EDUCATUM JSMT Vol. 9 Special Issue (2022)
ISSN 2289-7070 / e-ISSN 2462-2451 (38-47)
https://ejournal.upsi.edu.my/index.php/EJSMT/index

A Systematic Review on Instruments to Assess Critical Thinking & Problem-Solving Skills

Alysuliana Alias[1], Lilia Ellany Mohtar[2*], Shahrul Kadri Ayop[3] & Fanny Rahmatina Rahim[4]

[1,2,3]Department of Physics, Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, 35900 Tanjong Malim, Perak, MALAYSIA
[4]Department of Physics, Faculty of Mathematics and Natural Sciences, Universitas Negeri Padang, Jl. Prof Hamka, Padang 25131, INDONESIA

[*]*Corresponding author: lilia@fsmt.upsi.edu.my*

Published: 03 Jun 2022

To cite this article (APA): Alias, A., Mohtar, L. E., Ayop, S. K., & Rahim, F. R. (2022). A Systematic Review on Instruments to Assess Critical Thinking & Problem-Solving Skills. *EDUCATUM Journal of Science, Mathematics and Technology*, *9*, 38-47. https://doi.org/10.37134/ejsmt.vol9.sp.5.2022

To link to this article: https://doi.org/10.37134/ejsmt.vol9.sp.5.2022

Abstract

Critical Thinking and Problem Solving (CTPS) are soft skills essential to be equipped among students according to 21st-century learning. Several instruments have been developed to measure CTPS skills among students in various subjects. This review aimed to identify the type of instrument and the context they are used. The research methodology was based on the systematic literature search in online databases, mainly in Scopus and ERIC, complemented by Google Scholar and MyCite. The keywords used during the searching process included "instrument", "test", "assessment", "critical thinking", and "problem-solving". The screening process took into consideration the type of publication in which only the articles written in Malay or English language within the year 2017 to 2021 were considered. The result of this review revealed that the test, rubric, observation sheet, and questionnaire were the most used types of instruments. Meanwhile, the contexts of measurement were by group discussion, experimental report, quiz, and answer sheet. This review can help educators and researchers make informed decisions about choosing the appropriate CTPS instrument and context for their assessments.

Keywords: instrument, CTPS skills, assessment, critical thinking and problem-solving

INTRODUCTION

According to the survey by World Economic Forum, there are 15 top demanded skills that employers seek for in the years leading up to 2025. Among the skills are critical thinking and analysis and problem-solving. These two skills have consistently remained at the top of the chart since 2016 [1]. Critical thinking and problem-solving are fundamentally different skills. Critical thinking is the process of analysing, evaluating, or synthesising relevant information to form an argument or reach a conclusion supported by evidence [2].

Meanwhile, problem-solving is the ability to identify, analyse and solve the problem [3]. However, someone with critical thinking skills will have an impact on their problem-solving abilities. According to research, good critical thinkers make better decisions and judgements when dealing with complex problems [4]. As a result, some academicians combine critical thinking and problem-solving abilities in one phrase. In Malaysia, critical thinking and problem-solving (CTPS) was initially introduced by the Ministry of Higher Education (MOHE) through a soft skills-module in 2006 [5].

Most educational institutions do not teach CTPS skills directly. Instead, they are applied through specific subjects. One of the subjects that require mastery of CTPS skills is Physics. Abstract physics concept requires students to master CTPS skills to understand the subject [6]. Each subject or field has a different way of applying CTPS skills. For example, humanities field is not the same as science, although CTPS skills are essential for both [7].

Although the concept of CTPS skills was introduced centuries ago, assessing these skills is more complicated than expected [8]. The concept of CTPS has different definitions or characteristics depending on the particular context or discipline [9]. Mastery of CTPS skills in a subject is also challenging if it is not supported by measurement tools that can help educators assess the skills accurately [10].

There are various types of standardised tests for CTPS assessment available in the market, such as Cambridge Thinking Skills Assessment (TSA), Ennis-Weir Critical Thinking Essay Test, Watson-Glaser Critical Thinking Appraisal (WGCTA), California Critical Thinking Skills Tests (CCTST), Cornell Critical Thinking Test (CCTT) and Halpern Critical Thinking Assessment and others. Most of these standardised tests are available on relevant websites for free or require payment as listed in Table 1.

