Validation of A Critical Thinking Test in Mathematics for Senior High School

AUTHORS INFO

Ma. Rachel Kim L. Aure
Visayas State University, Baybay City, Leyte, Philippines
kim_aure@vsu.edu.ph

Enriqueta D. Reston
University of San Carlos, Cebu City, Cebu, Philippines
edreston@usc.edu.ph

Abstract

Critical thinking equips a person with the competencies needed and develops the power of the mind extensively. The development of critical thinking is known as an essential goal in educational systems. However, the assessment of critical thinking is not an easy feat. This study validated the critical thinking skills test in mathematics for senior high school students based on a model of critical thinking skills and learning areas of the K to 12 Mathematics Curriculum. The instrument validation process to 902 respondents through cluster sampling from different senior high schools. With an initial 45 items on three dimensions of critical thinking, the test's final form consisted of only 30 items after the items were screened based on the validation process. The content validation was carried out through an expert examination of the items and a comparison of the test results developed with Cornell's critical thinking for criterion-related validity evidence ($r = 0.299$, $p < 0.01$). Construct validation by factor analysis, and results generated 13 factors accounting for 58.51% of the variance. These results indicated that the developed mathematics critical thinking test measured other than the specified critical thinking skills model used. The developed test may try to assess students' critical thinking in the context of mathematics. Furthermore, this may contribute to promoting teaching and assessment strategies gears towards the development of students' mathematical critical thinking.

Keywords: critical thinking skills, test instrumentation, test validation

A. Introduction

Thinking is a cognitive exercise that every person does in all aspects of life. If a person can develop his skills in thinking, he can achieve many things, like being successful in social life, being emotionally and economically mature, and coping with life's uncertainties (Tang, 2014). One type of thinking is critical thinking. According to Ennis (1996), critical thinking is making reasonable decisions about what to do and what to believe. As a critical thinker, a person equipped with the competencies needed to make informed and sound decisions that are useful in day-to-day situations.

Critical thinking skills growth is now known as a significant target in education systems. Asian countries have supported the importance of this skill for students to hone and participate in society (Education Bureau, 2003). In the Philippines, the K to 12 mathematics curriculum indicates that the development of problem-solving and critical thinking skills are the twin goals for teaching the said subject. These two goals are necessary, covering all the learning domains (cognitive, affective, and psychomotor skills), taking into account the varied contexts of learners in the country (DepEd, 2010). Moreover, they are envisioned by developing both these skills and will relate it to real-life situations while studying mathematics.

The nature and assessment of critical thinking is a challenging task. Some educators in schools have commented that this ability was not performed rigorously in a systematic way (Pithers & Soden, 2010). The lack of proper assessment is one of the obstacles in effectively and objectively measuring the critical thinking skills of students (Halpern, 2003). If there is no appropriate assessment tool that would adequately measure students' critical thinking abilities, it will be complicated to examine the effectiveness of skills gained by the students. It is then very crucial to have a valid assessment that can help to identify curriculum enhancement, know student needs, and provide feedback for both teachers and students to devise teaching and learning plans altogether (Ennis, 1996; Retnawati, et al., 2017).

Moreover, there are researches done on critical thinking skills in other fields focusing on students (Aktas & Unlu, 2013), teaching and learning (Behar-Horenstein & Niu, 2011), curriculum (Butera et al., 2014), reading comprehension (Maandig, et al., 2017), and nursing education (Yildirim, 2011). In the country, there are recent researches on mathematics concerning the effect of using problem situations on students critical thinking skills in solving problems (Belecina, 2017), correlation of mathematics level of critical thinking and problem – solving skills (Alcantara & Bacsa, 2017) and a review on critical thinking in Philippine education (Marquez, 2017). However, with extensive search, there is no locally developed critical thinking skills test that may serve as a valuable assessment tool aligned with the learning domains of mathematics education.

In this context, the current study aimed to validate a test for assessing critical thinking in line with the goals of the K to 12 mathematics curriculum among senior high schools. Specifically, it sought to answer the validity of the developed test based on (a) construct validity, (b) content validity, and (c) criterion-related validity. Further, this can be very useful for mathematics teachers and researchers in assessing senior high school students' critical thinking skills for diagnostic purposes, improve classroom instruction, and research along this line. Critical thinking evaluation using this test can also be a basis for the development of mathematics teaching lessons that incorporate critical thinking development among students, thereby improving mathematics teaching and learning.

