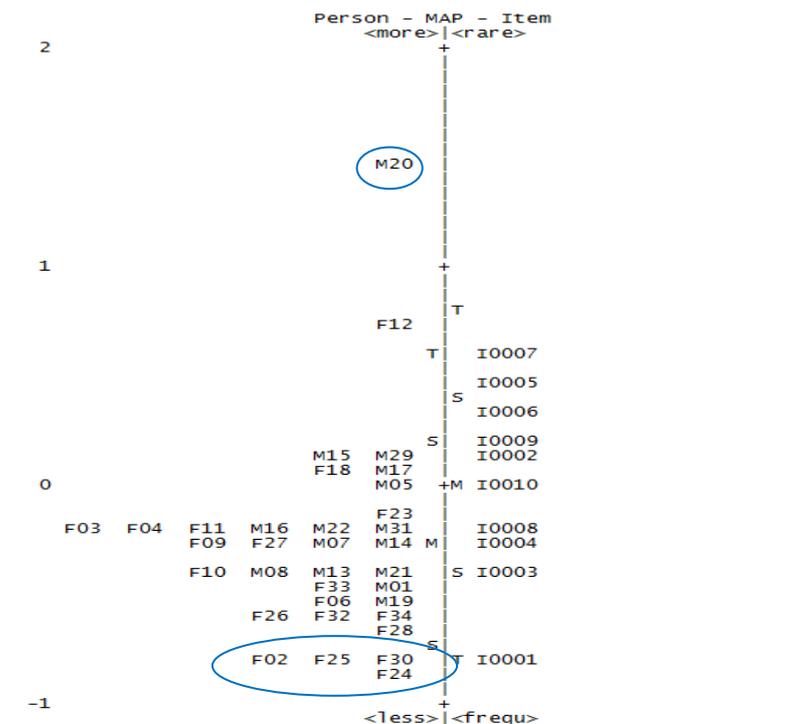


Figure 1. Person – Item map (Student-Question map) (PIDM) for the Entry Questions

The separation distance between the location of students and the questions on this map shows the level of one's ability, where the farther the separation of the responses given by the students, the more accurate the response would be. Meanwhile, the level of difficulty of the item or question is also shown, where the higher the location of the items from the mean, $Mean_{item} (+ M)$ means more difficult it is compared to the questions in the lower location. Thus, $Mean_{item}$ acts as a threshold and is set to be zero on the logit scale (Saifudin et al. 2008 & 2010).

Overall, the question is of uniform distribution, which is not too difficult and not too easy. Only one of the boys (M20) can answer all the questions the initial assessment. Through the admission application records, these students actually obtained very positive results during the pre-university examinations. Therefore, students will not have problems within the program and study in the department. However, girls achieve the lowest scores cluster as has been circled on the figure (bottom circle). This shows that girls only study topics in general and are less prepared when they are not given any guidance about the topics to be included in this evaluation question. Hence, these results can identify the capability of a particular student, and the members of the Student Development Committee in the department should take steps to prevent students from dropping out in their studies at university.

Figure 1 shows that almost half of the entry questions are found to be difficult by the students where item I0009 which is Question 9 Knowledge (K) - Current Knowledge is found to be the most difficult question. In this question, students were asked to describe the latest technology in the field of Electrical and Electronic Engineering. Meanwhile, item I0001 which is (Question 1 Analysis-understanding) is found to be the easiest question, as it is regarding the students themselves. This question can be ignored because it does not test the students' ability to think. Through question I0009, students' knowledge of the scenario occurring outside the scope of teaching in the university and students' attitude towards things that currently happen in the world can be assessed. This is important because the graduates of the department must meet the fourth program outcome (PO4) that is to be able to understand the professional and ethical responsibilities of environmental knowledge and contemporary issues.

Figure 2 shows the summary statistics for the category of students (people) and questions (items). From the table, the Cronbach α is 0.69, which exceeded the acceptable level of 0.6 (Saifudin et al. 2008 & 2010). Rasch analysis finds Person (Student) Reliability is low at 0.69, while Item (Question) Reliability is high at 0.89. This means that the instrument has high reliability in measuring what should be measured. Thus, the assessment questions outlined earlier is appropriate and effective in measuring the ability of the new students.