Table 1: Availability of standardised critical thinking & problem-solving instrument

Instruments	**Availability**	**URL**
Cambridge Thinking Skills Assessment (TSA)	Cambridge Assessment Admission Testing - open access	https://www.admissionstesting.org/for-test-takers/thinking-skills-assessment/tsa-cambridge/preparing-for-tsa-cambridge/
Ennis-Weir Critical Thinking Essay Test	Academia.edu - open access	https://www.academia.edu/1847582/The_Ennis-Weir_Critical_Thinking_Essay_Test_An_Instrument_for_Teaching_and_ Testing
Watson-Glaser Critical Thinking Appraisal (WGCTA)	Practice Aptitude Tests - open access (for 1st set only) - paid access (for complete 19 set)	https://www.practiceaptitudetests.com/testing-publishers/watson-glaser
California Critical Thinking Skills Tests (CCTST)	Insight Assessment - paid access	https://www.insightassessment.com/article/california-critical-thinking-skills-test-cctst-2
Cornell Critical Thinking Test (CCTT) Level Z	The Critical Thinking Co - available in book version	https://www.criticalthinking.com/cornell-critical-thinking-test-level-z.html
Halpern Critical Thinking Assessment	Vienna Test System - paid access	https://marketplace.schuhfried.com/en/hcta

However, one thing that should be acknowledged is not necessarily a standard test that is suitable for all measurement contexts [11]. In general, existing measurement tests are more convenient, but they are less suitable for use in teaching and learning in the classroom. Such tests do not match the content taught in the classroom [12]. To optimise CTPS skills, the measuring instrument and the context used need to be chosen precisely [13]. Therefore, some researchers have developed their own CTPS measurement instruments that are more specific and suitable for their subject or learning area.

This study presents a systematic review of instruments to assess critical thinking (CT) or critical thinking and problem-solving (CTPS). Although several reviews on CTPS skills have been conducted, most of them have focused on interventions and strategies to improve the skills. CTPS assessment must also be considered to ensure that CTPS is clearly embedded in schools or higher education institutions. Educators may face difficulty to decide how they can effectively assess CTPS among their students. Some instruments are too general and not specific for their course. Even if they decided on a particular type of instrument, the context of measurement may not align with the nature of their course or not be feasible to implement. Therefore, the purpose of this study was to identify the types of instruments and contexts used to assess CTPS skills that were reported over the last five years. This article reviews the benefits or drawbacks of

each type of instrument and the measurement context used. The outcomes of this discussion can help educators to make informed decisions on the suitable type of CTPS instrument that they can adapt or adopt and the appropriate measurement context.

METHODOLOGY

This study was conducted through a systematic review method. The method started with formulating the research questions, followed by three systematic searching strategies: identification, screening, and eligibility [14]. Finally, the data were extracted and analysed. Two research questions were formulated to frame this study: (1) What type of instrument does the researcher develop to assess CTPS skills? and (2) What is the context used to measure the CTPS skills using the instrument?

The keywords used to do the searching were "development", "instrument", "test", "assessment", "critical thinking", and "problem-solving". Researcher also used Malay keywords such as "pembangunan", "instrumen", "ujian", "penilaian", "pemikiran kritis" and "penyelesaian masalah" to find the articles coming from the researcher's country which is Malaysia. The selected main indexing databases were Scopus and ERIC. In addition, Google Scholar and MyCite databases complemented the search. The databases were chosen following the 14 leading databases suggested by [15].

The four stages of the searching process are shown in Figure 1, adapted from [14]. In the identification process, the result from the main database was 1035 and from manual searching in the complementary databases was 30. In total, 1065 articles were retrieved in the first stage. 14 duplicated articles from the main and complementary databases were removed.

The remaining 1051 articles were screened based on several inclusion and exclusion criteria. The first criterion was the literature type. This review focused on the reviewed journal article. The conferences paper, review, book chapter, book, conference review, and editorial were excluded. The second criterion was the subject area. This review excluded research area other than social sciences and education. The third criterion was the publication date. Articles published before the year 2017 were excluded. The last criterion in the screening process was the publication language. Only articles in English and Malay were included. 927 articles did not fit the criteria and were excluded.