B. Literature Review

1. Critical Thinking

Students in today's society are gathering tons of knowledge from a particular learning environment that requires a great deal of cognitive operation. Hence, critical thinking becomes an essential part of the education system (Angeli & Valanides, 2009; Listiani, 2016; Adnyani, 2020). The skills they learned needs to be accompanied by proper behaviors to make them useful in critically thinking about their daily life, monitor their thought processes, and act congruently with one's critical thinking (Paul, 1993). Nowadays, critical thinking is given more importance because of its nature and complexities that can be very useful in real-life situations. Further, critical thinking as developing skills in the 21stcentury (Halpern, 2003). Many models were built to capture what constitutes rational thinking, which could be the basis for developing methods for evaluating and quantifying the construct.

Dwyer, et al. (2013) contend that one model for critical thinking is Bloom's taxonomy of educational objectives, which consists of knowledge, comprehension, application, analysis, synthesis, and evaluation. These objectives are associated with critical thinking. It encompasses the thought processes specified in Bloom's taxonomy if it reaches the evaluation phase and, thus, belongs to the higher-order thinking skills. When deciding on what to do, this skill is also focusing on reasonable and reflective thinking (Ennis, 2005).

The teaching of critical thinking is an area identified that needs to be developed in the teaching-learning process because it allows individuals to go beyond simple recall of information to gain more complex understanding (Halpern, 2003). Students should have the ability, willingness, and foresight to critically handle new information so that they can solve problems, make reasonable conclusions, and decide accordingly (Dwyer, et al., 2014). Further, students would become adaptable to the environment and cope with the vast development of this generation.

2. *Validity*

Validity exists along a continuum in varying degrees, and it is inferred or judged from existing evidence, not measured or calculated directly (Colton & Covert, 2007). Therefore, validity exists in a continuum. Since validity is in a continuum, it should have sufficient evidence for the instrument to produce valid results and obtain the necessary information. A valid assessment is directly on the concepts of the measuring instrument.

In particular, this study uses construct validity, content validity, and criterion validity. Construct validity is necessary for theory development and testing (Peter, 1981). It is the appropriateness of inferences made based on test scores for its intended construct. On the other hand, The validity of the content requires test items, whether it covers all the general content areas, and the degree to which an instrument is reflective of the topic and mechanism under investigation (Colton & Covert, 2007). It is identifying as many factors as possible to operationalize the constructs. One way to build a test's content validity is to start with building a table of specifications. Table of specifications (TOS) is a test blueprint that teachers make to align objectives, topic content, and assessment tool (Notar, Zuelke, Wilson, & Yunker, 2004) as cited by (Fives & Barnes, 2013). When constructing a test, it should measure an adequate sampling of the topic content of the subject matter objectively (Fives & Barnes, 2013).

Moreover, content validity can give the readers confidence about the instruments used because construct validity is guaranteed given with utmost importance (Rubio, et al., 2003). It refers to the degree that the instrument covers the content that it is supposed to measure. Lastly, criterion-related validity is the extent to which a measure is related to an outcome (Salkind, 2010).

3. *Assessments of Critical Thinking*

There have been several definitions of critical thinking, and as such, assessing it will vary depending on its purpose, content, and format (Ku, 2009). Critical thinking is through the extrapolation of observation and information (Paul, 1991). It requires higher-order thinking over lower-order thinking (Moore, 2001). Critical thinkers can take responsibility for their thoughts, which may develop specific standards to evaluate and create, then eventually use them to solve problems they encounter (Elder & Paul, 1994). Moreover, it can be a tool that ignites the production of cognitive skills to use in the teaching-learning process (Kezer & Terker, 2012).

Critical thinking in mathematics depends on students' ability to think constructively and creatively, ranging from intuition to logic, and on the flexibility to argue a problem from many views (Dinuta, 2015). Further studies show that critical thinking in mathematics needs to be integrated and emphasized in the curriculum so that students can learn the necessary skills to apply and improve their performance and ability (Chukwuyenum, 2013). Many studies investigate critical thinking in general or in fields other than mathematics, but few discuss critical thinking in Mathematics (Aizikovitsh-Udi, 2012).

