International Journal of Innovation, Creativity and Change. www.ijicc.net
Volume 7, Issue 11, 2019

How Does the Rasch Model Justify Multiple Choice Question Items as a Measure of Student Understanding of Acid-Base Material at the Sub-microscopic Level?

Atiek Winarti[a*], Almubarak[b], Syubhan Annur[c], [a,b]Chemistry Education Department, Faculty of Teacher Training and Education, Lambung Mangkurat University, Indonesia, [c]Physics Education Department, Faculty of Teacher Training and Education, Lambung Mangkurat University, Indonesia,
Email: [a*]atiekwin_kimia@ulm.ac.id

The purpose of this study is to determine the item quality of acid-base assessment instruments on the sub-microscopic level using the Rasch model. This is achieved by identifying logit values and Rasch reliability components for each item (Multiple Choice Questions/MCQs). The method used was a quantitative approach. The data was collected by using the instrument which consisted of 15 items of acid-base assessment at the sub-microscopic level. Analysis of the data was conducted by using the Rasch model to probe the reliability of the item. Based on the data, the item reliability (0.99) of the assessment instrument on chemistry material at the sub-microscopic level can be stated as a measurement tool of the chemistry concept understanding especially on acid-base material. In addition, MCQs based at the sub-microscopic level was an alternative way of detecting student understanding in chemistry material. Therefore, the concept of justification of items with the Rasch model can be used as a reference for improving student understanding, quality of teaching and learning, and guidelines in designing chemistry learning strategies based on an analysis of student understanding.

Key words: *Rasch model, sub-microscopic level, multiple choice test, acid-base material.*

Introduction

Knowledge and understanding of chemistry according to Johnstone can be represented in three levels, namelyL macroscopic, sub-microscopic, and symbolic (Talanquer, 2018). The macroscopic level is an aspect of chemistry that is interpreted in a real and relevant way to everyday experience (Trivic & Milanovic, 2018). This macroscopic concept is generalised from the direct observation of natural phenomena or experimental results (Sihaloho, 2012). Observations and natural phenomena are called macroscopic such as burning candles, changes in colour, temperature, pH of the solution, gas formation and other phenomena (Rahmawan & Sukarmin, 2013).

Aside from the macroscopic level, there is the sub-microscopic level which includes particulate aspects as a principle in describing the dynamic movement of electrons such as atoms, ions, chemical bonds, and molecules symbolically in the form of microscopic images (Eliyawati, Rohman, & Kadarohman, 2018; Serobatse, Selvaratnam, & Drummond, 2014). The use of the sub-microscopic level, especially in learning chemistry, in addition to being able to accommodate students' understanding in describing microscopic content, can also show the ability of problem-solving in cases at certain levels, ranging from easy cases to the level of complicated problems. Unfortunately, the learning of chemistry mainly emphasises the macroscopic and symbolic levels, whereas the sub-microscopic level is less emphasised (Allred, 2019; Serobatse et al., 2014).

The symbolic level itself is a qualitative and quantitative representation of chemistry such as symbols, algebra, image shapes, chemical formulas, diagrams, and computerised chemistry such as media in describing something that happens (Rahayu & Nasrudin, 2014). The sub-microscopic level is part of the particulate level, which can be used to explain the dynamics of the movement of electrons, molecules, particles or atoms (Chittleborough, 2004). These three levels can be used to explain chemical concepts that are abstract, especially at the microscopic level, to include the small particles of a chemical (Allred, 2019; Chandrasegaran & Treagust, n.d; Gilbert, 2013).

The level described above is an important part that must be applied by teachers in learning chemistry. It was found that learning that did not emphasise the three levels was not able to foster students' critical thinking skills (Cloonan & Hutchinson, 2011; Liu, 2015). The impact is that students find it difficult to know and understand chemical concepts so they are prone to holding misconceptions (Indriyani, 2013). In other words, students easily interpret everything in the chemical aspect according to their own thinking without seeing whether the thought is scientific or not (Chandrasegaran & Treagust, n.; Treagust, Chittleborough, & Mamiala, 2003; Trivic & Milanovic, 2018).

Based on the analysis of the case revealed (Chandrasegaran, Treagust, & Mocerino, 2007), regarding the importance of deepening chemical material through microscopic, symbolic, and sub-microscopic knowledge levels, it is necessary to change the concept of chemistry learning both in terms of approaches, methods, application of learning models, strategies, and product-oriented development in an effort to reduce ambiguity and misconceptions for students. In another study, it was found that teachers must quickly take action so that misconceptions are not sustained; this was found to not only occur in schools, but also at the university level (Mubarak, 2016).

Identification of students' abilities in chemistry through multiple-choice tests is one appropriate strategy along with analytical techniques (Cheung, 2011; Cloonan & Hutchinson, 2011; Kilic, Sezen, & Sari, 2012; Mubarak, 2016; Sukor, Osman, & Abdullah, 2010; Villafañe, Loertscher, Minderhout, & Lewis, 2011). Assessment with the multiple-choice format is considered appropriate to be an instrument in diagnosing mental models and students' ability to understand the material as a whole. MCQs with distractors can further explore student understanding so that teachers know where misconceptions are occurring (Chandrasegaran et al., 2007; Lee, Liu, & Linn, 2011; Rauch & Hartig, 2010; Zamri bin Khairani & Bin Abd Razak, 2015).

