Developing a Higher Order Thinking Skill-Oriented and Metacognitive-Based Assessment for Vocational School Students

Wening Patmi Rahayu[1], Uswatun Hasanah[1], Rully Wiliandri[2]
[1]Faculty of Economics, Universitas Negeri Malang
[2]Leiden University, Netherlands
Email: wening.patmi.fe@um.ac.id

Abstract: This study aims to provide an assessment instrument oriented to Higher-Order Thinking Skills (HOTS) based on metacognitive skill. The research product was analyzed for feasibility using a validation test of constructs and contents. The assessment instrument was measured with the difficulty and different levels of the questions. A valid product is tested to acknowledge the user's response. The results showed that the construct and content validation is valid, with 94 percent and 91 percent, respectively. The analysis of the different levels of the question showed a good different result. It implies that the instrument points lead to HOTS, which can improve students' metacognitive skills. The revised instrument was tested to the users, with the response was about 86.17 percent. The validation result from the experts and the response of the user can be concluded that the instrument of the HOTS-oriented and metacognitive-based assessment is entirely feasible, so it can be used as a guideline and reference by the teachers in the vocational school to develop similar products.

Keywords: *Higher Order Thinking Skills, HOTS, Instruments Assessment, Metacognitive Skills*

INTRODUCTION

The theme of assessment for higher-order thinking skills (HOTS) has invited attention among scholars in the last decade. An assessment with a high difficulty level will increase the students' critical thinking, both cognitive and metacognitive skills (Narmaditya et al., 2018). It is prominent to research the development of assessment instruments oriented to HOTS based on metacognitive skill due to the result can function as a guideline for the teacher in knowing the high-level thinking skills of the student. It also allows the student ability to answer or solve the problem.

The idea is also supported by a research conducted by Lestari et al. (2016) that the test instrument oriented to HOTS is developed according to the quality criteria that have been set and effectivity level of use. Besides, the research of Budiman and Jailani (2014) produces a HOTS-based assessment instrument that can describe the test question quality of HOTS to measure the high-level thinking skills of students. Also, research by Yerizon (2019) found that the worksheet of math in junior high schools based on (metacognitive approach can effectively improve students' HOTS. Assessment is an activity undertaken to acquire, collect, analyze, and interpret information in any form, which is used as the basis for decision making about the performance of the students, the curriculum, learning programs, and cultural and school policy (Enger & Yager, 2009; Kusaeri, 2014; Uno & Koni, 2012).

The assessment activity could ideally measure the students' competence following the previous competency objectives. The instruments used in measuring the cognitive, affective, and psychomotor development should suit and be oriented to a higher level of difficulty. Besides, it must train the students to think critically. A HOTS-oriented and metacognitive-based assessment instrument is strictly necessary. As explained by Arifin (2014), the assessment activities should encourage the teachers to improve their teaching skills and help students develop optimally. The assessment goal can be used by teachers to monitor and evaluate the process, learning progress, and improvement of student learning outcomes on an ongoing basis (Widana, 2017).

Referring to the Regulation of the Minister of Education and Culture No. 23 of 2016, the implementation of assessment is conducted concerning the principles of whether it is valid, objective, fair, integrated, open, comprehensive, continuous, systematic, oriented to the criteria, and accountable. According to Curriculum 2013 of Indonesia, the revised edition, the assessment of student learning outcomes for secondary education is from the aspect of attitudes, knowledge, and skills. The assessment conducted in this study is the assessment of knowledge. The assessment of knowledge is an activity performed to measure the skills of the learners, such as factual, conceptual, procedural, metacognitive, and High-Order Thinking Skills (Bowler, 2010; Chevron, 2014; Higgins, 2017; Masari & Anghel, 2012; Zumbach et al., 2020).

The revised edition of Curriculum 2013 emphasizes on the Basic Competency (KD) used as the minimum competencies that must be accomplished by the learners. Teachers are required to formulate some indicators as a reference assessment to determine the target achievement of Basic Competence. Under the curriculum, some stages need to be done in developing an instrument of assessment: determining objectives, analyzing the necessary competencies, outlining the essential competencies into indicators, developing the points of the questions based on the indicators developed, creating guidelines for scoring and answer keys, reviewing and revising the questions, and investigating the questions.

To realize the assessment and improvement of the education quality standard in Indonesia based on the implementation goal of the revised edition of Curriculum 2013 as well as 21st-century learning, learners are required to improve their High-Order Thinking Skills (HOTS) in the cognitive domain. Curriculum 2013 Revised Edition requires the implementation of assessments that encourage students to develop answers, not only choosing the answers provided. This is in line with the statement of the Minister of Education and Culture that HOTS is deliberately implemented to encourage the critical culture of students, so their thinking process can be seen from the answer (Jarvis & Baloyi, 2020; Nguyen and Nguyen, 2017; Pilten, 2010; Yang, 2015; Yee et al., 2015). The students' high-order thinking skills are expected to apply in problem-solving to confront the real problems around them.