Most students' lecture-discussion in the classroom is a passive activity because they are only listening and will not have the opportunity to ask a question when the lecture is going on (Chukwuyenum, 2013). The acquisition of critical thinking must be learned by actually performing them (Schafersman, 1991). Therefore, the teacher should emphasize students' active intellect in teaching (Abiam & Odok, 2006). Training skills associated with critical thinking are vital, as some would call it the main objective of university education and skill (Tamayo, et al., 2014). Further, teaching critical thinking improves the students' quality of thinking about any

subject content and solve the problem by skillfully assessing it (Flores, 2016). In the result of Visande (2014), it is imperative to develop students' critical thinking skills in which there are various opportunities where students can enhance their cognitive abilities and develop skills that proceed from simple to complex.

C. Methodology

1. Research Design and Research Respondents

This research is a validation study using the non-experimental design that sought to validate a test for assessing critical thinking skills in mathematics among Grade 11 senior high students. In deciding the respondents, cluster sampling by class segment per secondary school was done. This kind of sampling breaks large segments of the population into smaller units (Colton & Covert, 2007).

Since the primary purpose of this study was to validate a test, as suggested by Burton & Mazerolle (2011), in determining the number of respondents, there would be ten respondents per item developed in the instrument. Since the test developed in this study consists of 45 items, 450 respondents utilized it. Not all respondents participated, a spillover of 20% of the total number of respondents was employed. Nine hundred two student respondents participated in this study. The pretesting phase was conducted in two public senior high schools to 66 (7.32%) Grade 11 students with one randomly chosen section per school. In the pilot testing phase, there were 481 (53.33%) student respondents in three different senior high schools and the second validation. There were 355 (39.35%) student respondents in four different senior high schools. Also, among the respondents, 449 (49.78%) were male, while 453 (50.22%) were female.

The development phases consisted of pretesting and pilot testing phases. The test's validity was achieved via the expert evaluation, followed by the refinement of the items based on the expert's suggestions. Construct validity is done by factor analysis. Meanwhile, the reliability of the instrument used the Cronbach Alpha formula. The value of the acceptable Alpha ranges from 0.60 – 0.80(UCLA Institute for Digital Research and Education Statistical Consulting, 2020).

2. Instruments

The instruments used in this study were the following: (a) Cornell Critical Thinking Skills Test Level X to assess students' critical thinking abilities that consist of 45 multiple choice items for assessing critical thinking based on the K to 12 (b) Critical Thinking Skills in Mathematics was a researcher-made questionnaire align with K to 12 Education Mathematics Curriculum learning domains vis-à-vis the critical thinking skills, (c) Evaluation on Content Validity by Experts was a researcher-made matrix for the content expert on their comments and suggestions for content validation about the items developed and it will serve as a basis whether an item will be accepted, revised, or rejected.

In this study, the Robert Ennis Cornell Critical Thinking Skills Test Level X is used to determine the students' critical thinking abilities. This test measures induction, deduction, observation, credibility, and assumptions. Furthermore, critical thinking is integrated into K to 12 mathematics learning domains and is most appropriate in mathematics since it encompasses all skills needed by a person to do the subject.

3. The Technique of Data Analysis

Descriptive statistics such as means, frequencies, standard deviation, and percentages have been used to provide a demographic profile for the respondents. The information gave additional data in the analysis of each item. Content validity of the items established by results of content expert evaluation of the alignment of test items with the table of specifications concerning the critical thinking skills and mathematics content was analyzed using the percentage of agreement and effect size measures. For criterion-related validity, the discrimination index is a measure together with concurrent validity. Concurrent validity evaluates by obtaining correlations between scores on the test with current behaviors (Warner, 2013). This study established through the use of a correlation coefficient with the senior high school students' scores in the Cornell Critical Thinking Skills as the external criterion.

The confirmatory-factor analysis was used to check the internal structure of the study as to the component constructs. Factor analysis is the study of the interdependence between variables to find a new set of variables, which is considerably less than the original variables (Rosnawati, et

al., 2015). Factor analysis can measure the validity of the developed test (Thompson & Vacha-Haase, 2000). It would also suggest how many factors require the correlations between variables included in the study (Warner, 2013).