Previous studies demonstrated that analysis using the Rasch model is considered an appropriate and effective measurement technique in representing both the ability of students to understand the material and the quality of questions created (Boone, 2016; Claesgens, Scalise, & Stacy, 2013; Johnson, 2013; Runnels, 2012; Talib, Alomary, & Alwadi, 2018). Thus, the Rasch model is the right strategy to use. The Rasch model is able to represent the ability of students and the level of difficulty of the problem so that it is a new strategy or technology in analysing student abilities and the quality of questions in one statistical data tabulation (Maier, 2007). Chiang’s study (2015) revealed that the Rasch model has advantages in which the tests carried out can be used as a basis for evaluating learning in the classroom. It is also able to analyse the ability of students through each item or answer patterns shown by students, and can identify the progress of learning of each student, so the teacher is able to know the weaknesses and strengths of each student in learning. It indicates that the Rasch model not only measures and or justifies students' cognitive abilities, but is also effective as a measurement tool so that this strategy can be used as a basis in improving the quality of teaching and learning chemistry (Amin et al., 2012; Chan, Ismail, & Sumintono, 2014; Chow , Tse, & Armatas, 2018; Claesgens et al., 2013; Herrmann-Abell & DeBoer, 2011; Johnson, 2013; Rabbitt, 2018; Runnels, 2012; Wei, Liu, Wang, & Wang, 2012).

Method

This study employed a descriptive, quantitative approach to diagnose the initial ability of the sample at the level of chemical sub-microscopic concept representation. From the pattern of answers and the structure shown through Rasch, it can be seen whether students' cognitive abilities have weak, moderate or high abilities at the sub-microscopic level. The position of the basic cognitive analysis specifies more how the character of students answers the questions and how the questions are able to be the basis of accurate measurements. The results of the Rasch model analysis can also be used to track the occurrence of students' misconceptions.

The Rasch model was used through WINSTEP software which detects in detail the ability and consistency of students' answer patterns in answering questions, so this greatly helps the teacher see the true abilities of the participants. The Rasch model was developed by Dr. George Rasch in the 1950s. In addition to ranking students' abilities, the Rasch model is also able to analyse the suitability between items and the participants' abilities, which can later provide information about misconceptions between items and participants (Sumintono, 2018).

The questions used as instruments for student assessments were in the form of multiple-choice with a total of 15 questions. In the Rasch Model, multiple-choice questions are called ‘dichotomies’, where Rasch modelling combines an algorithm which states the probabilistic expectations of items ‘i’ and respondent ‘n’, which are mathematically expressed as (Chan et al., 2014; Sumintono, 2018; Bambang & Wahyu, 2015):

$$\mathbf{P_{ni}\ (X_{ni} = 1 \mid b_n, d_i)} \quad = \quad \frac{e^{(bn-di)}}{1+e^{(bn-di)}}$$

Pni (Xni = 1| bn, di) is the probability of n respondents in item i to produce the correct answer (x = 1); with the respondent's ability, βn and item difficulty level of δi.

The above equation by Rasch can be further simplified by entering the logarithmic function and making it:

$$\mathbf{Log\ P_{ni}\ (X_{ni} = 1 \mid \beta_n, \delta_i) = \beta_n - \delta_i}$$

So, the probability of one success can be written as:
Probability of success = respondent's ability - a level of difficulty item.

The research sample for the trial of the instrument consisted of 191 students of the Chemistry Education Study Program, Faculty of Teacher Training and Education, Lambung Mangkurat University consisting of the first to the fourth-year students.

Research Findings and Discussion

The diagnosis made in this study detects students' abilities at the level of sub-microscopic representation as presented in Table 1.

Table 1: Statistics Items/ Measure Orders: Students' Sub-microscopic Understanding

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
15	17	191	3.74	.35	.88	-.4	.56	-.7	.49	.46	95.5	94.6	Q15
14	22	191	3.23	.29	1.00	.1	2.03	1.7	.44	.47	92.7	92.0	Q14
11	53	191	1.54	.20	.89	-1.2	.87	-.3	.55	.52	82.5	79.4	Q11
13	60	191	1.27	.19	.94	-.6	.75	-.7	.56	.53	78.5	77.8	Q13
12	65	191	1.08	.19	.85	-1.8	.66	-1.1	.59	.54	81.9	76.8	Q12
7	66	191	1.05	.19	1.01	.1	.81	-.5	.55	.54	73.4	76.7	Q7
9	86	191	.36	.18	1.01	.1	.85	-.6	.57	.57	71.8	74.5	Q9
6	87	191	.33	.18	.85	-2.0	.74	-1.1	.62	.57	78.0	74.5	Q6
8	69	154	.31	.21	1.02	.3	1.13	.5	.58	.59	79.3	75.3	Q8
4	100	191	-.11	.19	.85	-1.8	.71	-1.5	.63	.58	76.3	75.2	Q4
10	102	191	-.18	.19	.93	-.9	1.24	1.2	.60	.58	80.8	75.5	Q10
2	145	191	-1.93	.23	1.14	1.0	5.59	7.3	.50	.61	87.6	85.4	Q2
3	153	191	-2.38	.25	1.30	1.9	2.94	3.5	.50	.61	83.1	87.3	Q3
5	166	191	-3.35	.30	.63	-2.1	.46	-1.0	.67	.59	93.2	91.7	Q5
1	178	191	-4.96	.46	1.39	1.2	9.90	5.1	.40	.55	96.6	96.6	Q1
MEAN	91.3	188.5	.00	.24	.98	-.4	1.95	.8			83.4	82.2	
S.D.	48.0	9.2	2.25	.08	.18	1.2	2.48	2.5			7.7	7.9	

