International Journal of Evaluation and Research in Education (IJERE)
Vol.5, No.4, December2016, pp. 306~309
ISSN: 2252-8822

Validity and Realibility of Chemistry Systemic Multiple Choices Questions (CSMCQs)

Erfan Priyambodo, Marfuatun
Department of Chemistry Education, Universitas Negeri Yogyakarta, Indonesia

Article Info

Article history:

Received Sep 27, 2016
Revised Nov 15, 2016
Accepted Nov 24, 2016

Keywords:

Multiple choices question
Rasch model
Reliability
Systemic approach
Validity

ABSTRACT

Nowdays, Rasch model analysis is used widely in social research, moreover in educational research. In this research, Rasch model is used to determine the validation and the reliability of systemic multiple choices question in chemistry teaching and learning. There were 30 multiple choices question with systemic approach for high school student class XI. In this research, a 164 student as the subject of the research. The data was collecting during March-April 2015. The data was analysis using Winsteps application with Rasch model. The reseach shows that 29 question of 30 questions was valid through Rasch model analysis. The reliability of systemic multiple choices question was 0.93, which is very good chategory.

Corresponding Author:

Erfan Priyambodo,
Department ofChemistry Education,
Universitas Negeri Yogyakarta, Karangmalang, Yogyakarta, Indonesia.
Email: erfan@uny.ac.id

1. INTRODUCTION

Chemistry teaching and learning is a learning process in understanding the chemistry concepts. As well asother subjects, there willbe givenan evaluationto measure theacademicabilitiesof students.There are several instrumentsto measure the academic achievement of students, such as multiple choices question, true-false question, etc. These kinds of instruments can be categorized in two models, which are paper based question and computer based question.

Multiple choices questions is one formof evaluationthatthe answer can beobtainedby selectingone of the answersthathave been provided. This kind of evaluation consists of two parts, whish are a problem (stem) and a list of suggested solutions (alternatives) [1]. The basic form of a stemis is a question or an incomplete statement. Moreover, the list of the alternatives contains of the best answer (or the correct answer) and a number of incorrect answer or inferior alternatives (distractor).

Multiple choices question can be used to measure various levels of cognitive knowledge, which are related to Bloom's taxonomy of educational objectives (knowledge, comprehension, application, analysis, synthesis, and evaluation).The strengths and limitations of this type question are shown in Table 1.

Usually, multiple choices question has not beenable to reveal the chemistry conceptsin a comprehensive manner. As a result, studentsare lessable to thinkconstructivelyandtend to forgetthe basicchemistryconceptsthey have learnedpreviously [2]. Therefore, it is required a new approach in developing a multiple choices questions.

In this study, the multiple choices questions was developed using systemic approach. Thefore, this type of question is namely chemistry systemic multiple choices questions (CSMCQs). The systemic approach Is an approach in teaching and learning where multiple components are connected to each other byamutualrelationship [3]. The alternatives of SMCQs consists of three concepts in chemistry which are connected each other.

Table 1. The strengths and the limitations of multiple choices question [4]

Strengths	Limitations
1. Learning outcomes from simple to complex can be measured.	1. Constructing good items is time consuming.
2. Highly structured and clear tasks are provided.	2. It is frequentlyo difficult to find plausible distractors.
3. A broad sample of achievement can be measured.	3. It is ineffective for measuring some types of prolem solving and the ability to organize and express ideas.
4. Incorrect alternatives provide diagnostic information.	4. Scores can be influenced by reading ability.
5. Scores are more reliable then subjectively scored item (e.g. essay).	5. Often focus on testing factual information and fails to test higher level of cognitive thinking.
6. Scoring is easy, objective, and reliable.	6. Sometimes there is more than on defensible "correct" answer.
7. Item analysis can reveal how difficult each item was and how well it discriminated between the strong and weaker student in the class.	7. Does not provide a measure of writing ability.
8. Performance can be compared from class to class and year to year.	8. May encourage guessing.
9. Can cover a lot of material very efficiently.	

As an instrument to measure the academic abilities of student, this kind of evaluation should qualify the validity and reliability. Reliability is the degree to which measures are free from error and therefore yield consistent results (i.e. the consistency of a measurement procedure) [5]. The instrument is considered reliable if the instrument consistently assigns the same score to individuals or objects with equal values. The aim of the study was identifying the validity and reliability of CSMCQ though Rasch model analysis. The software for analysing the items was WINSTEPS software.

2. RESEARCH METHOD

This study is part of research and development aimed at developing an evaluation model based on a systemic approach. The kind of evaluation model was systemic multiple choice questions in chemistry teaching and learning. The CSMCQs were focussing in two main concepts, i.e. salt hydrolysis and buffer solution. The CSMCQs were consists of 30 multiple choices questions.

Validity determines whether the research trulymeasures that which it was intended to measureorhow truthful the research resultsare [6]. The process of validation involves accumulating evidence to providea sound scientific basis for the proposed score interpretations [7]. The CSMCQs should qualify the validity, so it needs to be tested to the students of class XI. The validity of the evaluation items were tested by application Rasch. Items are quantitatively analyzed using WINSTEPS software based on the Raschmodel to assess the validity and reliability of items.

