Research Journal of Applied Sciences 8 (10-12): 470-476, 2013
ISSN: 1815-932X

Investigation of Parallel Resistive Circuit PCUT-40 Instrument Construct Validity to Assess Engineering Students understanding in Electricity

[1]A.A. Azrilah, [2]A. Nazlinda, [1]M. Mohd Saidfudin and [1]A. Zaharim
[1]Centre for Engineering Education Research, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, Kebangsaan, Malaysia
[2]Faculty of Education, University Technology MARA, Selangor, Malaysia

Abstract: It has come to the attention of physics instructors that the engineering students enrolled in Malaysian universities do not possess sound conceptual understanding in basic electricity. Preconceptions and misconceptions regarding this subject matter especially on the topic of parallel resistors have been seen to exist amongst aspiring engineering degree students. Ironically, to date there are not many effective measures taken to arrest this issue. In addition, the existence of appropriate instruments within the local mainstream to measure these students' conceptual understanding of the related topic of parallel resistors is insufficient. Among the few claimed instrument was designed by Beh has been used to assess the students' conceptual understanding which has revealed some significant findings and led to some suggestions for the improvement of the current electricity syllabus and teaching instructional methods. However, recent developments within the Malaysian Institution of Higher Learning in science education system, call for a re-validation and improved measurement instrument. The aim of this research is to investigate the validity of this instrument, PCCUT_40-Parallel Resistive Circuit test using Rasch analysis in measuring the conceptual understanding of engineering students in institutions of higher learning in Malaysia on the topic of parallel resistors. The methods used are by purposive sampling on recently enrolled engineering students. The instrument focused on the six main domains which are "the meaning of parallel", practical knowledge of "current", "voltage", "resistance", followed by "circuit connection" and finally "the mental model". Data collected were analyzed using WINSTEPS version 3.71.0.1. The result shows that the instrument has 'excellent' item reliability of 0.94 logit, SE 0.33 and 'good' item separation of 3.82 logit, SE 0.30. Quality control procedures have resulted in an item reduction from 40 to only 28 items, thereby producing a better instrument in measuring the students' conceptual understanding of parallel resistors, the new PCCUT_28.

Key words: Learning outcomes, instructional objectives, performance assessment, engineering education, instrumentation, construct validity

INTRODUCTION

Physics educators, particularly in Malaysia have come to realize that engineering students in Malaysian Institutions of Higher Learning (IHL) do not have a sound basic academic background in electricity. They held on to the preconceptions and misconceptions that they have acquired throughout their secondary schooling days. Worst, these alternative conceptions were even seen amongst the local final year students and fresh young engineers. It seems that there were not enough measures taken to overcome the issue of students' lack of understanding in their academic subjects (McDermott and Shaffer, 1992; Metioui *et al*., 1996; Duit and von Rhoneck, 1998).

BACKGROUND OF STUDY

This gives a research motivation who is deeply concerned on the students' conceptual understanding on the topic of electricity with focus on parallel circuit. It has become obvious that there is a need for a method which can assist academia in identifying the students' level of understanding in basic electricity upon entrance into the respective IHL's. An effective method for this purpose would be a valid instrument which is able to assess and facilitate academicians in resolving the issue of incompetent engineering freshies in IHL. This instrument would be able to identify the areas of difficulties faced by the students. One such instrument that is available locally was designed by Beh, PCCUT_40-Parallel Circuit

Corresponding Author: A.A. Azrilah, Centre for Engineering Education Research, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, Kebangsaan, Malaysia

Understanding test with 40 items (Beh, 2000). However, in view of advances that took place in Malaysian education system overtime, there is a need for the instrument to be re-validated and tested for its reliability to be assure relevance for the current set of students. Hence, this study was conducted to re-validate PCCUT_40.

Respondents: The samples for this research were purposive where a selected group of engineering students studying in a local Malaysian university was identified. A random group of 60 students of mixed ability from poor to excellent from different engineering background and genders were chosen. There were 42 males (70%) and 18 females (30%) with 31 majoring in Electrical Engineering, 17 in Mechanical Engineering and 12 in Civil Engineering.