TABLE 3.1 C:\Users\USER\Desktop\try2.prn ZOU718WS.TXT Oct 24 10:23 2012
 INPUT: 34 Person 10 Item REPORTED: 34 Person 10 Item 6 CATS WINSTEPS 3.72.0

SUMMARY OF 34 MEASURED Person							
	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	ZSTD	OUTFIT MNSQ ZSTD
MEAN	17.9	10.0	-.27	.23	.96	.0	1.09 .2
S.D.	9.0	.0	.44	.06	.36	.8	.64 .8
MAX.	48.0	10.0	1.45	.51	1.88	1.9	3.18 2.3
MIN.	6.0	10.0	-.90	.19	.45	-1.5	-.37 -1.3
REAL RMSE	.25	TRUE SD	.37	SEPARATION	1.56	Person RELIABILITY	.69
MODEL RMSE	.23	TRUE SD	.37	SEPARATION	1.61	Person RELIABILITY	.72
S.E. OF Person MEAN	= .08						
Person RAW SCORE-TO-MEASURE CORRELATION = .99							
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY .69							
SUMMARY OF 10 MEASURED Item							
	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	ZSTD	OUTFIT MNSQ ZSTD
MEAN	60.9	34.0	.00	.12	1.00	-.3	1.09 .1
S.D.	33.3	.0	.41	.02	.39	1.7	.47 1.2
MAX.	132.0	34.0	.60	.17	1.61	2.2	1.99 1.7
MIN.	21.0	34.0	-.82	.10	.42	-4.0	.56 -1.8
REAL RMSE	.14	TRUE SD	.39	SEPARATION	2.85	Item RELIABILITY	.89
MODEL RMSE	.12	TRUE SD	.39	SEPARATION	3.14	Item RELIABILITY	.91
S.E. OF Item MEAN	= .14						
UMEAN=.0000 USCALE=1.0000							
Item RAW SCORE-TO-MEASURE CORRELATION = .99							
340 DATA POINTS. LOG-LIKELIHOOD CHI-SQUARE: 697.88 with 293 d.f. p=.0000							
Global Root-Mean-Square Residual (excluding extreme scores): 1.4455							

Figure 2. Summary statistics

Statistical analysis of the questions in Figure 3 shows that the entire entry question meets all the criteria as quality questions and thus no revision is required. This is because if the Point Measure Correlation (PT-MEASURE) = x ; $0.4 < x < 0.8$, the item shall be accepted. For item I0001 (Question 1) PT-Measure = $0.28 < 0.4$. This is expected to

be low because the question is about the students themselves. This question can thus be ignored and not included as the question in this assessment. It is only used as additional information to the department. Questions will only be considered to be unsuitable when they do not meet the range of PT-Measure, which are $0.5 < y < 1.5$ for Mean Square outfit (MNSQ) = y -value and the range $-2 < z < 2$ for the z -standard Outfit (ZSTD) (Saifudin et al. 2008). So, it can be seen from Figure 3, the initial assessment question has achieved the three criteria, thus fulfilling the suitability in assessing students, in this case after ignoring Question 1 that do not measure students' understanding and ability to think.

TABLE 13.1 C:\Users\USER\Desktop\try2.prn ZOU718WS.TXT Oct 24 10:23 2012
 INPUT: 34 Person 10 Item REPORTED: 34 Person 10 Item 6 CATS WINSTEPS 3.72.0

 Person: REAL SEP.: 1.50 REL.: .69 ... Item: REAL SEP.: 2.85 REL.: .89

 Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT		PT-MEASURE		EXACT MATCH		Item
					MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.	OBS%	EXP%	
7	21	34	.60	.17	.81	-.3	.66	-.5	.73	.60	64.7	63.2	I0007
5	26	34	.47	.15	1.61	1.5	1.99	1.7	.47	.59	41.2	49.2	I0005
6	34	34	.31	.13	.64	-1.2	.80	-.4	.59	.57	38.2	34.4	I0006
9	41	34	.20	.12	1.45	1.6	1.63	1.6	.34	.56	17.6	24.5	I0009
2	47	34	.11	.11	1.59	2.2	1.54	1.5	.42	.55	11.8	24.4	I0002
10	56	34	.00	.11	.87	-.6	.71	-1.0	.68	.53	26.5	22.2	I0010
8	75	34	-.20	.10	.81	-1.1	.72	-1.1	.66	.50	17.6	16.8	I0008
4	80	34	-.25	.10	.99	.0	1.36	1.4	.43	.49	23.5	19.1	I0004
3	97	34	-.42	.10	.42	-4.0	.56	-1.8	.56	.46	41.2	22.7	I0003
1	132	34	-.82	.12	.85	-.5	.94	.0	.28	.36	17.6	23.6	I0001
MEAN	60.9	34.0	.00	.12	1.00	-.3	1.09	.1			30.0	30.0	
S.D.	33.3	.0	.41	.02	.39	1.7	.47	1.2			15.4	14.1	

Figure 3. Question statistics

4. Discussion

The initial assessment questions used as an instrument in this study has been shown statistically by using Rasch analysis to be appropriate and has a high reliability in measuring the ability of new students enrolled in the department. The statistics has reflected on the academic achievements of the new students and also proved that the questions designed are effective in assessing the level of students' ability. The comparison between the results at pre-university and the early assessment questions can provide guidance in predicting the performance of students while studying at the department. Consequently, this enables the next steps that can be taken by the Student Development Committee of the department and student mentors in guiding students towards excellence in teaching and learning at university.