D. Findings and Discussion

1. Findings

A total of forty – five items were initially developed based on the critical thinking skills by Robert Ennis and the Department of Education mathematics learning areas. The items were equally divided based on their objectives—the developed mathematics critical thinking test items sent to mathematics education experts for face and content validation. The suggestion from content experts was to reject four items (items 16, 25, 37, and 44), which were incorporated accordingly. The critical thinking test items listed assessed for Grade 11 students who were not part of the target population for the validation process. The purpose of the pretesting was to determine the appropriate time allotment for the succeeding phase of this study and solicit initial feedback from student respondents for clarity of instructions. Based on the result of the pretesting phase, sixty minutes was the official allotted time of taking the developed test items.

During phase 2 – pilot testing, the forty – one item was administered to 481 students. After the test administration, item analysis was employed. Index of difficulty, index of discrimination, and index of effectiveness calculated per item, which used one of the bases for Acceptance (A), Revision (Rv), or Rejection (Rj) of the concerned item. Based on the result of the first item analysis, items number 7, 10, and 13 (induction), 19, 20, 21, 25, and 28 (deduction and assumption), 31, 34, 37, 38, 42, and 44 (observation and credibility) rejected because of the unacceptable value of discrimination index. Further, based on the results of item difficulty and discrimination index items, numbers 7, 10, 13, 19, 20, 21, 22, 25, 28, 31, 34, 38, 42, and 45 are also not included in the analysis because of these rejected items. All of the items had distractors that need to be revised, as evidenced by their respective values' positive values. All items with option D discarded, as demonstrated by the efficiency index zero value, during phase 3 – actual implementation, the same process employed to 312 respondents. The overall test results in terms of the item difficulty level are 0.38274 (average difficulty), and its discrimination index is 0.23077 (subject for improvement). Of the 45 objects, 30 are retained/revised, while 15 rejected. From the start of the conduct of this study, four items were rejected based on content mathematics experts, 15 items rejected, and 30 items retained/revised.

In terms of reliability, the equivalence of the developed mathematics critical thinking test to the Cornell Critical Thinking test and the overall reliability were moderately reliable, with a reliability value of $r = 0.507$ and $r = 0.525$, respectively. Nevertheless, before performing factor Analysis, a preliminary assessment of the data's suitability was done through the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy and Bartlett's Test of Sphericity (BTS). For KMO, more significant than 0.5 indicates that the sample size is adequate, and factor analysis may be useful. According to IBM Knowledge Center (2017), Bartlett's Test of Sphericity tests the hypothesis that a correlation matrix is an identity matrix that would indicate unrelated variables and, therefore, unsuitable for structure detection. Values less than the significance level indicate that factor analysis may be useful with the given data set.

Table 1. KMO Measure of Sampling Adequacy and BTS of the developed test

KMO and Bartlett's Test		1st validation	2nd validation
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.553	0.572
Bartlett's Test of Sphericity	Approx. Chi-Square	1395.549	778.486
	df	820	406
	p – value	< 0.0001	< 0.0001

Table 1 presents the results of KMO and BTS. The sample size used is adequate since KMO is $0.553 > 0.5$ and the factor model is appropriate with *p*-value < 0.05 (1st validation) and in increase of KMO is $0.572 > 0.5$ and the factor model is appropriate with *p*-value < 0.05 (2nd validation). It implies that the developed test items can is useful for factor analysis.

Table 2. Total variance explained of the developed test.

	Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
		Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1st validation	1	2.526	6.161	6.161	1.666	4.064	4.064	1.324	3.229	3.229
	2	1.815	4.428	10.588	.940	2.292	6.355	1.280	3.122	6.351
	3	1.594	3.888	14.476	.704	1.718	8.073	.706	1.722	8.073
2nd validation	1	2.652	9.146	9.146	1.811	6.246	6.246	1.213	4.182	4.182
	2	1.658	5.716	14.862	.791	2.727	8.974	1.130	3.896	8.079
	3	1.577	5.439	20.301	.700	2.414	11.388	.960	3.309	11.388