Consequently, 124 articles were proceeded for eligibility stage. The title, abstract and the main content of the articles were examined to ensure the conformation of the inclusion criteria to achieve the study objectives. Consequently, 108 articles were excluded as they were not focused on the type and context of instruments. Finally, a number of 16 articles were ready to be analysed. The data were extracted to answer the research questions. The finding is tabulated as in Table 2.

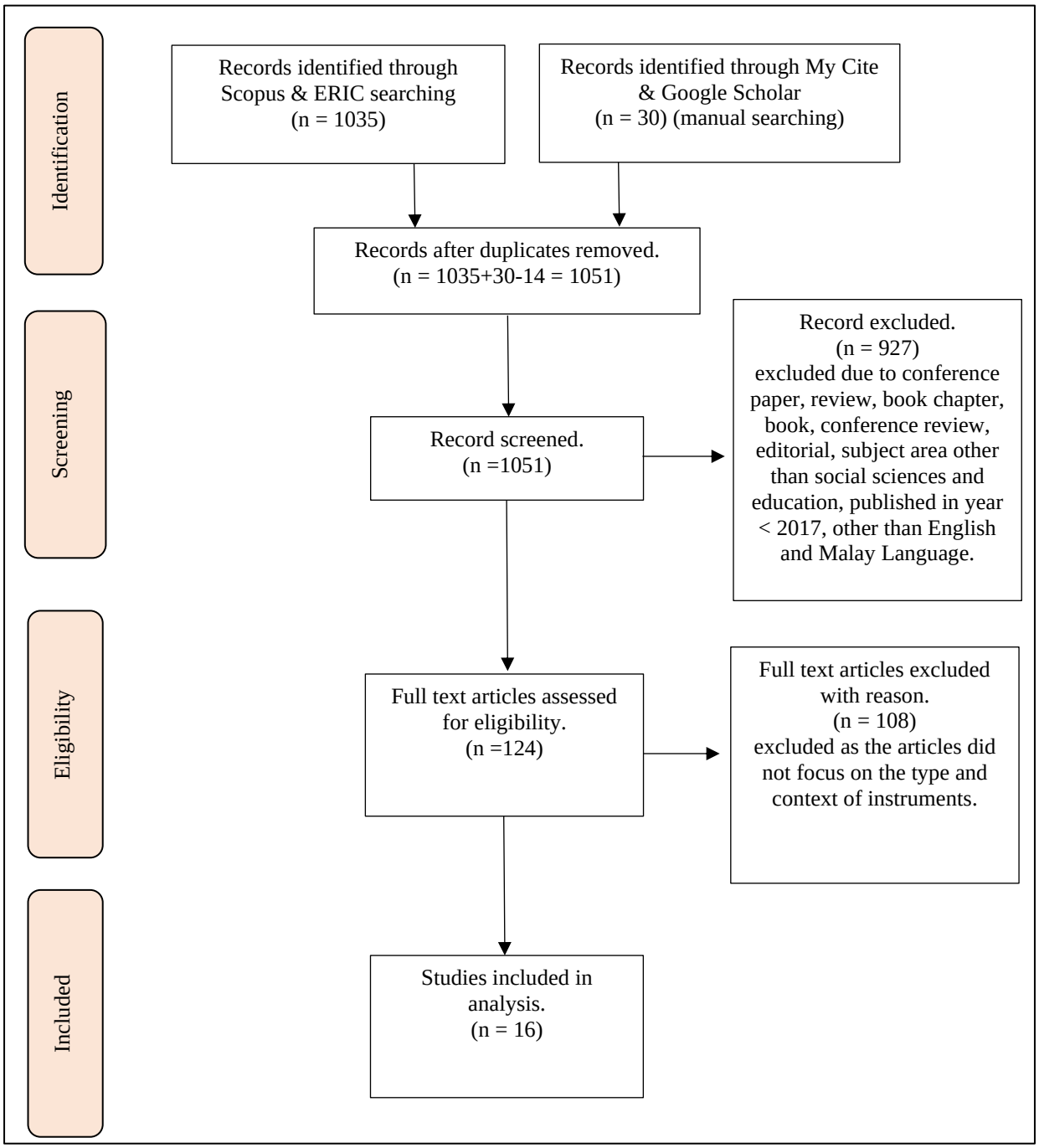

Figure 1 The flow diagram of the study (Adapted from [14])