HOTS is the ability to apply new information or prior knowledge and to manipulate it to reach the possible answers in the new situation (Budiman & Jailani, 2014; Pratiwi & Fasha, 2015; Roets & Maritz, 2017; Widana, 2017). The HOTS of students can be empowered by giving an erratic and unusual problem therefore students can solve it based on their knowledge and experience (Nofiana, 2016; Tal

& Hochberg, 2003; Wu, 2016). According to Kurniati et al. (2016), HOTS is the ability to think in a critical, logical, reflective, metacognitive, and creative way. When viewed from the knowledge dimension, generally, the questions related to HOTS measure metacognitive category, so they do not only measure the factual categories, conceptual, or procedural ones.

The metacognitive category describes the ability to connect several different concepts, interpret, solve problems, select problem-solving strategies, find new methods and arguments, and make the right decision (Widana, 2017). HOTS is a thought process of students in the cognitive and taxonomy level of learning, such as the methods of problem-solving, Bloom's taxonomy, and the taxonomy of learning, teaching, and assessment (Saputra, 2016). (Brookhart, 2010) reveals that HOTS is in "top-end" in Bloom's Taxonomy, which is analyzing (C4), evaluating (C5), and creating (C6). High-order thinking skills include problem-solving, creative thinking, critical thinking, arguing skills, and decision-making skills. Therefore, the processes are necessary to measure the student's knowledge.

In the higher-order thinking processes, the student demonstrates an understanding of information and reasoning, not merely recalling (Jarvis & Baloyi, 2020; Lee, 2014; Wing et al., 2014). According to Widana (2017), the questions of *HOTS* in the context of the assessment to measure skills of (1) transferring one concept to the other ones; (2) processing and applying the information; (3) searching the connection of the different variety of information; (4) using the information to solve problems; and (5) studying the ideas and information critically. The questions of HOTS have three characteristics: (1) measuring the HOTS; (2) using the assessment based on the contextual problem; and (3) using the form of several questions.

The interviews with the teachers of public administration at the State Vocational School 1 Pasuruan indicated that the students' skills are usually done in a written test. It is known that the questions have not been referred to as the students' HOTS thoroughly. The questions given to the students are generally still in the category of low-order thinking skills, which are in the domain of remembering (C1), understanding (C2), and applying (C3). In the multiple-choice questions, most questions given are taken directly from the material in the book and the examples that have been described by teachers, so students only need to remember back what they had learned. Whereas in the essay question, the given questions only ask students to explain the meaning and to mention something from the subject material. Therefore, the students' HOTS is less visible.

The results of the interviews above illustrate that HOTS's implementation for students has not been entirely performed as to what the Government is expected in the Curriculum 2013 Revised Edition. Therefore, it is necessary for assessment instruments that emphasize the creation of questions in the scope of analyzing (C4), evaluating (C5), and creating (C6) to stimulate metacognitive-based HOTS. Surasa et al. (2017) remarked that these studies' results could realize a harmonious correlation between aspects of metacognitive and cognitive. It also requires the contribution and active participation of the teacher.

A prior study by Wardany and Ramli (2017) on the development of assessment to measure Students' HOTS showed that the product of assessment instruments of HOTS is multiple choice test and a multiple-choice developed has confirmed for

validity and reliability with the very-high average interpretation, as well as the quality of the good questions in the level of difficulty and differentiation feature. The research by Arifin and Etnawati (2017); Susanto and Retnawati (2016) revealed that assessment instruments of HOTS developed can show and identify the students' HOTS. The assessment instruments developed by the researcher is the instrument of HOTS-oriented assessment based on Metacognitive, using two forms of the writing test, which is the form of the test is subjective and objective that every item comes with the answer key and the way of HOTS.