Extraction Method: Principal Axis Factoring

Table 3. Rotated factor matrix of the developed test

1st validation				2nd validation			
Test Item	Factor 1	2	3	Test item	Factor 1	2	3
OC_M 35	.451			OC_45	.481		
DA_G 23	.438			DA_NS 18	.394		
OC_PA 41	.405			I_G 11	.378		
OC_NS 33	.308			OC_40	.356		
DA_SP 30	.282			I_PA 14	.309		
DA_G 24				OC_33			
OC_M 36				DA_NS 29			
I_M 6				I_NS 1			
DA_NS 18				OC_35			
DA_SP 29				I_M 6		.359	
I_SP 15				I_G 12		.338	
DA_M 19				DA_NS 26		.317	
OC_PA 40				I_PA 15		.315	
I_M 5				DA_NS 24		.312	
I_PA 12		.386		OC_36			
OC_SP 43		.369		DA_NS 30			
OC_G 39		.355		I_G 9			
OC_NS 32		.336		DA_NS 23			
DA_NS 17		.283		I_M 8			.382
I_G 9				DA_NS 27			.324
I_PA 11				OC_39			.308
OC_NS 31				I_NS 3			
I_SP 14				I_M 5			
DA_G 22				OC_32			
I_NS 1				OC_43			
DA_M 20				I_NS 4			
DA_PA 26				OC_41			
OC_M 34				DA_NS 17			
I_NS 3				DA_NS 22			
OC_PA 42							
I_NS 2							
OC_G 38			.330				
OC_SP 45							
DA_PA 27							
I_M 4							
I_G 8							
I_M 7							
DA_M 21							
DA_SP 28							
I_SP 13							
I_PA 10							

For the explanatory power of the model, the variance and eigenvalues obtained from confirmatory factor analysis using three factors are in Table 2 for first and second validation. The Eigenvalue reflects the quantity of extracted factors whose sum is equal to the number of items subjected to factor analysis (Chetty & Datt, 2015). The most critical analytical decisions at this stage are how many considerations to keep (Hayton, et al., 2004). Three considerations have been kept in this analysis, as there are three aspects of critical thought. However, there is conceptual

and empirical evidence that both specifying too few factors and specifying too many are substantial errors that affect results (Hayton, et al., 2004).

All types of misspecifications have empirically been shown to result in poor factor-loading pattern replication and understanding (Velicer, et al., 2000). Table 2 also presents the three factors extracted from the developed test. The first factor accounts for 4.064%, the second factor 2.292%, and the third factor 1.718%, accounting for 14.476% of the variable for first validation and the first factor 9.146%, the second factor 5.716% and the third factor 5.439% for the 2nd validation. Moreover, using the Eigenvalue > 1 criteria, 18 and 13 factors were extracted, accounting for 59.393% and 58.513% of the 1st and 2nd validation variance, respectively. It indicates that the proposed mathematics critical thinking test measures other dimensions and factors aside from Induction, Deduction and Assumption, and Observation and Credibility.

In factor analysis, the extraction method was done through principal axis factoring in Varimax rotation with Kaiser Normalization. Also, the rotation was to reduce the number of factors on which the variables under investigation have high loadings (Chetty & Datt, 2015). Moreover, Table 3 presents the rotated factor matrix of the developed test. Factor 1 had five items, factor 2 5 items, and factor 3, only 1 item. In each factor, two/three critical thinking dimensions are in each grouping. It implies that the test items produced are not seen as one element but are multidimensional. Further, other factors may use variables for further analysis.

2. Discussions

This study considered multiple sources and pieces of evidence in achieving the validity of the developed test. Content and face validation were achieved through expert evaluation and utilizing the table of specifications. For content validation, it was done through the index of difficulty and index of effectiveness, for construct validation by factor analysis and criterion validation through correlation using the equivalence with the Cornell Critical Thinking Skills Test. Not all validity evidence yielded within the range of accepted standards, as provided in the relevant literature. The developed test on Mathematics Critical Thinking Test may be acceptable and valid for assessing critical thinking; moreover, further investigation to improve its reliability and strengthen its validity evidence is needed.

These results may be indicative of the interplay of several factors, including the types of students who answered the test. The nature of the respondents is one of the factors affecting the validity and reliability of a study (Fabrigar, et al., 1999). Also, another factor might be on the selection of the respondents. Guidelines on sample size calculations were followed based on the number of measure variables included in the analysis with more measure variables requiring larger sample sizes (Fabrigar, et al., 1999).

E. Conclusion

In this paper, we demonstrated that critical thinking in mathematics could be useful as an assessment tool in determining students' critical thinking skills, which would eventually be essential in the teaching-learning process.