RESULTS AND DISCUSSION

Table 2: Findings of the study

Study	Country	Subject (Level)	Type of instrument		Context of measurement	
			Name	Code	Name	Code
[16]	Indonesia	Physics (High School)	Observation Sheet, Test	T	Lab Report	LR
[2]	USA	STEM (Undergraduate)	Rubric	R	Lab Report	LR
[17]	India	Engineering (Undergraduate)	Rubric	R	Mind Mapping Exercise	AS
[18]	U.S.A.	Marketing (Undergraduate)	Homework, Quiz, Blueprint, Final Project	T	Composite scores (Answer sheet)	AS
[19]	Turkey	Mathematics (Fifth grades)	Observation Sheet, Rubric	OS R	Online Group Discussion (recording), quiz	GD QZ
[20]	Itali	English Language (Undergraduate)	Test	T	Answer Sheet	AS
[21]	Taiwan	Mechanical (High School)	Questionnaire	QS	-	-
[3]	Hong Kong	General (Undergraduate)	Questionnaire	QS	-	-
[22]	Malaysia	Accounting (Pre-university)	Rubric	R	Answer Sheet	AS
[23]	Malaysia	Co-curricular (Pre-university)	Rubric	R	Co-Curricular Activity	*
[24]	Malaysia	TESL (Undergraduate)	Observation Sheet	OS	Group discussion (Peer- Socratic Questioning)	GD
[25]	Belgium	Physics (Undergraduate)	Test	T	Answer Sheet	AS
[26]	Thailand	Computer Science (Undergraduate)	Rubric	R	Answer Sheet	AS
[27]	South-Eastern Country	General (Undergraduate)	Rubric	R	Online Group Discussion	GD
[28]	Indonesia	Chemistry (High School)	Test	T	Open-Ended Two-Tier Multiple-Choice Question Items	QZ
[29]	U.S.A	Business (Undergraduate)	Rubric	R	Answer Sheet, Role-Play Exercises (recording)	AS & *

Based on the Table 2, four types of instruments were developed to measure students' CTPS skills: a paper test, rubric, observation sheet, and questionnaire. The rubric type was the most reported. According to [30], rubric is the most suitable alternative assessment tool in 21st-century learning. The use of rubrics in assessment process allows educators to assess students more systematically [26], facilitates the process of providing feedback to students on their achievement [31], and as evidence-based tools in measuring student achievement levels [2]. Specifically, a rubric is defined as an assessment tool used as a guide to assess student progress through various contexts such as writing, oral presentation or group work skills [32].

The second most used type of instrument was the test. Although many different assessment methods are used in education, the test method is typically the most preferred method by educators [33]. There are several types of optional tests, including multiple-choice questions (MCQ) and subjective tests. The advantage of MCQ test is it is easier to administer [7]. However, from the perspective of the respondents and students, they only picked the correct answer but the researcher was not able to find out the reasons and criteria used by the students in formulating the chosen solution [34]. The challenge for subjective testing is to obtain high-reliability values using inter-rater reliability analysis, which necessitates strict procedures to minimise the differences between examiners [35]. According to [36], paper-pencil test only measures knowledge and neglects the aspect of skills. This statement is supported by [37], who explained that paper-pencil test is more likely to measure cognitive element.

The next type of instrument was the observation sheet. Sometimes people categorise the observation sheet as rubric. A rubric is more general as it can be used as a marking guideline for quiz and for measurement of students' skills or attitude during some task or activities [38]. Observation sheet is commonly used when the markers need to observe the attitude or the response of the students. Based on this review, two observation checklists were used for group discussion activities and observing lab activities. Direct observation requires educators to be physically present to observe and record events, both verbal and nonverbal, as they happen. The educators obtained first-hand information using the direct-observation method, which was also simple to use and verified from other sources [39]. Nevertheless, this method takes a long time because the educators have to observe each of their students, which must be difficult if the measurement occurs during class activities [40].