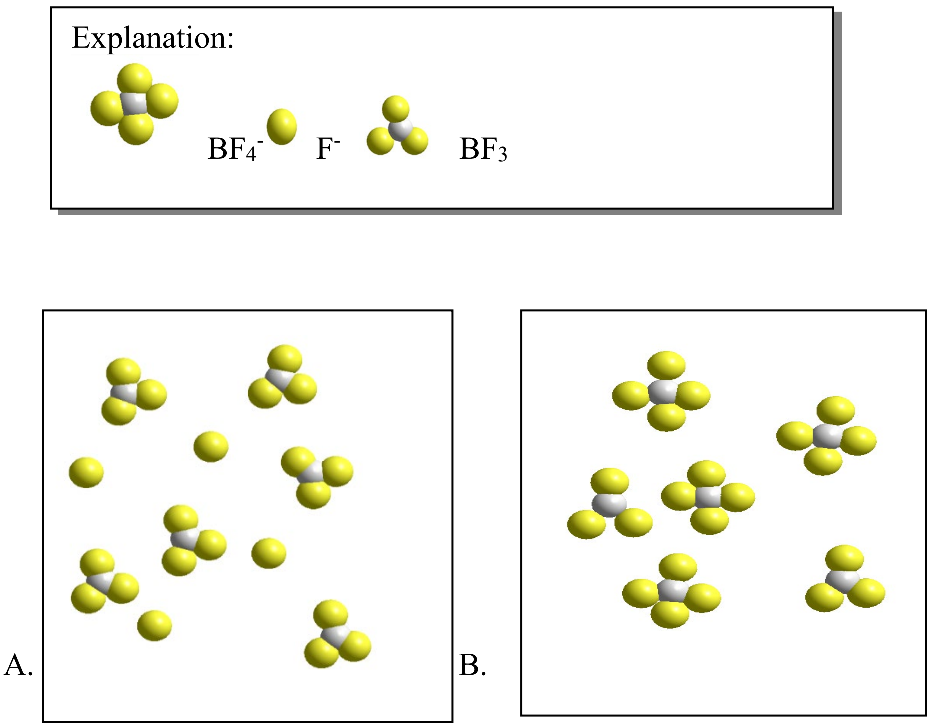
The data in Table 1 shows the level of difficulty of the questions by comparing the *logit* values that are owned by each dynamic. It appears that items of numbers 14 and 15 have the highest difficulty level with *logit* values of 3.27 and 3.74. In addition to the logit value data shown above, the 'total score' column is the number of samples that answered the item correctly so that it was found that the Q15 item and Q14 respectively only 17 and 22 students who answered correctly from 191 students. If examined in other numbers such as Q1, there are 178 students who answered correctly so that it is appropriate if items Q15 and Q14 are declared as items that have the highest level of difficulty. That is, students' sub-microscopic understanding of these items is considered weak so that this data can be a guideline for every instructor in designing learning concepts that can improve students' sub-microscopic understanding, especially at all aspects of multi-chemical representation.

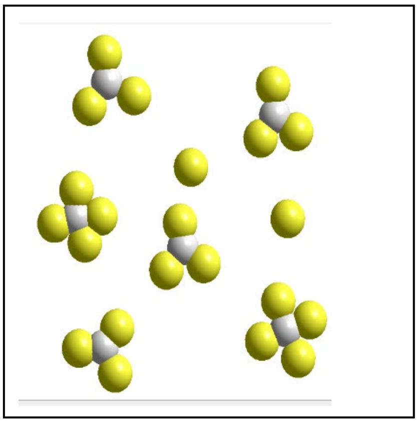
Low ability at the level of sub-microscopic representation can be a factor causing misconceptions when learning. This is because in learning chemistry, students are expected not only to know the theory but also to know in depth the concepts being learned and their relationship with daily life. Understanding at the level of sub-microscopic representation is the best way for students to interpret chemical material scientifically. Most students may already know the concept of chemistry but have not been able to connect between what they understand with actual theory. For example, students already know what and how kitchen salt, lime juice, detergent, are related to the daily life they experience, but they do not necessarily understand what is meant by sodium chloride, citrus acid, saponication, soda and other chemical terms others are learned in school. This means the constructivist theory (Bruce et al., 2016) is the basis of how prior knowledge that is known beforehand forms scientific concepts (Regan, Childs, & Hayes, 2011; Üce & Ceyhan, 2019).

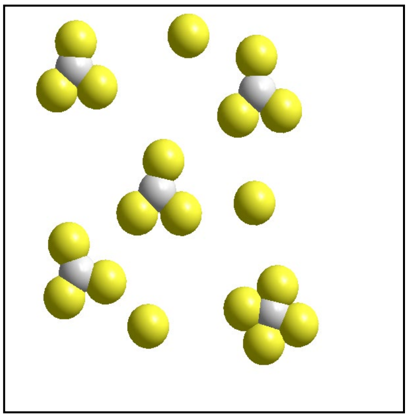
Misconceptions often occur because students do not understand chemistry at the level of representation (Anintia, Sadhu, & Annisa, 2018; Mubarak, 2016). In fact, the concept of multi-representation is proven to help students understand chemistry not only at the macroscopic level but also at the particulate and symbolic level (Chandrasegaran et al., 2007; Liang, Chou, & Chiu, 2011; Treagust et al., 2003; Trivic & Milanovic, 2018).