Acknowledgement

The author would like to thank UKM for providing the research grant (PTS-2012-102 and PTS-2012-091).

References

- Abbott-Chapman, J., Hughes, P., & Wyld, C. (1992). *Monitoring student progress: A framework for improving student performance and reducing attrition in higher education*. Hobart: National Clearinghouse for Youth Studies.
- Hafizah, H., Norbahiah, M., Norhana, A., Wan Mimi Diyana, W. Z., & Sarifah Nurhanum, S. S. (2011). Analisis Kuantitatif Pencapaian Akademik Pelajar JKEES. Seminar Pendidikan Kejuruteraan dan Alam Bina (PeKA'11). Retrieved from <http://www.ukm.my/p3k/images/sppb11/13.pdf>
- Lee, F. T., & Tan, J. H. (2003). *Project Management Skills Demand for Engineering Graduates in Malaysia*. Buletin Bulanan IJM: Jurutera. Edisi June 2003. Selangor: Dimension Publishing Sdn. Bhd. 16-25.
- Mohamed Amin Embi. (2010). Bimbingan Kemahiran Belajar. Pengalaman Pembelajaran Di UKM. Pengalaman Pembelajaran Di UKM 4, 57-67.
- Norhana, A., Noorfazila, K., Afida A., Sarzliana, S., & Hafizah, H. (2012). The Development of a Method for the Identification of Problematic Students during the Initial Stage of Study. *Asian Social Science*, 8(16), 1911-2017. Retrieved from <http://www.ccsenet.org/journal/index.php/ass/article/view/22704>
- Osman, S. A., Badaruzzaman, W. H. W., Hamid, R., Taib, K., Khalim, A. R., Hamzah, N., & Jaafar, O. (2011). Assessment On Students Performance Using Rasch Model In Reinforced Concrete Design Course Examination. *WSEAS Recent Researches in Education Proceeding*, 193-198. Retrieved from <http://www.wseas.us/e-library/conferences/2011/Penang/EDU/EDU-30.pdf>

- Pantages, T. J., & Creedon, C. F. (1975). Studies of college attrition: 1950–1975. *Review of Educational Research*, 48, 49-101. Retrieved from <http://www.jstor.org/discover/10.2307/1169909?uid=3738672&uid=2129&uid=2&uid=70&uid=4&sid=21102090504507>
- Saidfudin, M., Azlinah M., Azrilah A. A, Nor Habibah, A., & Sohaimi Z. (2007). Appraisal of Course Learning Outcomes using Rasch measurement: A case study in Information Technology Education. *International Journal of Systems Applications, Engineering & Development*, 4(1), 164-172. University Press, UK. Retrieved from <http://www.universitypress.org.uk/journals/saed/saed-22.pdf>
- Saidfudin, M., & Ghulman, H. A. (2009). Modern measurement paradigm in Engineering Education: Easier to read and better analysis using Rasch-based approach. *International Conference on Engineering Education. ICEED2009*. Retrieved from http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=5490624&url=http%3A%2F%2Fieeexplore.ieee.org%2Fexpls%2Fabs_all.jsp%3Farnumber%3D5490624
- Saidfudin, M., Azrilah, A. A., Rodzo'an, N. A., Omar, M. Z., Zaharim, A., & Basri, H. (2010). Easier Learning Outcomes Analysis using Rasch Model in Engineering Education Research. *EDUCATION'10 Proceedings of the 7th WSEAS International Conference on Engineering Education*, 442-447. Retrieved from <http://www.wseas.us/e-library/conferences/2010/Corfu/EDUCATION/EDUCATION-77.pdf>
- Saidfudin, M., Ghulman, H. A., Razimah, A., & Rozeha, A. (2008). Application of Rasch-based ESPEGS Model in Measuring Generic Skills of Engineering Students: A New Paradigm. *WSEAS Transactions on Advances in Engineering Education*, 8(5), 591-602. WSEAS Press. Retrieved from <http://trg.fke.utm.my/members/rozeha/publication2008/1j.pdf>