

# Critical sociology in the development of HOTS-oriented cognitive assessment instruments

Aris Martiana

*Department Sociology Education, Faculty of Social Science, Universitas Negeri Yogyakarta, Indonesia*  
Email: aris\_martiana@uny.ac.id

Edi Istiyono

*Department Educational Research and Evaluation, Universitas Negeri Yogyakarta, Indonesia*  
Email: edi\_istiyono@uny.ac.id

Widihastuti Widihastuti

*Department Educational Research and Evaluation, Universitas Negeri Yogyakarta, Indonesia*  
Email: widihastuti@uny.ac.id

## Abstract

The current curriculum focuses on three domains, one of which is cognitive. Cognitive abilities are very important for the development of individual competencies so that instruments to measure them must be carried out. High quality cognitive instruments are needed in the assessment. Assessment with High Order Thinking Skills (HOTS) orientation as an effort to implement 6c, namely communication, collaboration, critical thinking, creative thinking, also computational logic and concern by developing cognitive instruments in Sociology scientific assessment. The method used is Research and Development which is focused on development. Cognitive instruments were compiled and carried out content validity from 4 experts for Sociology, Education and Law which were then tested on 93 students through a questionnaire in the form of google form. Analysis of the data used is by using software iteman 4.3. The results of this study are cognitive instruments with 15 HOTS questions in the form of multiple choice using content validity with Aiken's count obtained a score of 0.75 to 0.9375 so it is included in the medium and high categories. The instrument has an Alpha score of 0.558 so it is included in the reliable category. For items that have a difficulty level of 0.637, it means that the difficulty of the questions is moderate and the positive discriminatory power with a number of 0.243 is in the medium category. Then the distractor on the answer choices in the question items can function well because most of them are  $> 0.05$ .

**Keywords:** *Development, Cognitive Instruments, HOTS, Sociology*

## Introduction

The independent curriculum that has been applied to higher education institutions currently provides opportunities for students to gain learning experiences outside their study programs and are oriented to acquire 21st century skills needed in the Industry 4.0 era, including communication, collaboration, critical thinking, creative thinking, as well as computational logic. and concern (Kemendikbud, 2020) skills in the 21st century are known as 6C. The curriculum contains many things such as planning, process and evaluation. The curriculum contains content and subject matter. The curriculum is a number of subjects that must be taken and studied by students to obtain subjects that must be taken and studied by students to gain some knowledge. Curriculum as a learning plan. Curriculum is an educational program that is provided to teach students. Based on this program, students carry out various learning activities, so that changes and developments in student behavior occur, in accordance with the objectives of education and learning. Curriculum is a learning experience. The curriculum

is a series of learning experiences (Oemar Hamalik, 2017). Evaluation is one component in the curriculum so that it has an important position. In developing curriculum evaluation is one of the important components and stages that must be taken by teachers to determine the effectiveness of the curriculum. The results obtained can be used as feedback for teachers in improving and perfecting the curriculum. (Zainal Arifin, 2017). When evaluation is carried out to make a decision that includes all components, then in the curriculum there are also assessment activities whose scope is more specific than evaluation even though they have similarities in making decisions.

Learning assessment standards are the minimum criteria for assessing student learning processes and outcomes in order to fulfill graduate learning outcomes. The assessment principles include educative, authentic, objective, accountable, and transparent principles which are carried out in an integrated manner. The assessment technique consists of observation, participation, performance, written test, oral test, and questionnaire (Permendikbud RI No. 2 of 2020). Assessment is one or more processes of identifying, collecting, and preparing data to evaluate the achievement of graduate learning outcomes (CPL), and curriculum objectives (ABET, 2016). Compulsory assessments contain motivational content, foster self-confidence to contribute to life choices as lifelong learners. Then use special skills to work in the superteam of his choice (Kemendikbud, 2020). Assessment is used to see the ability of students in mastering the material/concepts in learning. Assessment is an important component in the implementation of education. Efforts to improve the quality of education can be pursued through improving the quality of learning and the quality of the assessment system. Both are interrelated. A good learning system will produce good quality learning. The quality of this learning can be seen from the results of the assessment. Furthermore, a good assessment system will encourage educators to determine good teaching strategies and motivate students to learn better. (Mardapi, 2008). There are many types of assessment that can be done, one of which is a form of test to assess the cognitive abilities of students. Various cognitive assessment techniques can be used according to the competencies to be achieved. The techniques commonly used are written tests, oral tests and assignments. However, it is possible to use other appropriate techniques such as portfolio and observation. Assessment of written test knowledge such as true-false, multiple choice, matchmaking, filling/complete, description; oral test: quiz and question and answer; assignments such as assignments carried out individually or in groups in the education unit and/or outside of school; other techniques such as observation (Istiyono, 2020).