METHODS

This research was designed to produce a product of assessment instruments on Public Administration-oriented subjects to HOTS based on metacognitive for class X OTKP in the State Vocational School 1, Pasuruan. The research model employed Research and Development (R & D) by Borg and Gall (2003) and adapting eight of the ten steps, which illustrated in Figure 1.
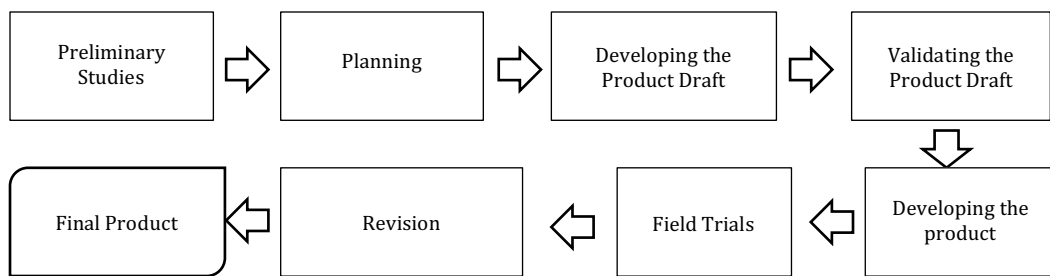
Figure 1. The Charts of the Stages of Research Instrument Development of HOTS-Oriented and Metacognitive-Based Assessment
Source: Adapted from Borg and Gall (2003)

The analyzed data were taken from the validation of the contents, validation of constructs, and user's response in the trial. Content validation was done by a teacher of General Administration in the State Vocational School 1, Pasuruan, while construct validation is conducted by a lecturer of the Management Department, Faculty of Economics, Universitas Negeri Malang. Validation was done by giving a questionnaire assessment of the product to the validator to acknowledge the feasibility of the instrument. The questionnaire assessment was equipped with column comments and suggestions used to revise the assessment instrument.

The results of the user response were taken when they carry out field tests to acknowledge the response of the students to the metacognitive-based HOTS questions presented by giving a questionnaire at the end of the learning as well as to know the level of difficulty and differentiation level of the question for the class X OTKP 2 with 30 students. The data analysis technique used to analyze the quantitative data of the validation assessment results and a subject test on the draft of the developed product uses the following formula (Akbar, 2016).

$$P = \frac{TSe}{TSh} \times 100\%$$

Description:
P   = Percentage
Tse   = The Empirical total Score
Tsh   = The Expected Total Score

The feasibility of the product results in the research and development are identified with a percentage score. The higher the score of the results of the data analysis shows the better level of the product feasibility results. The qualification rate of the product feasibility instruments is presented in Table 1.

Table 1. Feasibility of the Media

The Achievement Value	Validity Category	Description
25.00 – 40.00	Not feasible	not allowed to use
41.00 – 55.00	Less feasible	not allowed to use
56.00 – 70.00	Moderate	allowed to use after major revision
71.00 – 85.00	feasible	allowed to use after minor revision
86.00 – 100.00	Highly feasible	Excellent to use

Source: Adapted from Akbar (2016)

The analysis of the questions' quality employs the analysis of the difficulty level of questions and the analysis of the differentiation features. The difficulty level the question is associated with the number of students who can answer the item correctly. The points which are too easy are not stimulating students to increase the effort to solve it (Arikunto, 2015). The difficulty or easiness of a question is expressed as difficulty index. Thus, upon formulating the question's points, the level of difficulties needs to be empirically seen by the following formula (Arifin, 2014).

$$TK = \frac{(WL + WH)}{(nL + nH)} \text{ x } 100\%$$

Description:
WL   = the number of participants who answer any of the groups below
WH   = the number of participants who answer any of the groups below
nL   = the number of lower groups
nH   = the number of the upper groups

The difficulty level of the description question calculates how many percent of the students who failed to answer correctly or are below the passing grade of every question. The results of the calculations performed can be adjusted with the range of values in Table 2.

Table 2. Range of the Values of Difficulty Levels

Range	The Level of Difficulties
73% to the top	Difficult
28% – 72%	average
0 % – 27%	Easy

Source: Adapted from Arifin (2014)

The test of the differentiation is used to acknowledge whether the questions can distinguish learners with high skills and those with the low one. According to Daryanto (2012), the formula to find the value of distinguishing features.

$$D = \frac{\overline{x}_A - \overline{x}_B}{SMI}$$

Description:
$\overline{x}_A$: The average score of the Upper Groups
$\overline{x}_B$: The average score of the Lower Group
SMI   : Ideal Maximum Score
D   : Differentiation

Besides, the criteria for determining the differentiation are presented in Table 3.

Table 3. The Range of the Differentiation

Range	Differentiation
Of 0.70 < D ≤ 1.00	Excellent
0.40 < D ≤ 0.70	Good
0.20 < D ≤ 0.40	Moderate
0.00 < D ≤ 0.20	Sufficient
Negative (-)	Poor

Source: Adapted from Daryanto (2012)

RESULTS & DISCUSSION

Content Validity and Construct Validity Result
Content validation and construct validation are carried out to determine the HOTS-oriented and Metacognitive-based assessment instrument's content and construction feasibility, which can be accessed in Table 4.