The developed critical thinking skills test in mathematics has undergone a rigorous revision process based on mathematics content experts' suggestions, the result of pilot–testing, and the actual implementation phase. From an initial 45 items, the final form of the test consisted of 30 items with different levels of difficulty and an acceptable level of discrimination indices. The sources of validity included face and content validity by mathematics education experts, criterion-related evidence, and construct-related evidence using factor analysis. The developed Critical Thinking Test for mathematics may be used valid for assessing the first three dimensions of critical thinking based on Ennis' model, namely: Induction, Deduction and Assumptions, and Observation and Credibility. The results of factor analysis showed that other factors accounted for the variation in test scores; thus, the developed mathematics critical thinking test measures other dimensions of Critical Thinking aside from those measured in the Cornell Critical Thinking dimensions.

Since validity and reliability is evidence-based while Classical Test Theory is sample-dependent, this study recognized possible limitations on the nature and sampling of respondents, which was heterogeneous with different levels of academic performance and coming from various local public schools.

Despite the ease and practicality of using the test in gathering quantitative data for the development of a mathematics critical thinking test, several limitations were identified, which may have affected the results. Validity measures based on Classical Test Theory are sample dependent. Student respondents may not have been 100% focused on answering the test in the pilot administration of the test, as it was evident of skipped questions in their answer sheets. Further, most students finished and submitted at a very early time compared to its time allotment. Despite these limitations, results have shown various forms of validity evidence for assessing critical thinking.

F. Acknowledgment

G. References

Adnyani, L. P. A. P. (2020). Applying cognitive conflict strategy to develop mathematical critical thinking ability and character of students. *Journal of Mathematics Education*, *5*(1), 30–38. DOI: 10.31327/jme.v5i1.1174

Alcantara, E. C., and Bacsa, J. P. (2017). Critical Thinking and Problem-Solving Skills in Mathematics of Grades 7 Public Secondary Schools. *Asia Pacific Journal of Multidisciplinary Research*, Vol.5, No.4, 21-27

Aktas, G. S., & Unlu, M. (2013). Critical Thinking Skills of Teacher Candidates of Elementary. *Procedia-Social and Behavioral Science*, 93; 831-835

Angeli, C., & Valanides, N. (2009). Epistemological and methodological issues for the conceptualization, development, and assessment of ICT-TPCK: Advances in technological pedagogical content knowledge (TPCK). *SAGE Publication*, 52, 154–168.

Behar-Horenstein, L. S. & Niu, L. (2011). Teaching Critical Thinking Skills in Higher Education: A Review of the Literature. *Journal of College Teaching and Learning* Vol.8; 25-41

Belecina, R. R. (2017). Portfolio AS An Alternative Assessment: Effects on Problem-Solving Performance, Critical Thinking, and Attitude in Mathematics. *Philippine Normal University Journal on Teacher Education* Vol.3, No.1, 54 - 81

Colton, D., & Covert, R. (2007). *Designing and Constructing Instruments For Social Research and Evaluation* (pp. 321-325). San Francisco, California: Josey-Bass A Wiley Imprint.

DepEd. (2010). *DepEd K to 12 Mathematics Curriculum Guide.*

Dwyer, C. P., Hogan, M. J., & Stewart, I. (2014). *An Integrated Critical Thinking Framework for the 21st century*. Thinking Skills and Creativity, 12; 43–52.

Education Bureau. (2003). *Progress Report on the Education Reform.* Hong Kong: Education Bureau.

Ennis, R. (1996). *Critical Thinking.* Upper Saddle River, New Jersey: Prentice Hall Inc.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological methods*, *4*(3), 272. Retrieved from http://personal.psu.edu/jxb14/M554/articles/Fabrigaretal1999.pdf

Fives, H. & Barnes, N.D. (2013). Classroom Test Construction: The Power of a Table of Specifications. *Practical Assessment Research and Evaluation*. 18; 1-7

Flores, E. R. (2016). Thinking Skills Reflected in the Argumentative Essays of Freshmen College Students: A Descriptive Analysis. *The Asia-Pacific Education Researcher*, Vol. 16 No. 1; De La Salle University-Manila, Philippines

Halpern, D. F. (2003). *The "How" and "Why" of Critical Thinking Assessment.* New Jersey: Hampton Press.