The development of HOTS oriented cognitive instruments is seen from the aspects of reliability, validity, item difficulty level, differentiating power of questions and distractors. As in similar research, namely Student comprehensive test was constructed to evaluate the cognitive aspect of the students in the beginner string ensemble. The item analysis conducted in this study implicates three statistics to assists in analyzing the effectiveness of each of the comprehensive test questions specifically item difficulty, item discrimination and distractor analysis (Shafizan Sabri, 2013). Six important factors influencing classroom assessment include 21st-century knowledge, skills, and dispositions; technology; cognitive and sociocultural learning and motivation theory; standards-based education; high-stakes testing; and data literacy (McMillan, 2018). Assessment instruments or HOTS questions are questions that require higher-order thinking skills. In shaping the quality of students to be better, these kinds of questions must be developed by the teacher properly and applied in the class they are teaching. HOTS or higher order thinking skills are divided into four groups, namely problem solving, decision making, critical thinking, and creative thinking (Nitko & Brookhart.2011; Pratiwi.et.al.2017). Today, Higher Order Thinking Skills (HOTS) is needed by students. Students are not only required to apply what they have learned, but also to analyze, evaluate, and synthesize the knowledge they have gained to solve problems in everyday life. Thus, HOTS needs to be trained... (G.S. Pratama & H. Retnawati, 2018). It is important to give HOTS assessments to students in Sociology, especially in the Social Deviation, Crime and Law Enforcement courses with very good critical sociological analysis with the habit of students having the HOTS way of thinking in studying and implementing the science. Here Sociology is a science that studies the networks between humans in life so that being critical of the environment, both in terms of thought and behavior, is the goal of learning in this course. Therefore the development of cognitive instruments is important in order to improve cultural education.

## Methods

This study uses the development method, namely the research method used to produce certain products, and test the effectiveness of these products (Sugiyono, 2013). In this case, the product developed is a cognitive ability instrument oriented to HOTS (High Order Thinking Skills) which consists of a grid of questions, multiple choice questions and scoring guidelines. The development model used is the 4D Model which has stages, namely Define, Design, Develop and Disseminate. (Thiagarajan, Dorothy S. Semmel, and Melvyn I. Semmel, 1974). The 4D model is carried out to the 3D stage. This research refers to the development model as follows :

1. Define :
  - a. identify and analyze basic competencies
  - b. identify and analyze the characteristics of students
  - c. analyze student competence
  - d. analyzing material concepts
  - e. formulate learning outcomes
2. Design :
  - a. Preparation of the question grid
  - b. Preparation of cognitive instrument questions
  - c. Preparation of scoring guidelines
  - d. Preparation of validation instruments
3. Develop :
  - a. Assessment of cognitive instruments by experts
  - b. The trial is limited to respondents

Respondents in this study were students of Sociology Education FIS UNY who took courses in Social Deviation, Crime and Law Enforcement. Because this course is an elective course, it consists of 2nd, 4th and 6th semester students who can take it. For data collection is done through questionnaires and observations. Observations are made through observations in the classroom when teaching. Observations of students in each lesson with a total of 14 meetings. The analysis technique used is iteman software to help analyze cognitive instruments.

## RESULTS AND DISCUSSION

### Results

Based on the research that has been done using 3D development, the following are the results in developing a HOTS-oriented cognitive instrument :

#### a. Define

This stage is the initial defining stage related to the results of the analysis of the sociology of science in social deviation, crime and law enforcement courses. Information was obtained that the subject has never held a cognitive instrument in the form of multiple choice oriented to HOTS (High Order Thinking Skills). Assessment is usually done in the form of an essay so it is necessary to innovate learning in these subjects. The concept of the material in the course is analyzed so that multiple choice assessments can be used. The next analysis also includes the competencies that must be mastered by students which are adjusted to the Semester Learning Plan (RPS) by looking at the course descriptions and the learning outcomes of the courses.