Table 4. Content and Construct Validity Result

	Content Validity (1 teacher)				Construct validity (1 lecturer)		
No	Assessment Indicator	(%)	Criteria	No	Assessment Indicator	(%)	Criteria
1	Average Content	93.3	Very Feasible	1	Language	92	Very Feasible
2	Average Presentation	90	Very Feasible	2	Graphic	93.3	Very Feasible
3	Average Language	86.6	Very Feasible	3	Presentation	100	Very Feasible
	Average Validator Score	90.90	Very Feasible		Average Validator Score	94.54	Very Feasible

Source: Data Processed (2020)

The result of the content validation of the assessment instrument was very feasible with a score of 90.90 percent, and so was the construct validation of the assessment instrument with 94.54 percent. The results of the content validation and construct validation from the two expert validators, both qualitatively and

quantitatively, are used as a reference in making product improvements in instrument creation. Based on the validation result and field trial result, the developed HOTS-oriented and Metacognitive-based assessment instrument are developed content, and construction is feasible to use. However, improvements or revisions are still needed therefore the product developed becomes better and may meet the instrument criteria that may help students carry out high order thinking and meet the criteria for achieving metacognitive competence.

Field Trial Result

The student test results on the question items of the assessment instrument are done by filling in the user response questionnaire. User response questionnaire data taken from 30 students are provided in Table 5.

Table 5. Quantitative Data on User Response by Students

No.	Component	TSE	(%)	Criteria
1	Language	172.67	86.33	Excellent
2	Presentation	342	85.5	Good
3	Instrument utilization	174.66	87.33	Excellent
	Average Score		86.16	Excellent

Student test results on the question items of the HOTS-oriented and metacognitive-based assessment instrument was declared "excellent" with a percentage value of 86.16 percent.

Question Difficulty Level Result

The difficulty level data is taken from the user subject (students) when working on 50 question items, 30 multiple-choice questions, and 20 essay questions. The difficulty level result of the 1st question on multiple-choice is that there are questions with a low difficulty level, with a total of 10 questions. Besides, there are 17 questions with moderate difficulty. Meanwhile, the analysis result in the essay questions produces questions with two low-level difficulty questions and 17 moderate level difficulty questions. Based on these results, the researcher revised the questions on multiple-choice questions and essay questions in the category of moderate and low-level difficulty questions. This result still shows that the metacognitive-based and HOTS-oriented questions have not been achieved, because, in the results, there is still more than 50 percent of questions in low and moderate difficulty level. This means that the product orientation produced has not yet reached, and the questions' difficulty level still needs to be revised to produce questions that have a higher level of difficulty.

Question Differentiation Index Result

The first question's discrimination index result is taken from the user subject (students) when working on 50 questions given. The questions' discrimination index value that has been obtained is then compared with discrimination index criteria, which results in an analysis where several questions still hold a low discrimination index, which means that overlapping questions exist in the assessment instrument. It is necessary to revise questions that hold low discrimination index in the multiple-choice/objective test and essay/subjective test.

Based on these results, the researcher revised the second question on both multiple-choice/objective tests and essays that were moderate and low difficulty level questions and categorized as feasible and infeasible questions. The result is that there are 41 questions with high-level difficulty, six with moderate-level difficulty, and three with low-level difficulty. It means that the questions made have reached HOTS by not putting aside moderate and low-level questions (3 criteria) in order to meet students' ability characteristics, which are different from one another.

The second analysis of the question discrimination index leads to a high level of difficulty in multiple-choice questions and essays, which means that the questions in the second revision may help students think critically and have a better mastery of metacognitive competencies. From the discrimination index results, the questions have an excellent differentiation index. This means that the similarity/resemblance/overlap of questions may be insufficient, which is indicated by multiple-choice and essay question items of having good or even excellent differentiation index.

The Higher Order Thinking Skills (HOTS)-oriented and metacognitive-based assessment instruments developed based on content validation, construct validation, and user response tests met the eligibility criteria. The teacher's content validation results turned out that the average percentage calculation was 90.90 percent and stated "very feasible". It means that the instruments developed are very good in terms of its content. The construct validation results showed that the average percentage calculation was 94.54 percent that it was "very feasible". This means that the instruments developed are very good in terms of the construct (concept construction). Meanwhile, field trials that aimed to see user responses showed very decent results with a percentage calculation of 86.16 percent. Therefore, it concluded that the HOTS-oriented metacognitive-based assessment instrument could be proven valid in terms of language, systematic clarity, content suitability, and comprehensiveness.