The last reported type of instrument was the questionnaire. In general, questionnaire is one of the popular instruments chosen by the researcher. There are some benefits of questionnaire, such as it systematically focuses on the cognitive issues in the response process, it help educators provide more consistent feedback and low-cost enhancement [41]. However, in terms of the assessment related to teaching and learning, questionnaire is not suitable enough. Furthermore, when the students score for themselves, they may underestimate or overestimate the results [3].

The next part of this review is related to the context of measurement. Researchers or educators must determine the appropriate measurement context in addition to knowing the appropriate type of instrument to use. Based on the search, most of the measurement context is the answer sheet. This context is widely used as it can match two instruments: rubric and test. CTPS skills can be measured by looking at the strategy used by students to solve the problem in their answer sheet. For a test, educators usually have their marking scheme to mark students' answers [18]. In contrast with rubrics, the solution is evaluated using the rubric scoring and criteria [26].

The second context of measurement from this review is the group discussion. Group discussion is one of the student activities that can be held in class, outside of the class or during online learning. However, measuring CTPS skills using group discussion may take a long time because we need to measure the group of students simultaneously. According to [40], it is challenging to measure CTPS skills during face to face class, and one way to overcome it is by online discussion. Other than that, video recording can also help the assessment process, in which the students need to record their group discussion and submit it to the teacher for evaluation [19]. The recording context for performing the observations simplifies and eliminates the scoring errors for students. In addition, since the video recording can be replayed, it allowed the examiner to be careful while assigning the score to the students.

The third context is quiz. A quiz can be a type of instrument and, at the same time as the context of measurement. CTPS measurements obtained through quizzes are, in essence, no different from tests. It is just that quizzes are given out more frequently than tests. According to [42], quizzes are a measurement context that must be used daily. Students can answer the quiz right after the teaching and learning process has been completed. The advantage of this context is that educators can set the quiz questions related to the

subject content on the day and in line with CTPS skills. However, the challenges are definitely related to the time allocation for educator's preparation.

Lab report is also one of the contexts used to measure CTPS skills. There is a significant relationship between CTPS and the science process skills. It was proven that science process skills and critical thinking skills are inextricably linked because students who lack science process skills will lack critical thinking skills [16]. Science process skills are applicable during the practical mode when the students run the experiment and write the lab report. From this review, one of the lab reports was measured by using the observation sheet and test questions [16], and another one was by using rubrics [2]. All instruments have their strengths, but in terms of establishing science process skills with the score related to CTPS skills, it is more convenient to use a rubric with their flow of criteria and scoring strategies [38]. However, this context is limited to the subject related to science process skills only.

There were two other contexts of measurement used in assessing students' CTPS skills in this review. One was extra-curricular activity, and another one was role-playing exercise. Both were measured using rubrics. However, this context may seem challenging, as only one study had mentioned it. Co-curricular activities are specific to co-curricular subjects and should not be used for other subjects. Role-play activities are like group discussion activities. Educators will look at the recording version to analyse CTPS skills of the students. The concept of role-playing is broad, but it depends on the subject matter. [29] used the role-play exercise in a business study in which the students were required to work as a group to solve business-related problem. They had to agree on a solution and decide what steps they would take to resolve the problem. This method is very creative and fits the business topic. It can also be used in other subjects, but teachers must establish the goal of role-play and the problems that the students must solve.

Overall, based on the discussion of the four types of instruments developed by previous researchers, it is possible to conclude that each instrument has its benefits or drawbacks. Selecting an appropriate measurement context is also critical, as each measurement context has different strengths and challenges. Educators and researchers who need to decide a type of instrument and set the context of measurement should weigh the benefits, drawbacks, and challenges that have been discussed and then adapt to the subject of their respective fields. Suggesting the best method for assessing CTPS skill level is actually depends on the measurement objectives. Measurement is the process of obtaining a numerical explanation of how much an individual or student has characteristics to be measured using the instruments provided [43]. Therefore, the right instrument is needed to produce the right results.