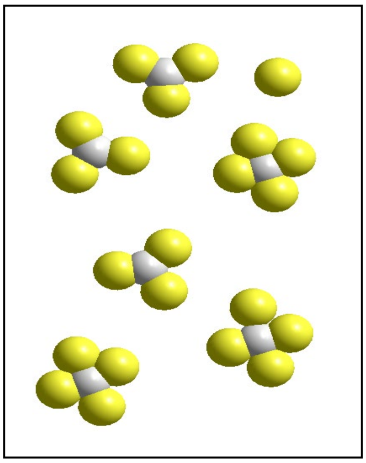
The following figure is the instrument numbers 14 and 15 which succeeded in becoming indicators to measure the occurrence of student misconceptions based on the results of the analysis using the Rasch model. Based on statistical data, it is known that only 5.23% or 10 people of the sample answered this item correctly. Microscopic picture of 6 molecules of BF3 and 4 molecules of F- with the reaction: BF3 + F- → BF4-
According to Lewis's acid-base theory below (see Figure 1a), the correct answer is option B.

Figure 1a. Formulation and answer key for question number 14



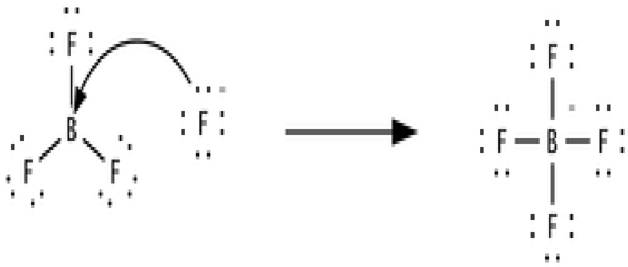
C. 

D. 

E. 

ANSWER: B

Explanation: $BF_3 + F^- \rightarrow BF_4^-$



Ion F- is a Lewis alkali because it can give PEB to BF3. After reacting 4 F-ions react with 6 BF3 molecules to form BF4- ions. Each F-ion reacts with one BF3 molecule, so there are 2 BF3 molecules that do not react with the F- ion.

15. Based on Question 14, the Lewis alkali from the reaction is...
A. F^-
B. BF_3
C. BF_4
D. BF_4 and BF_3
E. BF_3 and F^-
Answer: B
Explanation: BF_3 is the Lewis alkali after receiving PEB from F^-

Figure 1b. Formulation and answer key for question number 15

When viewed from the context of the problem, item numbers 14 and 15 have a link between the problem, for example, the reaction that occurs is $BF3 + F- \rightarrow BF4-$ with the description of the picture given. It indicates that F- is a Lewis alkali which gives a free electron pair (PEB) to BF3 so that if this item is not appropriately answered by students, then they experience misconceptions. This problem is very representative in knowing the ability of students' sub-microscopic level so that teachers can find out students' weaknesses and work to avoid misconceptions (Chandrasegaran et al., 2007; Trivic & Milanovic, 2018).

From the data of the logit item value and the student, logit can also be compared to the level of students' ability to answer questions. The average value of the item logit (item measure, Table 1) is 0.0 while the average value of the student logit (person measure) is -0 10. The average value of students' logit below the value of item/item logit. This means the ability of students is below the ability of the item/question so that it reinforces the reason for the low percentage of students who answered correctly on the item in question (5.23).

Table 2: Rasch Test Sample Results: Person-Measure, Students Symbolic Microscopic Understanding

Entry Number	Total Score	Total Count	Measure	Model S.E.	Infit Mnsq	Infit Zstd	Outfit Mnsq	Outfit Zstd	Pt-Measure Corr.	Pt-Measure Exp.	Exact Match Obs%	Exact Match Exp%	Person
12	15	15	5.58	1.91	MAXIMUM MEASURE				.00	.00	100.0	100.0	12PB
32	15	15	5.58	1.91	MAXIMUM MEASURE				.00	.00	100.0	100.0	32PB
35	15	15	5.58	1.91	MAXIMUM MEASURE				.00	.00	100.0	100.0	35PB
41	15	15	5.58	1.91	MAXIMUM MEASURE				.00	.00	100.0	100.0	41PB
59	15	15	5.58	1.91	MAXIMUM MEASURE				.00	.00	100.0	100.0	59PB
61	15	15	5.58	1.91	MAXIMUM MEASURE				.00	.00	100.0	100.0	61PB
64	15	15	5.58	1.91	MAXIMUM MEASURE				.00	.00	100.0	100.0	64PB
45	14	15	4.13	1.16	1.69	1.0	9.90	4.9	-.23	.32	93.3	93.3	45PB
142	14	15	4.13	1.16	1.66	1.0	5.08	2.0	-.02	.32	93.3	93.3	142PB
69	13	15	3.09	.91	2.24	1.8	9.90	3.1	-.14	.43	73.3	89.2	69PB
144	13	15	3.09	.91	.36	-1.3	.15	-.5	.61	.43	100.0	89.2	144PB
2	12	15	2.37	.80	.59	-.8	.35	-.2	.61	.48	93.3	86.7	2PB
5	12	15	2.37	.80	.81	-.2	1.02	.5	.50	.48	93.3	86.7	5PB
13	12	15	2.37	.80	.59	-.8	.35	-.2	.61	.48	93.3	86.7	13PB
24	12	15	2.37	.80	.89	-.1	5.08	2.0	.37	.48	93.3	86.7	4PB
27	12	15	2.37	.80	1.21	.6	.70	.2	.45	.48	80.0	86.7	27PB
30	12	15	2.37	.80	.59	-.8	.35	-.2	.61	.48	93.3	86.7	30PB
58	12	15	2.37	.80	.76	-.4	.66	.2	.54	.48	93.3	86.7	58PB

If we take one sample of student, for example, student with the 144PB code who answered incorrectly on items 14 and 15 above, where the student logit score of 144PB is 3.09 while the logit value of items 14 & 15 is 3.27 & 3.74. From these data, it can be seen that 144PB student cannot answer both items because the student logit value is below the item logit value or 144PB student ability is below the item ability (questions). When compared with other students, for example, students 12PB and 45PB both of which correctly answered the two-item questions, where the student logit score of 12PB and 45PB were 5.58 and 4.13, respectively. Their logit score data shows that 12PB and 45PB students have abilities above the ability of items (questions), so it is appropriate if the student answers both items correctly.