The validation results indicate that the test results of the difficulty level and differentiation index of the questions which have been revised twice and improvised. It has also produced HOTS question items to train students to think at a high level and meet the achievement of metacognitive competence. The instrument in the form of question items produced in the second revision has produced questions that hold a high-level difficulty and decent discrimination index. This means that the instrument in the form of question items produced in General Administrative Subjects of Vocational Schools can have students to think at a high level and be able to train students in having the ability in connecting several different concepts, interpreting, solving problems, arguing and making the right decisions. The instrument question items produced by the researchers have met the metacognitive category. This is consistent with the opinion of Widana (2017), which explains about the metacognitive category.

The instrument in the form of question items produced by the researchers, proven by the high-level difficulty and decent discrimination index results, has been oriented as HOTS. HOTS questions generated in the instrument of General Administrative Subjects of Vocational Schools have three characteristics, namely (1) measuring high-level thinking skills, (2) using contextual problem-based assessments; and (3) using various forms of questions. The way to solve high-level thinking questions provided by researchers is not only by and answer key. It is

beneficial for students to practice solving questions with a high level of difficulty without feeling hopeless. 21st-century students need Metacognitive-based and HOTS-oriented assessment questions because they have the characteristics of critical thinking, problem-solving, creativity, communication skills, and the ability to work. This inherent characteristic of 21st-century students is what made the assessment must also be oriented to students' needs, which means that the assessment must be HOTS and metacognitive-based. This is what has been produced by researchers.

Characteristics of a student who have mastered metacognitive competence are that the students can (1) use evidence proportionally and precisely; (2) manage thoughts and interpret them concisely and clearly; (3) distinguish between giving reasons and rationalizing them; (4) understand ideas or opinions at a certain level of trust; (5) anticipate the risks of an action; (6) learn independently and be confident in doing so; (7) implement problem-solving strategies in completing material; (8) apply the material taught by teachers in daily life; (9) understand and criticize their own opinion as well as the risk of their opinion; (10) know the possibility of error from their opinion and; (10) aware of the fact that someone has limited understanding (Fitriana, 2016; Nickerson, 1988; Sudjana & Wijayanti, 2018).

Students who have metacognitive competencies in themselves will be able to manage and control learning activities independently. If students can independently manage their learning activities, they will be able to create questions that must be solved on their own and can evaluate themselves. The process of finding questions and answers that arise and self-evaluation will improve students' higher-order thinking skills and ultimately affect the learning outcomes (Iskandar, 2014). Students need to have metacognitive competences in their learning activities. This is reinforced by the results of some researches which states that metacognitive competencies have an essential role in managing and controlling the cognitive processes of individuals in learning activities and in having them to think more effectively and efficiently, also to train students to develop learning strategies and determine solutions in solving problems (Bakar and Ismail, 2019; Bran and Balas, 2011; Chen et al., 2019; Iskandar, 2014; Kartika and Sujadi, 2015; Christian, 2015; Künsting et al., 2013; Kyriakides et al., 2020; Pintrich, 2002; Romli, 2012; Sastrawati and Rusdi, 2011; Taub et al., 2020; Vrieling et al., 2012; Zheng et al., 2019)

The result of this study supports a prior study by Wardany and Ramli (2017), which showed that HOTS assessment instrument products in the form of developed multiple choice and essays have validity and reliability with very high average interpretation and good quality questions in terms of the level of difficulty and discrimination index. Indeed, Sari (2018) also revealed that the feasibility of metacognitive-based and HOTS-oriented Student Activity Sheet, which assessed material aspects, linguistic aspects, and visual aspects, met decent criteria even the product was suitable for use, both for different learning or research. Content criteria, presentation criteria, language criteria, and graphic criteria meet a very well category. Research by Netri et al. (2018) also shows that HOTS-oriented test instruments meet valid criteria based on material, construction, and language aspects. The test instrument is also declared reliable within a high-value category, so the product was suitable to be used. In addition, research and development from

Sari et al. (2018) state that expert validation data result meets the valid criteria therefore it is suitable to be used. The ability to have a higher level of thinking in a student after using the developed assessment instruments can be categorized as a good result.

From the content and construct validation and user trials result, the HOTS-oriented and metacognitive-based assessment instrument are feasible and applicable to be implemented. Therefore, this assessment instrument can be used by the General Administrative teacher in knowing students' high-level thinking skills and assessing students' abilities in the metacognitive mastery of the questions provided and can also be used as a guide in making the same instrument for other essential competencies. Also, students can use the question items in the instrument as learning material and practice their high-level understanding and thinking skills.

HOTS-based and metacognitive-based assessment instruments are able to improve students' critical thinking ability. Furthermore, these instruments allow students to enhance their creative thinking. By using these instruments, students are habituated to creatively and critically answer HOTS-based questions based on the recent 21st century skills.