The instrument: PCCUT-40 was designed to test the students' understanding on the topic of basic parallel resistors. The instrument was divided in to six sections which basically are the major problems encountered by most engineering students. The six sections are:

- Section 1: meaning of the term parallel
- Section 2: practical knowledge of current
- Section 3: practical knowledge of voltage in parallel circuit
- Section 4: practical knowledge of resistance
- Section 5: practical knowledge of circuit connection: parallel and series
- Section 6: Mental Model

OVERVIEW OF RASCH MODEL

It is common practice for academicians to assess students' ability on an examination or test by performing calculations based on their correct responses and hence producing total raw scores (Saidfudin *et al*., 2011). These scores will provide the ranking order of the students who took the examination. A student who obtained 90% in an examination ranks higher than those who achieved 70% thus indicating that these scores are of ordinal data. Data placed in ranking order are not linear but is of continuum in nature. They do not have equal intervals which is necessary for statistical analysis. Hence, it must be put to practice in reading the score of 90% as the odd of success being 90:10. This type of data which is the ratio data is more appropriate for measurement purposes (Stevens, 2010).

Ratio data has a distinguishing feature which is the possession of a non-arbitrary zero value. The Rasch Measurement Model is a formulation that stipulates the relationship between a person and an item based on a mutual latent trait. To be more precise, it is able to predict the likelihood of a person of a given capability to correctly respond to an item of a certain difficulty level.

The probability of success depends on the difference between the ability of the person and the difficulty of the item (Bond and Fox, 2007). The Rasch Measurement Model is based on two fundamental theorems which says that:

- A person who is more capable has a greater likelihood of correctly answering all the items given
- An easier item is more likely to be answered correctly by all persons

In other words, the Rasch Model assumes that the item difficulty is the attribute that is influencing the person responses while the person ability is the attribute that is influencing the item difficulty estimates (Linacre, 1998). Fit statistics by Rasch analysis enable the researchers to see whether the data they are using is feasible; specifically higher ability students should be more likely to answer items of greater difficulty correctly than the lower ability students (Bradley *et al*., 2007). It includes the outfit and infit (mean square and standardized values) of the persons and items.

According to Greene and Frantom (2002), the term fit refers to "infit" (weighted by the distance between the person position and item difficulty) and "outfit" (an unweighted measure). If the data fit the Rasch Model then the expected values of the mean square and the standardized fit indices are 1.0 and 0.0, respectively. Most of the times, outfit is more sensitive to extreme responses compared to the infit.

Since, Rasch takes into account the relationship between the person's capability and item difficulty hence it is desirable to look at the person fit as well as the item fit. The person fit refers to an index which signifies the responses of an individual. An irregular or erratic response could be a sign of a misfit. Similarly, an item fit refers to an index which implies the functionality of the item. A misfit item means that the particular item is either too difficult or too easy for the respondent or it could mean that the item is not really testing on the desired latent trait. There are means of checking for quality control in rasch. In order to verify for fit and misfit items or persons, the following criteria must be satisfied:

- Point measure correlation: $0.32<x<0.8$
- Outfit mean square, $0.5<y<1.5$
- Outfit Z standard, $-2.0<Z<+2.0$

Method: The questionnaires of 40 items were administered to the students and the data were collected. The analysis of the data was done using rasch analysis Software (WINSTEPS 3.71.0.1). The idea of the test is to investigate on the instrument construct and investigate whether all 40 items are required during the final conduct of the research. This research carries all the necessary procedure in order to test for construct validity.