3. RESULTS AND ANALYSIS

Rasch model with the application of WINSTEPS was used to analyze the datas as well as to test the validity and reliability of the instrument. Rasch analysis offers a method of ensuring that key measurement assumptions are tested and, where data fit the model, arithmetic operations may be undertaken [8]. The Rasch model considers the ability of the students, consists of 164 students, who answered the CSMCQs.

The items of CSMCQs were valid based on the Rasch model analysis if it complies with the following criteria.

a. The value of Outfit Mean Square (MNSQ) in the range of 0.5 <MNSQ <1.5
b. The value of Outfit Z-Standard (ZSTD) in the range of -2.0 < ZSTD< +2.0
c. The value of Point Measure Coorelation (Pt Mean Corr) in the range of 0,4 <Pt Measure < 0,85

Items was valid should meet at least one of the criteria above [9].

The validity of an instrument of evaluation is the degree that shows where a test measures what it intends to be measured. The value of each item through Rasch model analysis is shown in Table 2. Based on these criteria above, it was 1 item, which is number 14, from 30 items which not valid because the value of MNSQ, ZSTD and Pt Mean Corr out of the range that is allowed.

The reliability analysis of CSMCQs performed on 29 valid items, which is shown in Table 3. Based on Rasch model analysis using WINSTEPS software, the value of item reliability is 0.93. This value can be chategorized as very good chategory [9].Reliability, the consistency of a test or measurement, is frequently quantified in the movement sciences literature [10].The realibility of this instrument only conducted in this population under certain condition.

Chemistry is a very complex subject is shown from the research on problem solving and misconceptions [11]. Therefore, we need to use new approaches in chemistry teaching and learning so that students can understand the chemistry comprehensively. An approach that can be used is a systemic approach, where some multiple componentsare connectedto each other byamutualrelationship. It required the

evaluation model that supports such approach.The chemistry systemic multiple choices questions (CSMCQs) is an instrument to measure the academic abilities of student class XI. The CSMCQs is a set of question which has three chemistry concepts connected each other byamutualrelationship in every alternative. The CSMCQs consists of 30 items of buffer solution and salt hydrolysis concepts. This material was chosen because many students who have misconceptions on this matter [12].

Table 2. Validity of items CSMCQs trough Rasch Model Analysis

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
14	30	164	90.94	2.27	1.73	4.6	2.28	3.9	A	-.11	.43	70.1	83.8	14
7	156	164	36.96	3.71	.99	.1	1.60	1.0	B	.16	.18	95.1	95.1	7
13	142	164	48.86	2.42	1.00	.1	1.38	1.1	C	.24	.29	86.6	86.6	13
27	163	164	15.29	10.08	1.02	.4	1.28	.6	D	.02	.07	99.4	99.4	27
21	153	164	40.52	3.22	1.05	.3	1.26	.6	E	.15	.21	93.3	93.3	21
12	101	164	65.34	1.80	1.14	1.9	1.04	.4	F	.36	.44	64.6	70.5	12
29	115	164	60.63	1.88	1.13	1.7	1.03	.3	G	.32	.41	75.0	73.5	29
4	114	164	60.98	1.87	1.12	1.6	1.00	.1	H	.33	.41	69.5	73.1	4
16	155	164	38.26	3.52	.97	.0	1.11	.4	I	.20	.19	94.5	94.5	16
11	136	164	52.07	2.22	1.09	.7	1.05	.3	J	.25	.32	82.9	83.1	11
3	147	164	45.64	2.68	1.04	.3	1.03	.2	K	.22	.26	89.6	89.6	3
19	158	164	33.83	4.24	1.03	.2	.89	.0	L	.13	.16	96.3	96.3	19
17	154	164	39.44	3.36	1.01	.1	1.02	.2	M	.19	.20	93.9	93.9	17
10	142	164	48.86	2.42	1.01	.1	.97	.0	N	.27	.29	86.6	86.6	10
20	158	164	33.83	4.24	.99	.1	.99	.2	O	.17	.16	96.3	96.3	20
2	140	164	49.99	2.34	.99	.0	.87	-.3	o	.32	.30	86.0	85.4	2
22	84	164	70.75	1.78	.97	-.4	.98	-.2	n	.49	.47	72.6	70.7	22
25	135	164	52.55	2.19	.98	-.2	.77	-.8	m	.37	.33	82.3	82.5	25
18	155	164	38.26	3.52	.96	.0	.82	-.1	l	.23	.19	94.5	94.5	18
23	148	164	44.90	2.74	.95	-.2	.83	-.9	k	.32	.25	90.2	90.3	23
30	138	164	51.06	2.27	.95	-.4	.75	-.8	j	.37	.31	84.8	84.2	30
5	98	164	66.31	1.79	.95	-.7	.90	-.8	i	.49	.45	77.4	70.3	5
1	135	164	52.55	2.19	.95	-.4	.72	-1.0	h	.39	.33	82.3	82.5	1
6	154	164	39.44	3.36	.93	-.2	.51	-.9	g	.29	.20	93.9	93.9	6
8	139	164	50.53	2.31	.93	-.5	.82	-.5	f	.36	.31	85.4	84.8	8
26	117	164	59.91	1.90	.90	-1.2	.78	-1.2	e	.48	.40	76.2	74.3	26
9	160	164	29.53	5.13	.89	-.1	.25	-1.3	d	.27	.13	97.6	97.6	9
15	144	164	47.65	2.51	.88	-.7	.51	-1.5	c	.41	.28	87.8	87.8	15
28	69	164	75.58	1.82	.81	-2.5	.83	-1.6	b	.62	.49	76.2	72.4	28
24	118	164	59.55	1.91	.76	-3.0	.60	-2.4	a	.58	.40	80.5	74.6	24
MEAN	131.9	164.0	50.00	2.92	1.00	.0	.96	-.2			85.4	85.4	
S.D.	30.0	.0	14.90	1.58	.16	1.3	.36	1.1			9.3	9.1	