CONCLUSION

The research and development concluded that metacognitive-based and HOTS-oriented assessment instruments could be implemented for appropriate and applicable General Administrative subjects. To produce HOTS-based and metacognitive-based instrument question items, several revisions are needed because the product instrument can train students' critical thinking. Further researches on the development of metacognitive-based and HOTS-oriented assessment instrument need to concern the following points: (1) it is necessary to continue the research until the dissemination, and product implementation stages; (2) metacognitive-based and HOTS-oriented assessment instruments that will be developed can add psychomotor aspects, so it is not limited to cognitive or knowledge aspects; (3) it is necessary to continue to do several question analysis in measuring the difficulty level and differentiation index of the questions therefore they can produce HOTS questions; and (4) the questions in each basic competency must be tested thoroughly, hence it can know which questions are suitable for use or which question must be revised or discarded.

REFERENCES

Akbar, S. (2016). *Instrumen perangkat pembelajaran.* Bandung: PT. Remaja Rosdakarya.

Arifin, Z. (2014). *Evaluasi pembelajaran: prinsip, teknik, dan prosedur*. Bandung: PT. Remaja Rosdakarya.

Arifin, Z., & Retnawati, H. (2017). Pengembangan instrumen pengukur higher order thinking skills matematika siswa SMA Kelas X. *Pytha Jurnal Pend. Math 12*(1), 98. https://Doi.Org/10.21831/Pg.V12i1.14058

Arikunto, S. (2015). *Dasar-Dasar Evaluasi Pendidikan Edisi 2.* Jakarta: PT. Bumi Aksara.

Bakar, M. A. A., & Ismail, N. (2019). Metacognitive learning strategies in mathematics classroom intervention: a review of implementation and operational design aspect. *International Electronic Journal Mathemathics Education, 15*(1), 1-9. https://doi.org/10.29333/Iejme/5937

Borg, W. R., & Gall, R. (2003). *Educational research: an introduction, 7th Ed*. New York: Logman Inc.

Bowler, L. (2010). A taxonomy of adolescent metacognitive knowledge during the information search process. *Library & Information Science Research 32*(1), 27–42. https://doi.Org/10.1016/j.Lisr.2009.09.005

Bran, C.-N., & Balas, E.-C. (2011). Metacognitive regulation and in-depth learning. a study on the students preparing to become teachers. *Procedia - Social and Behavioral Sciences 11*, 107–111. https://doi.org/10.1016/j.sbspro.2011.01.043

Brookhart, S. A. (2010). *How to assess high-order thinking skills in your classroom*. Virginia: Ascd.

Budiman, A., & Jailani, J. (2014). Pengembangan instrumen asesmen higher order thinking skill (HOTs) pada mata pelajaran matematika SMP kelas VIII semester 1. *Jurnal Riset Pendidikan Matematiika 1*(2) 139-151. https://doi.org/10.21831/jrpm.V1i2.2671

Chen, J. H., Björkman, A., Zou, J. H., & Engström, M. (2019). Self–regulated learning ability, metacognitive ability, and general self-efficacy in a sample of nursing students: a cross-sectional and correlational study. *Nurse Education in Practice 37,* 15–21. https://doi.org/10.1016/j.nepr.2019.04.014

Chevron, M. P. (2014). A metacognitive tool: theoretical and operational analysis of skills exercised in structured concept maps. *Perspectives in Science 2*, 46–54. Https://doi.org/10.1016/j.pisc.2014.07.001

Daryanto. (2012). *Evaluasi pendidikan.* Jakarta: PT. Rineka Cipta.

Enger, K.S., & Yager, E.R. (2009). *Assesing student understanding in science, 2nd ed.* California: Corwin A Sage Company.

Fitriana, M. (2016). Penggunaan strategi pembelajaran inkuiri untuk meningkatkan metakognisi siswa SMA 10. *Jurnal Inovasi Pendidikan Kimia 10*(1), 1702-1711.

Higgins, S. (2017). Managing higher-order thinking skills, in: managing academic libraries. Elsevier. https://doi.org/10.1016/b978-1-84334-621-0.00004-2

Iskandar, S. M. (2014). Pendekatan keterampilan metakognitif dalam pembelajaran sains di kelas. *Erudio Journal of Educational Innovation, 2*(2), 13–20. https://doi.org/10.18551/erudio.2-2.3

Jarvis, M. A., & Baloyi, O. B. (2020). Scaffolding in reflective journaling: a means to develop higher order thinking skills in undergraduate learners. *International Journal of Africa Nursing Sciences 12*, 100195. https://doi.org/10.1016/j.ijans.2020.100195

Kartika, D. L., & Sujadi, I. (2015). Proses metakognisi dalam pemecahan masalah matematika pada siswa kelas XI di SMA Negeri Banyumas. *Jurnal Pembelajaran Matematika, 3*(9), 1021-1034.