OVERALL FINDINGS

The first part of the analysis looked at the instrument fit to test whether the data obtained fit those that of the Rasch Model. In order to do so the item outfit statistics values were observed. The expected mean square value was found to be at 0.92 which is very close to the expected value of '1' and within the expected range of $0.5<x<1.5$ while the outfit Z-standard value for normality was found to be at -0.3, very closed to the expected value of 'zero' and within the normality range of $-2<Z<+2$. Hence, it can be said that the instrument has 'excellent' fit with item reliability of 0.94 (Fisher, 2007). The high item reliability of 0.94 indicates that the replicability of the items could occur if these items are to be given to another sample of the same size. The instrument yields an item separation of 3.82 which can be round of to 4 different levels of item difficulty: very difficult, difficult, easy and very easy. As for the item mean, it is set at 0.00 logit to ensure that each student has a 50:50 chance of success in responding to the item that matches their ability.

The Cronbach Alpha (KR-20) person raw score test reliability was used to test for the internal consistency of the respondents' responses to all the items in a measure (Sekaran, 2010). It was found that the Cronbach Alpha value for this pilot test was found to be at 0.98 which is considered as 'excellent' according to Fisher. This is an indication of an excellent internal consistency of the instrument in measuring the conceptual understanding of the respondents in the topic of parallel resistors. The summary statistics displays a person separation of 3.18. It means that the instrument objectively separate the students into 3 different heterogeneous profiling. As suggested by Greene and Frantom, the measured persons obtained are on a continuum rather than on a dichotomy.

The next part of analysis looked at the person mean. It was found that the person mean for the pilot test was found to be at -0.05 logit, SE = 0.30 which means that these students have slight difficulty in understanding the subject matter in general. To ensure that the person fit the rasch model reasonably well, the data need to fulfil the quality fit test. The Pilot test indicates that the outfit mean square value is at MNSQ = 0.99, S.D = 0.39 which is very close to the expected value of '1' and the person Z-standard value is exactly at '0' which is to be expected at the norm. Hence, it can be said that the persons or respondents for this pilot test do fit the rasch model very well. The above information can be extracted from Table 1.

The person-item distribution map: It has been said that the Person-Item Distribution Map (PIDM) is the heart of analysis in the rasch measurement model. Figure 1 displays the distribution of the persons and the items along a vertical ruler (dashed line) measured in logit. The

Table 1: Summary statistics for person and item

					Infit		Outfit	
Summary	Total score	Count	Measure	Model error	MNSQ	ZSTD	MNSQ	ZSTD
60 measured person								
Mean	38.1	80.0	-0.05	0.30	1.00	-0.1	0.99	0.0
SD	12.1	0.0	1.05	0.02	0.23	1.5	0.39	1.1
Maximume	64.0	80.0	2.39	0.36	1.60	3.6	1.99	2.9
Minimume	17.0	80.0	-2.01	0.28	0.63	-3.0	0.48	-1.9
Real RMSE	0.31	True SD	1.00	Separation	3.18	Person reliability		0.91
Model RMSE	0.30	True SD	1.00	Separation	3.34	Person reliability		0.92
SE of person mean	0.14							
36 measured item								
Mean	29.4	60.0	0.00	0.33	0.97	-0.3	0.92	-0.3
SD	13.7	0.0	1.36	0.06	0.19	1.5	0.34	1.6
Maximume	55.0	60.0	3.04	0.53	1.58	4.2	1.90	4.7
Minimume	4.0	60.0	-2.85	0.29	0.65	-3.3	0.48	-3.2
Real RMSE	0.34	True SD	1.31	Separation	3.82	Item	Reliability	0.94
Model RMSE	0.33	True SD	1.32	Separation	3.93	Item	Reliability	0.94
SE of item mean	0.23							

Electric circuits_pilot test (summary stats) Mar. 25, 2011; Input: 60 person 80 item reported: 60 person 40 item 2cats winsteps3.71.0.1; Cronbach alpha (KR-20) person raw score "Test" reliability: 0.98; Maximum extreme score: 4 item