In the context of measuring CTPS skills through 21st-century learning, authentic assessment is more encouraged. Authentic assessment is an assessment conducted continuously and is realistic that it does not only rely on answer sheets alone. Instead, this assessment involves skills that students highlight during activities in or outside the classroom. For the mentioned case, rubrics are the recommended instrument [30]. The rubric provides structured standards and ensures educators conduct assessments systematically. However, the rubric also minimises a student's true potential because it limits the degree to which an individual conforms. Rubric, without realising it, becomes an instrument that forms the standard of assessors and can be subjective [44].

CONCLUSIONS

This review summarised reported studies on the CTPS measurement in specific subject using various types of instruments within a context of measurement. In conclusion, the four common types of instruments to assess CTPS skills in education are paper test, rubric, observation sheet and questionnaire. Among the context of measurement were student answer sheet, quiz/ test, experiment report and group discussion. Rubric was the most used instrument to measure the CTPS due its adaptation flexibility on the wide range of the context of measurement such as group discussion, role-play, and student process skills. The selection of appropriate type of instrument and correct match with the context of measurement can influence the success in the evaluation process of CTPS skills among students. The most important is the application of CTPS in education and the suitable measurement tools that can help improve students' CTPS skills.

ACKNOWLEDGEMENT

The first author would like to acknowledge the Ministry of Education for providing Hadiah Latihan Persekutuan (HLP). However, this research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

[1] World Economic Forum. (2020). *The future of jobs report*. https://www.weforum.org/reports/the-future-of-jobs-report-2020

[2] Reynders, G., Lantz, J., Ruder, S. M., Stanford, C. L., & Cole, R. S. (2020). Rubrics to assess critical thinking and information processing in undergraduate STEM courses. *International Journal of STEM Education*, *7*(9), 1–15. https://doi.org/10.1186/s40594-020-00208-5

[3] Chan, C. K. Y., & Luk, L. Y. Y. (2020). Development and validation of an instrument measuring undergraduate students ' perceived holistic competencies. *Assessment & Evaluation in Higher Education*, *46*(3), 467–482. https://doi.org/10.1080/02602938.2020.1784392

[4] Dwyer, C. P., Hogan, M. J., & Stewart, I. (2014). An integrated critical thinking framework for the 21st century. *Thinking Skills and Creativity*, *12*(June 2014), 43–52. https://doi.org/10.1016/j.tsc.2013.12.004

[5] Kementerian Pengajian Tinggi Malaysia. (2006). *Modul pembangunan kemahiran insaniah (soft skills) untuk IPT Malaysia*. Universiti Putra Malaysia.

[6] Docktor, J. L., Strand, N. E., Mestre, J. P., & Ross, B. H. (2015). Conceptual problem solving in high school physics. *Physical Review Special Topics - Physics Education Research*, *11*(2), 1–13. https://doi.org/10.1103/PhysRevSTPER.11.020106

[7] Rear, D. (2019). One size fits all? The limitations of standardised assessment in critical thinking. *Assessment & Evaluation in Higher Education*, *44*(5), 664–675. https://doi.org/10.1080/02602938.2018.1526255

[8] Wren, D., & Cashwell, A. (2018). Mission Possible: Measuring Critical Thinking and Problem Solving. *Educational Leadership*, *75*(5), 70–75.

[9] Mueller, J. F., Taylor, H. K., Brakke, K., Drysdale, M., Kelly, K., Levine, G. M., & Ronquillo-Adachi, J. (2020). Assessment of Scientific Inquiry and Critical Thinking: Measuring APA Goal 2 Student Learning Outcomes. *Teaching of Psychology*, *47*(4), 274–284. https://doi.org/10.1177/0098628320945114

[10] Mabruroh, F., & Suhandi, A. (2017). Construction of Critical Thinking Skills Test Instrument Related The Concept On Sound Wave. *Journal of Physics: Conference Series*, *812*(012056). https://doi.org/10.1088/1742-6596/755/1/011001

[11] Alsaleh, N. J. (2020). Teaching Critical Thinking Skills: Literature Review. *TOJET: The Turkish Online Journal of Educational Technology*, *19*(1), 21–39. http://www.tojet.net/articles/v19i1/1913.pdf

[12] Larsson, K. (2021). Using Essay Responses as a Basis for Teaching Critical Thinking-a Variation Theory Approach. *Scandinavian Journal of Educational Research*, *65*(1), 21–35. https://doi.org/10.1080/00313831.2019.1650824