Table 3.The reliability of SMCQs

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	24.0	29.0	74.34	7.44				
S.D.	3.5	.0	14.23	4.36				
MAX.	29.0	29.0	103.64	18.59				
MIN.	15.0	29.0	50.91	4.29	.63	-1.5	.15	-1.0
REAL RMSE	8.81	TRUE SD	11.17	SEPARATION	1.27	Person RELIABILITY		.62
MODEL RMSE	8.62	TRUE SD	11.32	SEPARATION	1.31	Person RELIABILITY		.63
S.E. OF Person MEAN = 1.11								

Person RAW SCORE-TO-MEASURE CORRELATION = .94
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .73

SUMMARY OF 29 MEASURED Item

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	135.4	164.0	50.00	2.96	.99	.0	1.07	.2
S.D.	23.7	.0	13.24	1.59	.09	1.0	.39	1.1
MAX.	163.0	164.0	78.11	10.08	1.18	2.6	2.31	2.6
MIN.	69.0	164.0	16.54	1.84	.78	-3.0	.29	-2.6
REAL RMSE	3.39	TRUE SD	12.80	SEPARATION	3.77	Item	RELIABILITY	.93
MODEL RMSE	3.36	TRUE SD	12.81	SEPARATION	3.81	Item	RELIABILITY	.94
S.E. OF Item MEAN = 2.50								

Before it can be used, this kind of evaluation should qualify the validity and reliability. The instrument validity determines whether the instrumet trulymeasures that which it was intended to measure [6]. The instrument reliability is the degree to which measures are free from error and therefore yield consistent results. Based on the analysis, it is show that 29 itemswas valid and it can be used to measureacademic abilities in chemistry teaching and learning with very good reliability.

4. CONCLUSION

The result of this study showed that 29 question of 30 questions was valid through Rasch model analysis. The reliability of chemistry systemic multiple choices question (CSMCQs) was 0.93, which is very good chategory.

REFERENCES

[1] R. R. Sudweeks, *et al.*,"How to prepare better multiple-choice test items: Guidelines for university faculty," Brigham Young University Testing Services, 1991.
[2] J. J. Lagowski, *"SATL, Learning Theory, and the Physiology of Learning*," 20th ICCE Proceeding, Mauritius, pp.3-8, 2008.
[3] Z. A. Bashaireh, "Systemic Approach Effect on Achievement of Tafila Schools Students in Science," *International Journal of Humanities and Social Science*, vol/issue: 3(1), pp.47-52, 2011.
[4] D. M. Zimmaro, "Writing good multiple-choice exams," Measurement and Evaluation Center.The University of Texas. Austin. Retrieved on, 2004.
[5] G. Thanasegaran, "Reliability and validity issues in research," *Integration & Dissemination*, vol. 4, pp. 35-40, 2009.
[6] M. Joppe, "The Research Process," 1998. Retrieved February 25, 2000, from http://www.ryerson.ca/~mjoppe/rp.htm
[7] S. Messick, "Vailidty," In R.L. Linn (Ed.). A nation at risk: The imperative for education Reform Washington, DC: U.S. Government Printing Office, 1989.
[8] A. Tennant, et al.,"Application of Rasch analysis in the development and application of quality of life instruments," *Value in Health*, vol/issue: 7(s1), pp. S22-S26, 2004.
[9] B. Sumintono and W. Widhiarso, *"*Aplikasi Model Rasch untuk penelitian ilmu-ilmu sosial," Trimkom Publishing House, Bandung, 2013.
[10] J. P. Weir, "Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM," *The Journal of Strength & Conditioning Research*, vol/issue: 19(1), pp. 231-240, 2005.
[11] D. Gabel, "Improving teaching and learning through chemistry education research: A look to the future,"*J. Chem. Educ*, vol/issue: 76(4), pp. 548, 1999.
[12] M. Orgill and A. Sutherland, "Undergraduate chemistry students' perceptions of and misconceptions about buffers and buffer problems," *Chemistry Education Research and Practice*, vol/issue: 9(2), pp. 131-143, 2008.