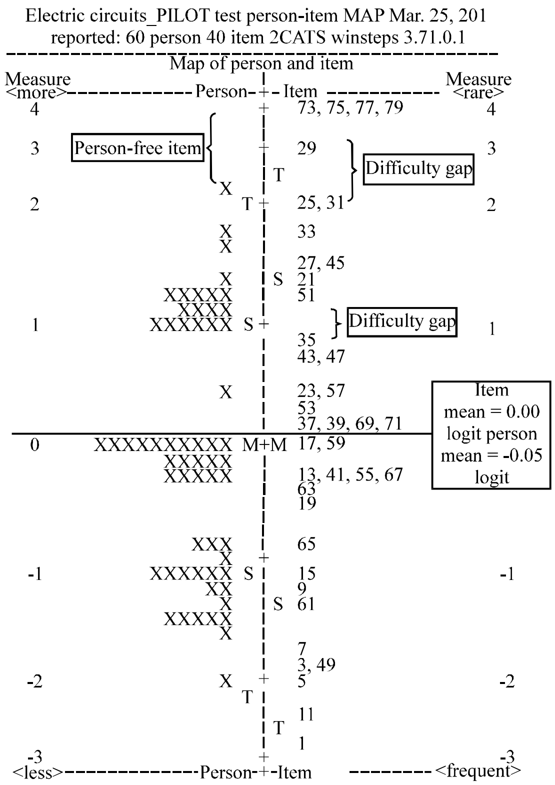

Fig. 1: PIDM item deletion using Point Measure Correlation (PMC), outfit mean square and Z-standard

vertical dashed line represents the ordering of the persons and items from less to best (bottom to top). The items are situated on the left and arranged from the easiest (bottom) to the most difficult (top). As for the persons, it can be seen that the ordering goes from the less smart (bottom) to the smartest (top of the vertical line). At the centre of the vertical line is the letter "M" which denotes the average of the mean for the item and the persons. The letter "S" reflects one standard deviation away from the mean while "T" indicates two standard deviations away from the mean. In Rasch Model, the scale has been set to zero for the item mean when the person's has a 50:50 likelihood of answering successfully (Saidfudin *et al*., 2011).

The items' spread is about +7 logit while the persons' ability stretches about +4 logit. This is an indication that this instrument contains some items which are beyond the current students' ability. As can be seen in Fig. 1, items 73, 75, 77 and 79 are situated at extremely difficult level with no respondents to the left of the map. Items 1 and 11 are situated at extremely easy level with no respondents as well. The distance between item 29 and items 25 and 31 is quite large (about +1.0 logit) with no item in between. However, there is a student situated on the right but with no item compatible to his/her ability. Therefore, there might be a need to add an item there. According to Baghaei (2008), in order to perform a uniformly precise measurement, the difficulty of the items should match the ability of the persons and the items should be reasonably spaced. Another obvious gap is between item 51 and 35. Perhaps more items can be added in this region to obtain a more precise estimate of the students' ability. As for the set of students, it is obvious that there are vacant spaces between them which mean that the instrument can even cater to more students.

Item deletion: Figure 2 displays all the items in the order of decreasing level of difficulty. There were 40 items originally but the analysis indicated that there are four extreme score items found. These four are those entry number 73, 75, 77 and 79. They can be eliminated since they are not assessing the students' understanding at all. A high stake test is a test with important outcome towards the student who took it. As suggested by linacre and bond, the cut off Point Measure Correlation (PMC) for a high stake test is at 0.32 or rather the range is $0.32<x<0.80$.

Point measure correlation is used to check that the string of response do make sense and reflecting the correct order of expected response. In other words, it is use to check on response validity. Hence, any items that do not meet these standards do not qualify to be in a high stake test and can be discarded. The items that have been identified not meeting the standards are identified in blue in Fig. 2. It can be seen, all 10 items have the PMC values out of the expected range. However, there are a few that can be highlighted upon removal.

For instance, item 55 has a PMC of -0.09 logit. The negative value gave an indication that there might be some strange responses pattern occurring for this item. An inspection on the scalogram confirmed this assumption where majority of the smart students gave incorrect responses to this easy item (-0.31 logit) while the mediocre and weak students are those who were able to provide the correct answers. Hence, item 55 can be discarded.