[13] Istiyono, E., Dwandaru, W. S. B., Lede, Y. A., Rahayu, F., & Nadapdap, A. (2019). Developing IRT-based physics critical thinking skill test: A CAT to answer 21st century challenge. *International Journal of Instruction*, *12*(4), 267–280. https://doi.org/10.29333/iji.2019.12417a

[14] Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Guidelines and guidance preferred reporting items for systematic reviews and meta-analyses:The PRISMA statement. *PLOS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

[15] Gusenbauer, M., & Haddaway, N. R. (2020). Which academic search systems are suitable for systematic reviews or meta-analyses ? Evaluating retrieval qualities of Google Scholar , PubMed , and 26 other resources. *Research Syntheses Methods*, *11*(2), 181–217. https://doi.org/10.1002/jrsm.1378

[16] Kurniawan, W., Pathoni, H., Muliawati, L., Kurniawan, D. A., Romadona, D. D., Ningsi, A. P., &

Dari, R. W. (2020). Relationship of science process skills and critical thinking of students in physics Subject. *Universal Journal of Educational Research*, *8*(11), 5581–5588. https://doi.org/10.13189/ujer.2020.081162

[17] Priyaadharshini, M., & Sundaram, B. V. (2018). Evaluation of higher-order thinking skills using learning style in an undergraduate engineering in flipped classroom. *Computer Applications in Engineering Education*, *26*(6), 2237–2254. https://doi.org/10.1002/cae.22035

[18] Nonis, S. A., & Hudson, G. I. (2018). Developing and assessing critical thinking skills in marketing students : The power of making explicit problem-solving processes. *Journal of Education for Business*, *94*(3), 195–203. https://doi.org/10.1080/08832323.2018.1504737

[19] Dolapcioglu, S., & Doganay, A. (2020). Development of critical thinking in mathematics classes via authentic learning: An action research. *International Journal of Mathematical Education in Science and Technology*. https://doi.org/10.1080/0020739X.2020.1819573

[20] Re, M. R., Amenduni, F., Medio, C. De, & Valente, M. (2019). How to use assessment data collected through writing activities to identify participants' critical thinking levels. *Journal of E-Learning and Knowledge Society*, *15*(3), 117–132. https://doi.org/10.20368/1971-8829/1135051

[21] Kuang, C. Y., Kuen, Y. L., & Shu, F. C. (2017). The development and validation of a mechanical critical thinking scale for high school students. E*URASIA Journal of Mathematics Science and Technology Education*, *13*(5), 1361–1376. https://doi.org/10.12973/eurasia.2017.00675a

[22] Uvathi Mariappan, Arsaythamby Veloo, & Kanageswari Suppiah Shanmugam. (2018). Pembinaan dan pengesahan instrumen kemahiran insaniah bagi subjek perakaunan di matrikulasi. *Jurnal Kurikulum & Pengajaran Asia Pasifik*, *6*(4), 29–39.

[23] Aris Hanafi Abu Jalil, Gunathevan Elumalai, & Mohd Izwan Shahri. (2019). Pembinaan rubrik kemahiran insaniah bagi matapelajaran Kokurikulum pelajar program matrikulasi Kementerian Pendidikan Malaysia. *Jurnal Penyelidikan Dedikasi*, *17*, 126–145.

[24] Zainab Ab Rahman, Hoon, T. S., & Sidhu, G. K. (2019). Using the peer socratic questioning (PSQ) technique to develop critical thinking skills in group discussion. *Jurnal Penyelidikan Tempawan*, *XXXVI*, 91–103.

[25] Tiruneh, D. T., Cock, M. De, Weldeslassie, A. G., Elen, J., & Janssen, R. (2017). Measuring critical thinking in physics : Development and validation of a critical thinking test in electricity and magnetism. *International Journal of Science and Mathematics Education*, *15*(4), 663–682. https://doi.org/10.1007/s10763-016-9723-0

[26] Lertyosbordin, C., Maneewan, S., Yampinij, S., & Thamwipat, K. (2019). Scoring rubric of problem-solving on computing science learning. I*nternational Education Studies*, *12*(8), 26–32. https://doi.org/10.5539/ies.v12n8p26

[27] Bernstein, A. G., & Isaac, C. (2018). Critical thinking criteria for evaluating online discussion. *The International Journal for the Scholarship of Teaching and Learning*, *12*(2), 1–8.