Item Fit / Misfit Order

Boone, et al (2014) and Bond & Fox (2015) explain that outfit means square, z-standard outfit, and measure correlation points are criteria used to look at patterns or contributions of student knowledge to items and item conformity levels (items -fit). It was concluded that item fit aims to identify the feasibility of distributed items so that they will represent whether each item has the potential to be used as a measurement tool or not. The criterion stating that the items are in the pattern according to the model is when they meet the following conditions (Sumintono, 2018):

- The received Outfit Mean Square (MNSQ) value: 0.5 <MNSQ <1.5
- The received Outfit Z-Standard (ZSTD) value: -2.0 <ZSTD <+2.0
- Point Measure Correlation (Pt Mean Corr) Value: 0.4 <Pt Measure Corr <0.85

Based on these criteria, item questions that can be analysed are presented in Table 3.

Table 3: Item Fit: Appropriateness Level of the Items

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
1	178	191	-4.96	.46	1.39	1.2	9.90	5.1	A .40	.55	96.6	96.6	Q1
2	145	191	-1.93	.23	1.14	1.0	5.59	7.3	B .50	.61	87.6	85.4	Q2
3	153	191	-2.38	.25	1.30	1.9	2.94	3.5	C .50	.61	83.1	87.3	Q3
14	22	191	3.23	.29	1.00	.1	2.03	1.7	D .44	.47	92.7	92.0	Q14
10	102	191	-.18	.19	.93	-.9	1.24	1.2	E .60	.58	80.8	75.5	Q10
8	69	154	.31	.21	1.02	.3	1.13	.5	F .58	.59	79.3	75.3	Q8
9	86	191	.36	.18	1.01	.1	.85	-.6	G .57	.57	71.8	74.5	Q9
7	66	191	1.05	.19	1.01	.1	.81	-.5	H .55	.54	73.4	76.7	Q7
13	60	191	1.27	.19	.94	-.6	.75	-.7	g .56	.53	78.5	77.8	Q13
11	53	191	1.54	.20	.89	-1.2	.87	-.3	f .55	.52	82.5	79.4	Q11
15	17	191	3.74	.35	.88	-.4	.56	-.7	e .49	.46	95.5	94.6	Q15
12	65	191	1.08	.19	.85	-1.8	.66	-1.1	d .59	.54	81.9	76.8	Q12
4	100	191	-.11	.19	.85	-1.8	.71	-1.5	c .63	.58	76.3	75.2	Q4
6	87	191	.33	.18	.85	-2.0	.74	-1.1	b .62	.57	78.0	74.5	Q6
5	166	191	-3.35	.30	.63	-2.1	.46	-1.0	a .67	.59	93.2	91.7	Q5
MEAN	91.3	188.5	.00	.24	.98	-.4	1.95	.8			83.4	82.2	
S.D.	48.0	9.2	2.25	.08	.18	1.2	2.48	2.5			7.7	7.9	

The question with the code of Q15 has an MNSQ outfit value of 0.56, an outfit value of ZSTD -0.7, and a Pt Measure Corr outfit value of 0.49. Seen from the criteria above, although the ZSTD outfit value is smaller than the criteria, Q15 needs to be maintained in measurements with values that match the criteria. Whereas Q14 has MNSQ, ZSTD, and Pt Measure Corr values of 2.03, respectively; 1.7; and 0.44. Even though the MNSQ outfit value is more than the specified criteria, Q14 is still maintained. This is influenced by the number of samples that the greater the number of samples the ZSTD value will also have an effect or can be greater than 3.0. Therefore, experts recommend not using the ZSTD value with very large sample conditions (N> 500). In other words, Q15 and Q14 can be used as a measuring tool in measuring the level of sub-microscopic understanding of students on acid-base material. This data recapitulation also helps teachers in assessing the appropriateness of items in measuring students 'understanding and basis in making teaching strategies that fit students' learning needs.

Table 4a: Statistics Summary: Students Symbolic Microscopic Understanding

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	7.2	14.8	-.10	.86				
SD	3.6	.4	2.34	.33				
MAX.	15.0	15.0	5.58	1.98				
MIN.	.0	14.0	-6.29	.66	.30	-2.0	.08	-.7

REAL RMSE .97 TRUE SD 2.13 SEPARATION 2.20 Person RELIABILITY .83 |
MODEL RMSE .92 TRUE SD 2.15 SEPARATION 2.33 Person RELIABILITY .84 |
S.E. OF Person MEAN = .17

Person RAW SCORE-TO-MEASURE CORRELATION = .98
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .84

Table 4b: Statistics Summary of 15 Tested Questions

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	91.3	188.5	.00	.24	.98	-.4	1.95	.8
SD	48.0	9.2	2.25	.08	.18	1.2	2.48	2.5
MAX.	178.0	191.0	3.74	4.6	1.39	1.9	9.90	7.3
MIN.	17.0	154.0	-4.96	.18	.63	-2.1	.46	-1.5