Another item to be discarded is item 29, a difficult item (measure order +3.04 logit) where the PMC is +0.23. An inspection of the scalogram confirmed the fact that the 4 students who answered correctly for item 29 were not the bright students but rather the mediocre ones. The other 2 items which are eligible for elimination are items 49 and 1. Both have low PMC values (+0.15 and +0.31, respectively). The scalogram shows that the low PMC value of item 49 was contributed by the mistakes done by the bright students while the weaker students were the

Electric circuits_PILOT test — ZOU383WS.TXT Mar. 25 21:59 2011

Input: 60 person 80 item — Reported: 60 person 40 item — 2 CATS WINSTEP 3.71.0.1

Person: Real SEP.: 3-18 REL.: .91...Item: Real SEP.: 3.48 REL.: .92

Item Statistics: Measure Order

Entry number	Total Score	Total Count	Measure	Model SE	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD	PT-meas corre.		Item
73	0	60	5.76	1.83	Maximum measure				0.00	0.00	ES6_a_M_P
75	0	60	5.76	1.83	Maximum measure				0.00	0.00	ES6_b_M_P
77	0	60	5.76	1.83	Maximum measure				0.00	0.00	ES6_c_M_P
79	0	60	5.76	1.83	Maximum measure				0.00	0.00	ES6_d_M_P
29	4	60	3.04	0.53	1.03	0.2	0.76	-0.1	0.23	0.23	ZS2_C1_bA3_P
25	9	60	2.04	0.39	1.17	0.8	1.02	0.2	0.21	0.33	ZS2_c1_bA1_P
31	9	60	2.04	0.39	0.92	-0.3	0.67	-0.7	0.42	0.33	US2_c2_cA1_P
33	11	60	1.77	0.36	0.93	-0.3	0.72	-0.7	0.44	0.36	US2_c2_cA2_P
27	13	60	1.52	0.34	1.01	0.1	0.83	-0.4	0.40	0.38	AS2_c1_bA2_P
45	13	60	1.52	0.34	0.95	-0.2	0.78	-0.6	0.44	0.38	US3_c_V1V2_P
21	14	60	1.41	0.33	0.69	-2.1	0.49	-2.0	0.67	0.39	ZS1_q1b_p1_P
51	16	60	1.20	0.32	1.27	1.7	1.63	2.1	0.13	0.41	ZS4_q2_i_P
35	19	60	0.90	0.31	1.14	1.0	1.31	1.4	0.27	0.42	KS2_c2_cA3_P
43	21	60	0.72	0.30	0.78	-1.8	0.66	-1.9	0.64	0.43	AS3_b_V1V2_P
47	21	60	0.72	0.30	1.08	0.6	1.28	1.4	0.34	0.43	US3_d_V1V2_P
57	23	60	0.54	0.30	1.27	2.0	1.47	2.4	0.18	0.44	ZS4_q3_ii_P
23	24	60	0.45	0.29	0.91	-0.7	0.86	-0.8	0.53	0.44	US2_c1_a_P
53	25	60	0.37	0.29	1.15	1.2	1.26	1.5	0.30	0.45	AS4_q2_ii_P
37	26	60	0.28	0.29	0.84	-1.4	0.84	-1.0	0.58	0.45	KS3_a_V1V2_P
39	26	60	0.28	0.29	0.92	-0.7	0.92	-0.5	0.52	0.45	AS3_b_V1_P
69	26	60	0.28	0.29	0.65	-3.3	0.57	-3.2	0.75	0.45	ZS6_c_P
71	26	60	0.28	0.29	0.69	-2.8	0.61	-2.8	0.72	0.45	ZS6_d_P
17	30	60	-0.06	0.29	0.83	-1.5	0.79	-1.5	0.60	0.45	KS1_1a_c9_P
59	30	60	-0.06	0.29	1.05	0.5	1.02	0.2	0.41	0.45	AS5_a_P
13	32	60	-0.22	0.29	0.82	-1.5	0.77	-1.6	0.60	0.45	KS1_1a_C7_P
41	32	60	-0.22	0.29	0.69	-3.0	0.64	-2.7	0.71	0.45	AS3_b_V2_P
55	33	60	-0.31	0.29	1.58	4.2	1.90	4.7	-0.9	0.44	AS4_q3_i_P
67	33	60	-0.31	0.29	0.80	-1.8	0.76	-1.6	0.62	0.44	ZS6_b_P
63	34	60	-0.39	0.29	1.09	0.7	1.08	0.5	0.37	0.44	ZS5_c_P
19	35	60	-0.47	0.29	1.02	0.2	0.97	-0.1	0.43	0.44	ZS1_q1b_srs_P
65	39	60	-0.82	0.30	0.74	-2.3	0.62	-2.1	0.66	0.42	ZS6_a_P
15	42	60	-1.10	0.31	1.01	0.1	0.83	-0.7	0.43	0.40	KS1_1a_ct_P
9	43	60	-1.20	0.31	0.88	-0.9	0.82	-0.7	0.50	0.40	KS1_1a_c8_P
61	45	60	-1.40	0.32	1.03	0.3	1.00	0.1	0.36	0.38	AS5_b_P
7	48	60	-1.73	0.35	0.95	-0.2	0.88	-0.2	0.39	0.35	KS1_1a_c4_P
3	49	60	-1.86	0.36	1.08	0.5	1.04	0.2	0.28	0.34	7 KS1_1a_c2_P
49	49	60	-1.86	0.36	1.16	0.8	1.67	1.6	0.15	0.34	AS4_q1_P
5	50	60	-1.99	0.37	0.85	-0.7	0.62	-0.9	0.47	0.32	KS1_1a_c3_P
11	54	60	-2.64	0.45	0.88	-0.3	0.48	-0.9	0.41	0.26	KS1_1a_c6_P
1	55	60	-2.85	0.48	0.94	0.0	0.64	-0.4	0.31		KS1_1a_c1_P
Mean	26.5	60.0	0.57	0.48	0.97	-0.3	0.92	-0.3			
SD	15.7	0.0	2.16	0.45	0.19	1.5	0.34	1.6		8	