Halpern, D. F. (2003). *Thought and knowledge: An introduction to critical thinking (4th ed.).* New Jersey: Laurence Erlbaum Associates.

Halpern, D. (2010). *Halpern Critical Thinking Assessment*. Retrieved from sites.google.com: http://www.schuhfried.com/vienna-test-system-vts/all-tests-from-a-z/test/hcta-halpern-critical-thinking-assessment-1/

Hayton, James & Allen, David & Scarpello, Vida. (2004). *Factor Retention Decisions in Exploratory Factor Analysis: A Tutorial on Parallel Analysis*. Organizational Research Methods. 7. 191-205. DOI: 10.1177/1094428104263675.

Listiani, W. (2016). The enhancement of mathematical critical thinking skills and self-efficacy at senior high school students through learning- based problems contextual model. *Journal of Mathematics Education*, *1*(2 July), 55–61. DOI: 10.31327/jomedu.v1i2 July.48

Maandig, R. B., Lomibao, L. S. & Luna, C. A. (2017). Structure Content Reading Instruction vs. Direct Instruction: Their Implication on Students' Achievement, Reading comprehension, and Critical Thinking in Mathematics. *American Journal of Educational Research*, Vol.5, No.5, 574-578

Marquez, L. (2017). *Critical Thinking in Philippine Education: What We Have and What We Need.* University of the Philippines, 272-302

Paul, R. (1993). *What Every Person Needs to Survive in a Rapidly Changing World.* USA: Foundation for Critical Thinking.

Pithers, R. T., & Soden, R. (2010*).* Critical Thinking in Education: A Review. *Educational Research*, 42, 237-249.

Retnawati, H., Hadi, S., Nugraha, A. C., Ramadhan, M. T., Apino, E., Djidu, H., Wulandari, N. F., & Sulistyaningsih, E. (2017). *Menyusun laporan hasil asesmen pendidikan di sekolah: referensi untuk pendidik, mahasiswa, & praktisi pendidikan*. UNY Press.

Rosnawati, R. Kartowagiran, B. & Jailani (2015). A Formative Assessment Model of Critical Thinking in Mathematics Learning in Junior High School*. Research and Evaluation in Education Journal*, Vol.1, No.2, 186-198

Rubio, D. M., Berg-Werg, M., Tebb, S. S., Lee, E., & Rauch, S. (2003). Objectifying Content Validity: Conducting a Content Validity Study in Social Work Research*. Social Work Research*, 27(2), 94-104.

Salkind, N. J. (2010). *Encyclopedia of research design* (Vols. 1-0). Thousand Oaks, CA: SAGE Publications, Inc. DOI: 10.4135/9781412961288

Schafersman, S. D. (1991). *Introduction to Critical Thinking*. Retrieved from https://vcut.org/893-index.pdf

Tamayo, J. D., Tayao, J-r, C., Lopez, M. Y. & Mangalile, M. J. (2014). A Cross-Sectional Analysis of the Level of Critical Thinking of Students in the College of Management and Technology at CEU Malolos*. International Journal of Scientific and Research Publications*, Vol.4, Issue 5 ISSN 2250-3153

Tang Keow Ngang a*, S. N. (2014). Developing instruments to measure thinking skills and problem-solving skills among Malaysian primary school pupils. *Procedia Social and Behavioral Sciences*, 3 7 6 0 – 3 7 6 4.

Thompson, B. & Vacha-Haase, T. (2000). *Psychometrics is data metrics: The Test is Not Reliable*. Educational and Psychological Measurement. 60, 174-195

UCLA Institute for Digital Research & Education Statistical Consulting (2020). *What does Cronbach's Alpha Mean?.* Retrieved from: https://stats.idre.ucla.edu/spss/faq/what-does-cronbachs-alpha-mean/

Velicer, W. F., Eaton, C. A., & Fava, J. L. (2000). Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. In *Problems and solutions in human assessment* (pp. 41-71). Springer, Boston, MA. DOI: 10.1007/978-1-4615-4397-8_3

Yildirim, B., & Ozkahraman, S. (2011*).* Critical Thinking Theory and Nursing Education*. International Journal of Humanities and Social Science*, Vol. 1 No. 17.