REAL RMSE .27 TRUE SD 2.24 SEPARATION 8.38 Item RELIABILITY .99 |
MODEL RMSE .25 TRUE SD 2.24 SEPARATION 8.86 Item RELIABILITY .99 |
S.E. OF Item MEAN = .60

Tables 4a and 4b provide statistical summaries showing the overall information about the quality of students response patterns or microscopic-symbolic understanding, the quality of instruments used, as well as interactions between people and items. Person measure = -0.10 shows the average value of all students working on the items given. The average value is greater than 0.0 indicates that the student's ability is considered good. Alpha Cronbach (measuring reliability, for instance, is the interaction between person and item as a whole), obtained 0.84 meaning ‘very good’ because the value is more than 0.5. This means the ability of students to answer the items is considered sufficient because based on the statistics above, the ability of students is below the ability of the items. In other words, if the average value is smaller than the logit value of 0.0, it indicates the tendency of students' ability to be smaller than the difficulty level of the items. This is confirmed by the person reliability and item reliability from Table 3 that the person (student) reliability is smaller than the item reliability.

In Table 3, the Person Reliability is 0.83 while the Item Reliability is 0.99. It can be concluded that the consistency of the students’ answers is ‘good’ (0.8-0.9 = Good). Likewise, with the items which were made to obtain ‘special’ reliability criteria (> 0.94 = Special). It is seen from the ‘scalogram’, that most of the students are able to answer the questions given, but the weak responses consistently show that the students are still experiencing a weak understanding of the basic concepts of chemical materials or misconceptions that can also occur. In fact, the items show that the instrument is in a good category with a value of 0.89.

It seems that the content of acid-base materials used to assess student understanding is considered suitable. In addition to the contribution of conceptual material, acid-base material also contains many interpretations of theory such as understanding pH, application of formulas and chemical reactions (acid-base), pH strength, concentration based on acid-base concepts, and continuity of acid-base material starting from Arrhenius, Browsted-Lowry to Lewis. As a result, microscopic-symbolic exploration in this material can be used in the analysis of the extent to which students touch the content of chemical material. In addition, the application of the concept of acid-base material is also commonly found in everyday life so that it can make students' understanding more complex and scientific. Misconceptions that occur can be concluded to occur because students do not learn the concept of acids and bases at a sub-microscopic-symbolic manner. This means that with a good understanding of the sub-microscopic-symbolic, it can be ascertained that misconceptions will not occur. Studying chemistry with sub-microscopic symbols also has great potential to hone students' critical thinking skills and their problem-solving abilities.

Based on the data presented, either the statistics summary or previous data provides information for teachers that they need to design learning that is suitable for students. From the data and analysis of learning styles obtained, one of the learning strategies that is considered suitable to be applied is problem-based learning (Erdogan & Senemoglu, 2014) with the brainstorming method. This is due to the consistency of students' answers which show where some students answered correctly on items with a high level of difficulty, but wrong on problems with a low level of difficulty. Then, there are questions that have the same logit value so that the need to provide questions that vary both the design of the problem, work procedures, or the level of difficulty. The aim is that students are challenged in completing and proving their answers scientifically and precisely. With questions that have varying degrees of difficulty, students will form their knowledge gradually and indirectly increase their academic abilities. This improvement in academic ability helps students to solve the next problem at a high level.

Further analysis of the acid-base material that has been presented shows the potential for misconceptions. Low cognitive achievement can be caused by students' understanding of being wrong, not knowing, or even by guessing the correct answer. In fact, the instruments developed have met the measurement criteria so that if students get low scores, there is a misconception in students' understanding of chemistry. Misconceptions that occur can be due to (1) students learning the concept of acid-base only at the macroscopic level, (2) lack of teaching materials that introduce concepts to the microscopic level, and (3) lack of discussion on the concepts at the microscopic level. As a result, studying chemistry with sub-microscopic symbols also give great potential to hone students' critical thinking skills (Rodzalan & Saat, 2015) and their problem solving (Maha, Brush, Haack, & Ho, 2018; Naiker & Wakeling, 2015).

Conclusion

Based on the results of this study, it can be concluded that the ability of the sub-microscopic level is a method that is able to be used by teachers in improving students' understanding of chemical concepts. Distractor-based multiple-choice diagnostic tests are alternative ways to find the extent to which students understand sub-microscopic chemical content. This test is also a basis for analysis to find whether the students experience misconceptions or not. The research presented shows that the use of the Rasch model can be used as an appropriate, effective, and systematic strategy in justifying the measurement of each item made. With a value of item reliability of 0.99 (> 0.95 = special category), the items tested were included as the right questions for use as a measurement tool. This means, in addition to showing the pattern of each student's answer, the Rasch model is also able to assess the quality of each item made so that the progress of student learning, teaching quality, and analysis of student understanding can be known.

Acknowledgement
The researchers' sincerest appreciation goes to the Rector of Lambung Mangkurat University and the Dean of the Faculty of Teacher Training and Education Lambung Mangkurat University Indonesia, who supported this research by providing the funding.