Künsting, J., Kempf, J., & Wirth, J. (2013). Enhancing scientific discovery learning through metacognitive support. *Contemporary Educational Psychology, 38*(4)*,* 349–360. https://doi.org/10.1016/j.cedpsych.2013.07.001

Kurniati, D., Harimukti, R., & Jamil, N. A. (2016). Kemampuan berpikir tingkat tinggi siswa smp di kabupaten jember dalam menyelesaikan soal berstandar pisa. *Jurnal Penelitian dan Evaluasi Pendidikan, 20*(2), 142-155. https://doi.org/10.21831/pep.v20i2.8058

Kusaeri. (2014). *Acuan dan teknik penilaian proses dan hasil belajar dalam kurikulum 2013.* Yogyakarta: Ar-Ruzz Media.

Kyriakides, L., Anthimou, M., & Panayiotou, A. (2020). Searching for the impact of teacher behavior on promoting students' cognitive and metacognitive skills. *Studies in Educational Evaluation, 64,* 100810. https://doi.org/10.1016/j.stueduc.2019.100810

Lee, S. M. (2014). The relationships between higher order thinking skills, cognitive density, and social presence in online learning. *The Internet and Higher Education 21*, 41–52. https://doi.org/10.1016/j.iheduc.2013.12.002

Lestari, C. F., Kristiana, A. I., & Kurniati, D. (2016). Pengembangan paket tes matematika berbasis kemampuan berpikir tingkat tinggi siswa kelas X TKJ SMK materi sistem persamaan linier. *Jukasi 3*, 34. https://doi.org/10.19184/jukasi.v3i2.3527

Masari, G. A., & Anghel, O. (2012). Comparative study on developing metacognitive abilities of students from technical, vocational and human sciences. *Procedia - Social and Behavioral Sciences 46*, 4418–4422. https://doi.org/10.1016/j.sbspro.2012.06.266

Narmaditya, B. S, Wulandari, D., & Sakarji, S. R. B. (2018). Does problem-based learning improve critical thinking skill?. *Cakrawala Pendidikan, 37*(3), 378-388

Netri, N., Holiwarni, B., & Abdullah. (2018). Development of test instruments based higher order thinking skill (HOTs) on chemical equilibrium at second grade in senior high school. *JOM FKIP, 5*(2), 1-11.

Nguyen, T. M. T., & Nguyen, T. T. L. (2017). Influence of explicit higher-order thinking skills instruction on students' learning of linguistics. *Thinking Skills and Creativity, 26,* 113–127. https://doi.org/10.1016/j.tsc.2017.10.004

Nickerson, R. S. (1988). Improving thinking through instruction. *Review of Research in Education 15*(1), 3–57. https://doi.org/10.3102/0091732x015001003

Nofiana, M. (2016). Pengembangan instrumen evaluasi higher order thinking skills pada materi kingdom plantae. *Pedagogi Hayati, 1*(1), 46-53.

Pilten, G. (2010). Evaluation of the skills of 5th grade primary school students' high-order thinking levels in reading. Procedia - *Social and Behavioral Sciences, 2*, 1326–1331. https://doi.org/10.1016/j.sbspro.2010.03.195

Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory Into Practice, 41*(4), 219–225.

Pratiwi, U., & Fasha, E. F. (2015). Pengembangan instrumen penilaian hots berbasis kurikulum 2013 terhadap sikap disiplin. *Jurnal Penelitian Pembelajaran IPA 1*(1), 123-142. https://doi.org/10.30870/jppi.v1i1.330

Roets, L., & Maritz, J. (2017). Facilitating the development of higher-order thinking skills (hots) of novice nursing postgraduates in africa. *Nurse Education Today 49*, 51–56. https://doi.org/10.1016/j.nedt.2016.11.005

Romli, M. (2012). Strategi membangun metakognisi siswa sma dalam pemecahan masalah matematika. *Aksioma: Jurnal Matematika dan Pendidikan Matematika, 17,* 1-16.

Saputra, H. (2016). P*engembangan mutu pendidikan menuju era global: penguatan mutu pembelajaran dengan penerapan HOTS (higher order thinking skills).* Bandung: Smile's Publishing.

Sari, A. T. I. (2018). Pengembangan lembar kegiatan peserta didik (LKPD) berbasis higher—order thinking skills (HOTS) pada mata pelajaran akuntansi perbankan syariah kelas XI semester II. *Jurnal Pendidikan Akuntansi, 6*(3), 230-238.