Fig. 2: Item measure table

ones who answered them correctly. The Outfit Mean Square (Outfit MNSQ) value of item 49 shows an unexpected value of 1.67 which is not within the range of 0.5<y<1.5. Hence, it does not comply with the fit behaviour. As for item 1, it can be discarded since it is an extremely easy item (-2.85 logit) which is not really assessing the students' ability.

Items 25, 51, 57, 35, 53 and 3 are the other items possessing PMC values not within the accepted range of 0.32-0.80. They are eligible for elimination but an investigation conducted on each item proved otherwise. The removal of these items would disrupt the content validity of the particular section.

Redundant measurement: The analysis done by Rasch analysis has a mean of showing redundancy of items in the result. As suggested by Greene and Frantom, redundancy enables items reduction in order to reduce the length of instrument. Raw score residual correlations are used to detect dependency between pairs of items or persons (Table 2). Yen (1984, 1993) suggested a small positive adjustment to the correlation of size 1/(L-1) where L is the test length. A large positive correlation would mean local dependency.

Highly locally dependent items (Correlation>0.7), such as items entry number "31" and "33" share more than half their "random" variance, suggesting that only one of the two items is needed for measurement. Item pairs due for consideration are those having large standardised residual correlations such as items 31 and 33.

Items 39 and 43 are also scrutinized due to high proximity of Residual Correlation to +0.7 (these values are highlighted in the Fig. 2). An inspection on the PMC, Outfit MNSQ and Outfit Z-Standard shows that these items (items 31 and 33 and items 39 and 43) satisfy the quality requirements and hence do not need to be discarded.

Another form of redundancy can be seen from items that have “same measure and same domain” (Fig. 2). Any two or more items that have the same measure order and also testing on the same domain are not allowed to co-exist with each other since they are performing the same task at the same difficulty level. To avoid the redundancy, the item that has the lower quality (or lower point measure correlation) needs to be eliminated from the instrument.