[28] Sadhu, S., & Laksono, E. W. (2018). Development and validation of an integrated assessment for measuring critical thinking and chemical literacy in chemical equilibrium. *International Journal of Instruction*, *11*(3), 557–572. https://doi.org/10.12973/iji.2018.11338a

[29] Bandyopadhyay, S., & Szostek, J. (2019). Thinking critically about critical thinking: Assessing critical thinking of business students using multiple measures. *Journal of Education for Business*, *94*(4), 259–270. https://doi.org/10.1080/08832323.2018.1524355

[30] Azizi Alias, & Kamisah Osman. (2018). *Pentaksiran alternatif: Pembinaan dan pelaksanaan rubrik dalam pendidikan Sains*. Penerbit Universiti Kebangsaan Malaysia.

[31] Huang, B., & Jong, M. S. Y. (2020). Developing a generic rubric for evaluating students' work in STEM education. *2020 International Symposium on Educational Technology (ISET)*, 210–213. https://doi.org/10.1109/ISET49818.2020.00053

[32] Boughey, S. (2019). *Developing rubric to assess 3rd - 5th grade students understanding of science concepts via screencast models*. Indiana University, Bloomington.

[33] Tadele, K., & Sitotaw, B. (2018). Perception and trends in assessment of students' learning in Physics courses. *Latin-American Journal of Physics Education*, *12*(1), 1307–1315. https://doi.org/10.17576/jpen-2018-43.02-09

[34] Fukuzawa, S., & DeBraga, M. (2019). Graded response method: does question type influence the assessment of critical thinking? *Journal of Curriculum and Teaching*, *8*(1), 1–10. https://doi.org/10.5430/jct.v8n1p1

[35] Poce, A., Medio, C. De, Amenduni, F., & Re, M. R. (2019). Critical thinking assessment : A first approach to the automatic evaluation. *18th International Conference on Information Technology Based Higher Education and Training (ITHET)*, 331–338.

[36] Harahap, D. E., Festiyed, & Djamas, D. (2019). Preliminary study on development of assessment performance instruments on physics learning to improve students' critical thinking ability. *Journal of Physics: Conference Series*, *1185*(012128), 1–7. https://doi.org/10.1088/1742-6596/1185/1/012128

[37] Tangkin, W. P. (2019). Pemanfaatan rubrik sebagai instrumen penilaian alternatif. *Scholaria: Jurnal Pendidikan Dan Kebudayaan*, *9*(1), 29–39. https://doi.org/https://doi.org/10.24246/j.js.2019.v9.i1.p29-39

[38] Brookhart, S. M. (2018). Appropriate criteria : Key to effective rubrics. *Frontiers in Education*, *3*(22), 1–12. https://doi.org/10.3389/feduc.2018.00022

[39] Afifah Fadhlullah, & Nurbarirah Ahmad. (2017). Thinking outside of the box: Determining students' level of critical thinking skills in teaching and learning. A*sian Journal of University Education*, *13*(2), 51–70.

[40] Rathakrishnan, M., Ahmad, R., & Choo, L. S. (2017). Online discussion : Enhancing students' critical thinking skills. *AIP Conference Proceedings*, *1891*(1), 020120.

[41] Beatty, P. C., Collins, D., Kaye, L., Padilla, J. L., Willis, G. B., & Wilmot, A. (2019). *Advances in questionnaire design, development, evaluation and testing*. John Wiley & Sons.

[42] Batsell Jr, W. R., Perry, J. L., Hanley, E., & Hostetter, A. B. (2016). Ecological validity of the testing effect : The use of daily quizzes in introductory psychology. *Teaching of Psychology*, *1*(6), 18–23. https://doi.org/10.1177/0098628316677492

[43] Ahmad, A., & Awang, M. I. (2016). *Pengukuran dan penilaian pendidikan*. Dewan Bahasa dan Pustaka.

[44] Tenam-Zemach, M., & Flynn, J. E. (2015). *Rubric nation: Critical inquiries on the impact of rubrics in education*. Information Age Publishing Inc