Sari, D. R. U., Wahyuni, S., & Bachtiar, R. W. (2018). Pengembangan instrumen tes multiple choice high order thinking pada pembelajaran fisika berbasis e-learning di sma. *Jurnal Pembelajaran Fisika, 7*(1), 100-107. https://doi.org/10.19184/jpf.v7i1.7231

Sastrawati, E., & Rusdi, M. (2011). Problem-based learning, strategi metakognisi, dan keterampilan berpikir tingkat tinggi siswa. *Teknopedagogi*, 1, 14.

Sudjana, D., & Wijayanti, I. E. (2018). Analisis keterampilan metakognitif pada materi kelarutan dan hasil kali kelarutan melalui model pembelajaran pemecahan masalah. *Educhemia, 3*(2), 206-221. https://doi.org/10.30870/educhemia.v3i2.3729

Surasa, N. N., Witjaksono, M., & Utomo, S. H. (2017). Proses belajar siswa dalam meningkatkan kemampuan berpikir kritis mata pelajaran ekonomi SMA. *Jurnal Pendidikan: Teori, Penelitian, dan Pengembangan, 2*(1), 78-84

Tal, R., & Hochberg, N. (2003). Assessing high order thinking of students participating in the "wise" project in Israel. *Studies in Educational Evaluation, 29*(2), 69–89. https://doi.org/10.1016/s0191-491x(03)00016-6

Taub, M., Sawyer, R., Smith, A., Rowe, J., Azevedo, R., & Lester, J. (2020). The agency effect: the impact of student agency on learning, emotions, and problem-solving behaviors in a game-based learning environment. *Computers & Education, 147*, 103781. https://doi.org/10.1016/j.compedu.2019.103781

Uno, H. B., & Koni, S. (2012). *Assessment pembelajaran*. Jakarta: Bumi Aksara.

Vrieling, E. M., Bastiaens, T. J., & Stijnen, S. (2012). Effects of increased self-regulated learning opportunities on student teachers' metacognitive and motivational development. *International Journal of Educational Research, 53*, 251–263. https://doi.org/10.1016/j.ijer.2012.03.014

Wardany, K., & Ramli, M. (2017). Pengembangan penilaian untuk mengukur higher order thinking skills siswa. *Inkuiri: Jurnal Pendidikan IPA, 6*(2), 1-12.

Widana, I. W. (2017). *Modul penyusunan soal hots tahun 2017*. Jakarta: Departemen Pendidikan Dan Kebudayaan.

Wing, C. K., Piaw, C.Y., & Chang, P. K. (2014). Effects of aural-imitative and aural-motivic analyses on higher-order thinking skill and creative musical product in music improvisation. *Procedia - Social and Behavioral Sciences, 116,* 5130–5134. https://doi.org/10.1016/j.sbspro.2014.01.1086

Wu, L. Y. (2016). Children's Formation and representations of money-related thinking in graphical complexes: compound relation operations and creative high-order thinking. Thinking *Skills and Creativity, 19,* 232–245. https://doi.org/10.1016/j.tsc.2015.12.001

Yang, Y. T. C. (2015). Virtual ceos: a blended approach to digital gaming for enhancing higher order thinking and academic achievement among vocational high school students. *Computers & Education, 81*, 281–295. https://doi.org/10.1016/j.compedu.2014.10.004

Yee, M. H., Yunos, J. Md., Othman, W., Hassan, R., Tee, T. K., & Mohamad, M. M. (2015). Disparity of learning styles and higher order thinking skills among technical students. *Procedia - Social and Behavioral Sciences, 204*, 143–152. https://doi.org/10.1016/j.sbspro.2015.08.127

Yerizon, Y. (2019). Pengembangan lembaran kerja matematika smp berbasis pendekatan metakognisi untuk meningkatkan higher order thinking skill peserta didik. *Jurnal Pendidikan Matematika dan Sains Matematika, 4,* 143–153. https://doi.org/10.31629/jg.v4i2.1418

Zheng, L., Li, X., Zhang, X., & Sun, W. (2019). The effects of group metacognitive scaffolding on group metacognitive behaviors, group performance, and cognitive load in computer-supported collaborative learning. *The Internet and Higher Education, 42,* 13–24. https://doi.org/10.1016/j.iheduc.2019.03.002

Zumbach, J., Rammerstorfer, L., & Deibl, I. (2020). Cognitive and metacognitive support in learning with a serious game about demographic change. *Computers in Human Behavior, 103*, 120–129. https://doi.org/10.1016/j.chb.2019.09.026