In the case of item 25 and 31, both are displaying the same measure of +2.04 logit. Both of them are testing on two different conceptual understanding of electrical current. Item 25 is asking the students to provide the reading of the ammeters when one of the parallel resistors is removed while item 31 needs the students to identify the reading of the ammeter when a battery of similar voltage is added parallel to the existing one. By rule of thumb, item 25 which has the lower PMC of +0.21 compared to item 31 has to be discarded. However, an earlier analysis has identified item 25 to be one of the items to be retained due to its important contribution to the content validity of section 2. Therefore, in this case for items 31 and 25 are retained for the instrument.

The next pair of items that has same measure are items 27 and 45. Both are at +1.52 logit but these 2 items can be retained since item 27 is testing on electrical current while item 45 is about electrical voltage. Hence, there is no redundancy here. The situation is similar for items 43 and 47 where both have the same measure order at +0.72 logit. However, each of the items is testing on different conceptual understanding of the electrical

Table 2: Largest standardized residual correlations used to identify dependent items

Correlation	Entry number item	Entry number item
0.84	31 US2_circuit2_cA1_P	33 US2_circuit2_cA2_P
0.71	39 AS3_b_V1_P	43 AS3_b_V1V2_P
0.56	15 KS1_1a_circuit8_P	17 KS1_1a_circuit9_P
0.55	1 KS1_1a_circuit1_P	11 KS1_1a_circuit6_P
0.54	25 ZS2_circuit1_bA1_P	29 ZS2_circuit1_bA3_P
0.52	41 AS3_b_V2_P	43 AS3_b_V1V2_P
0.47	7 KS1_1a_circuit4_P	1 KS1_1a_circuit6_P
0.43	49 AS4_q1_P	53 AS4_q2_ii_P
0.41	43 AS3_b_V1V2_P	45 US3_c_V1V2_P
-0.41	55 AS4_q3_i_P	69 ZS6_c_P

Largest standardized residual correlations used to identify dependent items

voltage. Item 43 requires the readings of the two voltmeters when one of the parallel resistors is disconnected while item 47 requires the reading of the same voltmeters when the parallel resistors are replaced by parallel capacitors. Since, this instrument is testing on the conceptual understanding of parallel resistors and not on capacitors, it would be more appropriate that the item 47 which has a lower PMC of +0.34 than item 43 with PMC at +0.64 to be discarded.

As for items 37, 39, 69 and 71 all four of them have the same measure of +0.28 logit. A scrutiny of the measure order table (Fig. 2) shows that items 69 and 71 are both items related to the topic of mental models and in fact are the same circuit, drawn in a different manner. It was decided that item 71 is to be retained since it is more relevant to assess students’ capability compared to item 69.

As for items 37 and 39, both are items testing on the students’ understanding of the electrical voltage. The one that is not quality compliance needs to be taken out and in this case it would be item 39 since its PMC is only +0.52 compared to item 37 with +0.58 correlation value.

Another pair of items that needs scrutiny is items 17 and 59. Both have same measure at -0.06 logit. However, each of them relates to different domain with item 17 on parallel circuit diagram while item 59 on parallel circuit connection. Therefore, both can be preserved. Similarly, items 13 and 41 are also testing on two different domains with item 13 testing on the combine diagram of resistors while item 41 is testing on the electrical voltage. Both items need not be taken out from the instrument but it has been identified earlier that item 41 displayed a misfit behaviour where it did not comply with the normal response range and had to be discarded.

Both items 55 and 67 have same measures at -0.31 logit but they are not measuring the students’ understanding on the same domain. Item 55 is related to the electrical resistance while item 67 is on electrical mental model. According to this requirement, none should be discarded but item 55 has been identified as a misfit item earlier due to its irregularity of responses PMC = -0.09 and need to be taken out from this instrument.