REFFERENCES

Allred, Z. D. R. (2019). Research and Practice multiple representations of the helium atom. https://doi.org/10.1039/c8rp00296g

Amin, N., Abdullah, H., Hashim, F. H., Arsad, N., Ali, S. H., & Aziz, N. A. (2012). Evaluation of Students' Achievement in the Final Exam Questions for Microelectronic (KKKL3054) using the Rasch Model. *Procedia - Social and Behavioral Sciences*, *60*(c), 119–123. https://doi.org/10.1016/j.sbspro.2012.09.356

Anintia, R., Sadhu, S., & Annisa, D. (2018). Identify Students' Concept Understanding Using Three-Tier Multiple Choice Questions (TTMCs) on Stoichiometry. *International Journal of Science and Applied Science: Conference Series*, *2*(1), 308. https://doi.org/10.20961/ijsascs.v2i1.16734

Boone, W. J. (2016). Rasch analysis for instrument development: Why,when,and how? *CBE Life Sciences Education*, *15*(4). https://doi.org/10.1187/cbe.16-04-0148

Bruce, M. R. M., Bruce, A. E., Avargil, S., Amar, G., Wemyss, T. M., & Flood, V. J. (2016). Polymers and Cross-Linking: A CORE Experiment To Help Students Think on the Submicroscopic Level. https://doi.org/10.1021/acs.jchemed.6b00010

Chan, S. W., Ismail, Z., & Sumintono, B. (2014). A Rasch Model Analysis on Secondary Students' Statistical Reasoning Ability in Descriptive Statistics. *Procedia - Social and Behavioral Sciences*, *129*, 133–139. https://doi.org/10.1016/j.sbspro.2014.03.658

Chandrasegaran, A. L., & Treagust, D. F. (n.d.). Emphasizing Multiple Levels of Representation To Enhance Students ' Understandings of the Changes Occurring during Chemical Reactions, *86*(12), 1433–1436.

Chandrasegaran, A. L., Treagust, D. F., & Mocerino, M. (2007). The development of a two-tier multiple-choice diagnostic instrument for evaluating secondary school students' ability to describe and explain chemical reactions using multiple levels of representation. *Chemistry Education Research and Practice*, *8*(3), 293–307. https://doi.org/10.1039/B7RP90006F

Cheung, D. (2011). Using diagnostic assessment to help teachers understand the chemistry of the lead-acid battery. *Chemistry Education Research and Practice*, *12*(2), 228–237. https://doi.org/10.1039/c1rp90028e

Chiang, W.-W. (2015). Ninth Grade Student' Self-assessment in Science: A Rasch Analysis Approach. *Procedia - Social and Behavioral Sciences*, *176*, 200–210. https://doi.org/10.1016/j.sbspro.2015.01.462

Chow, J., Tse, A., & Armatas, C. (2018). Comparing trained and untrained teachers on their use of LMS tools using the Rasch analysis. *Computers and Education*, *123*, 124–137. https://doi.org/10.1016/j.compedu.2018.04.009

Claesgens, J., Scalise, K., & Stacy, A. (2013). Learning mapping in chemistry: The perspectives of chemists. Chemical Education, 24 (4), 407-415. https://doi.org/10.1016/S0187-893X(13)72494-7

Cloonan, C. A., & Hutchinson, J. S. (2011). A chemistry concept reasoning test. *Chemistry Education Research and Practice*, *12*(2), 205–209. https://doi.org/10.1039/c1rp90025k

Eliyawati, Rohman, I., & Kadarohman, A. (2018). The effect of learning multimedia on students' understanding of macroscopic, sub-microscopic, and symbolic levels in electrolyte and nonelectrolyte. *Journal of Physics: Conference Series*, *1013*(1). https://doi.org/10.1088/1742-6596/1013/1/012002

Erdogan, T., & Senemoglu, N. (2014). Problem-Based Learning in Teacher Education : Its Promises and Challenges. *Procedia - Social and Behavioral Sciences*, *116*, 459–463. https://doi.org/10.1016/j.sbspro.2014.01.240

Gilbert, J. K. (2013). Multiple Representations in Biological Education Models and Modeling in Science Education. https://doi.org/10.1007/978-94-007-4192-8

Herrmann-Abell, C. F., & DeBoer, G. E. (2011). Using distractor-driven standards-based multiple-choice assessments and Rasch modeling to investigate hierarchies of chemistry misconceptions and detect structural problems with individual items. *Chemistry Education Research and Practice*, *12*(2), 184–192. https://doi.org/10.1039/c1rp90023d

Johnson, P. (2013). Una progresión de aprendizaje para la comprensión del cambio químico. *Educacion Quimica*, *24*(4), 365–372. https://doi.org/10.1016/S0187-893X(13)72489-3

Kilic, D., Sezen, N., & Sari, M. (2012). A Study of Pre-Service Science Teacher's Graphing Skills. *Procedia - Social and Behavioral Sciences*, *46*, 2937–2941. https://doi.org/10.1016/j.sbspro.2012.05.593

Lee, H. S., Liu, O. L., & Linn, M. C. (2011). Validating measurement of knowledge integration in science using multiple-choice and explanation items. *Applied Measurement in Education*, *24*(2), 115–136. https://doi.org/10.1080/08957347.2011.554604

Liang, J. C., Chou, C. C., & Chiu, M. H. (2011). Student test performances on behavior of gas particles and mismatch of teacher predictions. *Chemistry Education Research and Practice*, *12*(2), 238–250. https://doi.org/10.1039/c1rp90029c

Liu, X. (2015). Difficulties of items related to energy and matter: Implications for learning progression in high school chemistry. *Educación Química*, *24*(4), 416–422. https://doi.org/10.1016/s0187-893x(13)72495-9

Maha, P. G., Brush, E. J., Haack, J. A., & Ho, F. M. (2018). Journal of Chemical Education Call for Papers ☐ Special Issue on Reimagining Chemistry Education: Systems Thinking, and Green and Sustainable Chemistry, 8–10. https://doi.org/10.1021/acs.jchemed.8b00764

Maier, K. S. (2007). A Rasch Hierarchical Measurement Model. *Journal of Educational and Behavioral Statistics*, *26*(3), 307–330. https://doi.org/10.3102/10769986026003307

Mubarak, S. (2016). Journal of Innovative Science Education Pengembangan Tes Diagnostik Three Tier Multiple Choice. *Journal of Innovative Science Education*, *5*(2), 101–110.