The last pair of items that needs to be examined is items 3 and 49 where both items have a measure order of -1.86 logit. Item 3 is related to the parallel diagrams while item 49 is about the electrical resistance. Since, they are measuring the students' abilities on different domains, hence they do not need to be discarded. On the other hand, item 49 has been discarded earlier since it did not comply to the quality control by having a low value of PMC of +0.15 and a high Outfit MNSQ value of 1.67.

CONCLUSION

The results of analysis indicate that the person distribution fairly matched the items that measured the students' conceptual understanding of parallel resistors. In terms of refinements, several actions were conducted. Some of the existing items which were identified as "misfit" were eliminated in order to match the students' capability. A few of the items which do not fulfil the quality requirement were retained to ensure that the content validity of that specific domain has not been disrupted. In conclusion, the original PCCUT-40 test items has been identified as having 12 "misfit" items which do not fit the Rasch Measurement Model. The new validated instrument version; PCCUT-28 now consists only 28 items but having better measurement accuracy with smaller SE.

ACKNOWLEDGEMENTS

The researchers wish to acknowledge the financial support received from the Centre of Academic Advancement (PTS-2012-088) and Centre for Engineering Education Research, Universiti Kebangsaan Malaysia as research grant (OUP-2012-126) in the effort to improve the quality of teaching and learning in engineering education.

REFERENCES

Baghaei, P., 2008. The Rasch model as a construct validation tool. Rasch. Measure. Trans., 22: 1145-1146.

Beh, K.L., 2000. An investigation of the effect of instruction in physics on the formation of mental models for problem-solving in the context of simple electric circuits. Ph.D. Thesis, The University of Iowa, Iowa, USA.

Bond, T.G. and C.M. Fox, 2007. Applying the Rasch Model. Fundamental Measurement in the Human Sciences. 2nd Edn., Lawrence Erlbaum Associates Publishers, New Jersey, ISBN: 0805854622.

Bradley, K.D., J.D. Cunningham and S.O. Sampson, 2007. Evaluating and restructuring science assessment: An example measuring students conceptual understanding of heat. University of Kentucky 144 Taylor Education Building, Lexington.

Duit, R. and C. von Rhoneck, 1998. Learning and understanding key concepts of electricity. Connecting Research in Physics Education with Teacher Education. An International Commission on Physics Education. http://pluslucis. univie. ac.at/ Archiv/ICPE/C2.html.

Fisher Jr., W.P., 2007. Rating scale instrument quality criteria. Rasch Measur. Trans., Vol. 21.

Greene, K.E. and C.G. Frantom, 2002. Survey development and validation with the rasch model. Proceedings of the International Conference on Questionnaire Development, Evaluation and Testing, Charleston, November 14-17, 2002, South Carolina, pp: 42.

Linacre, J.M., 1998. Detecting multidimensionality: Which residual data-type works best? J. Outcome Measure., 2: 266-283.

McDermott, L.C. and P.S. Shaffer, 1992. Research as a guide for curriculum development: An example from introductory electricity. Part I. Investigation of students understanding. Am. J. Phys., 60: 994-1003.

Metioui, A., C. Brassard, J. Levasseur and M. Lavoie, 1996. The persistence of student's unfounded beliefs about electrical circuits: The case of Ohm's law. Int. J. Sci. Educ., 18: 193-212.

Saidfudin, M., A.A. Azrilah, N.A. Rodzo'an, M.Z. Omar, A. Zaharim and H. Basri, 2011. Use of Rasch Analysis to Measure Students Performance in Engineering Education. In: Latest Trends on Engineering Education, Dondon, P. and O. Martin (Eds.). WSEAS Press, Greece, ISBN: 978-960-474-202-8, pp: 435-441.

Sekaran, U., 2010. Research Methods for Business. 9th Edn., Wiely, New York.

Stevens, S.S., 2010. On the theory of scales of measurement. Science, 103: 677-680.

Yen, W.M., 1984. Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. Applied Psychol. Measur., 8: 125-145.

Yen, W.M., 1993. Scaling performance assessments: Strategies for managing local item. J. Educ. Measur., 30: 187-213.