Naiker, M., & Wakeling, L. (2015). Evaluation of group based inquiry oriented learning in undergraduate chemistry practicals. *International Journal of Innovation in Science and Mathematics Education*, *23*(5), 1–17.

Rabbitt, M. P. (2018). Causal inference with latent variables from the Rasch model as outcomes. *Measurement: Journal of the International Measurement Confederation*, *120*(February), 193–205. https://doi.org/10.1016/j.measurement.2018.01.044

Rauch, D. P., & Hartig, J. (2010). Multiple-choice versus open-ended response formats of reading test items : A two-dimensional IRT analysis. *Pschological Test and Assessment Modeling*, *52*(4), 354–379.

Regan, Á., Childs, P., & Hayes, S. (2011). The use of an intervention programme to improve undergraduatestudents' chemical knowledge and address their misconceptions. *Chemistry Education Research and Practice*, *12*(2), 219–227. https://doi.org/10.1039/c1rp90027g

Rodzalan, S. A., & Saat, M. M. (2015). The Perception of Critical Thinking and Problem Solving Skill among Malaysian Undergraduate Students. *Procedia - Social and Behavioral Sciences*, *172*(2012), 725–732. https://doi.org/10.1016/j.sbspro.2015.01.425

Runnels, J. (2012). Using the Rash model to validate a multiple choice English achievement test. *International Journal of Language Studies*, *6*(4), 141–155.

Salam, M., Iqbal, J., Hussain, A., & Iqbal, H. (2018). The Determinants of Services Sector Growth: A Comparative Analysis of Selected Developed and Developing Economies. *The Pakistan Development Review*, 57(1), 27-44.

Samad, G., Ahmed, V., & Khalid, R. (2018). Economic Contribution of Copyright-based Industries in Pakistan. *The Pakistan Development Review*, 57(1), 99-114.

Sambo, A. S., & Lucky, A. O. (2016). Undergraduates Perception and Utilization of Serial Materials in Federal University of Petroleum Resources Library. *American Journal of Social Sciences and Humanities*, 1(1), 1-9.

Serobatse, B. M., Selvaratnam, M., & Drummond, H. P. (2014). Students' conceptions about the sub-microscopic approach to explanations in chemistry throughout their BSc degree course : research article. *South African Journal of Chemistry*, *67*, 40–44. Retrieved from http://reference.sabinet.co.za/webx/access/electronic_journals/chem/chem_v67_a7.pdf

Sukor, N. S., Osman, K., & Abdullah, M. (2010). Students' achievement of Malaysian 21st Century Skills in Chemistry. *Procedia - Social and Behavioral Sciences*, *9*, 1256–1260. https://doi.org/10.1016/j.sbspro.2010.12.316

Sumintono, B. (2018). Rasch Model Measurements as Tools in Assesment for Learning, (October 2017). https://doi.org/10.2991/icei-17.2018.11

Talanquer, V. (2018). Importance of Understanding Fundamental Chemical Mechanisms. https://doi.org/10.1021/acs.jchemed.8b00508

Talib, A. M., Alomary, F. O., & Alwadi, H. F. (2018). Assessment of Student Performance for Course Examination Using Rasch Measurement Model: A Case Study of Information Technology Fundamentals Course. *Education Research International*, *2018*, 1–8. https://doi.org/10.1155/2018/8719012

Treagust, D. F., Chittleborough, G., & Mamiala, T. L. (2003). The role of submicroscopic and symbolic representations in chemical explanations. *International Journal of Science Education*, *25*(11), 1353–1368. https://doi.org/10.1080/0950069032000070306

Trivic, D. D., & Milanovic, V. D. (2018). The macroscopic, submicroscopic and symbolic level in explanations of a chemical reaction provided by thirteen-year olds. *Journal of the Serbian Chemical Society*, *83*(10), 1177–1192. https://doi.org/10.2298/JSC171220055T

Üce, M., & Ceyhan, İ. (2019). Misconception in Chemistry Education and Practices to Eliminate Them : Literature Analysis, *7*(3), 202–208. https://doi.org/10.11114/jets.v7i3.3990

Villafañe, S. M., Loertscher, J., Minderhout, V., & Lewis, J. E. (2011). Uncovering students' incorrect ideas about foundational concepts for biochemistry. *Chemistry Education Research and Practice*, *12*(2), 210–218. https://doi.org/10.1039/c1rp90026a

Wei, S., Liu, X., Wang, Z., & Wang, X. (2012). Using Rasch Measurement To Develop a Computer Modeling-Based Instrument To Assess Students ' Conceptual Understanding of Matter.

Zamri bin Khairani, A., & Bin Abd. Razak, N. (2015). Modeling a Multiple Choice Mathematics Test with the Rasch Model. *Indian Journal of Science and Technology*, *8*(12). https://doi.org/10.17485/ijst/2015/v8i12/70650