Based on the initial analysis, a HOTS oriented cognitive instrument was developed in the Deviation, Crime and Law Enforcement elective courses. An important assessment is carried out in order to determine

the competence of students in mastering the concepts and application of their knowledge in life and even the HOTS assessment is needed to be able to stimulate students to think more holistically and comprehensively.

## b. Design

At the design stage, it is the stage of preparing the instrument properly so that it can be applied to students so that it is in accordance with the objectives. Feasibility in designing this instrument is necessary so that several things are done, among others: first is in the preparation of a grid of questions that must be seen from the Semester Learning Plan (RPS) so that the material to be tested is in accordance with what is taught to students. The preparation of the question grid is also adjusted to the objectives of this study and also the considerations for the allocation of working time and the level of difficulty in the material which is of course HOTS-oriented. So that it can be said that this is an adjustment to learning needs that must be done. After compiling a grid of questions, it is done by compiling a HOTS-oriented cognitive instrument. Questions are arranged by adjusting the grid and creating a stimulus from data, cases and images that are in accordance with the questions to be given. This suitability is important so that students have an understanding of what will be tested. Likewise in compiling answer choices that must have original answers as well as four answer choices by having good distractors.

Scoring guidelines are also needed in multiple-choice cognitive assessments so that they can conduct assessments objectively and fairly for students. Then an instrument for expert validators was also compiled with 14 criteria, both scientifically, grammatically and linguistically used in the instrument. The assessment instrument for the validator uses a Likert scale of 1 to 5 with the criteria of very good, good, good enough, not good and not good which must be given to each criterion and suggestions are given regarding validated cognitive instruments. The validator must also provide conclusions and recommendations for the cognitive instrument to be tested.

## c. Develop

This develop stage is the development stage of the instrument to be tested. After all the assessment tools from the grid, questions, scoring guidelines, and validation sheets are ready, they are given to the validator to validate the contents. There are 4 validators selected with expertise in education, sociology and law. The following table calculates the results of Aikens :

Table 1. Aikens Results

Butir ke	Rater-1	Rater 2	Rater 3	Rater 4	si	N(c-1)	V	Keterangan
1	3	5	4	5	13	16	0.8125	Tinggi
2	5	5	4	5	15	16	0.9375	Tinggi
3	5	5	4	4	14	16	0.875	Tinggi
4	4	4	4	4	12	16	0.75	Sedang
5	4	4	3	5	12	16	0.75	Sedang
6	5	5	4	4	14	16	0.875	Tinggi
7	3	5	4	4	12	16	0.75	Sedang
8	5	5	4	4	14	16	0.875	Tinggi
9	5	5	4	4	14	16	0.875	Tinggi
10	5	5	3	4	13	16	0.8125	Tinggi
11	5	5	4	4	14	16	0.875	Tinggi
12	5	5	4	4	14	16	0.875	Tinggi
13	5	5	4	5	15	16	0.9375	Tinggi
14	5	5	4	5	15	16	0.9375	Tinggi

Based on the aiken calculation results table, it can be explained that the results of 4 expert raters/validators were processed into the Aikens test with the results ranging from 0.75 to 0.9375 in the medium and

high categories. After that the instrument was revised according to what was noted by the rater/validator. The conclusions and recommendations of the rater/validator are appropriate for testing after revision. So there are several inputs related to the suitability of the grid, RPS and the nature of HOTS that must be raised in the question. Then a limited trial was conducted on 93 students as respondents. Students as students are students who take courses in Social Deviation, Crime and Law Enforcement. The trial was assisted by using google forms. The arrangement on the Google form is made well so that the substance that is the goal is conveyed and also makes the design on the google form to be attractive to respondents.

Cognitive instruments that have been answered by respondents through google form are then made on a coding table so that analysis can be carried out with the help of iteman software 4.3. The results of this item analysis can be seen, among others:

1. Reliabilitas

**Tabel 2. Reliabilitas**

Score	Alpha	SEM	Split-Half (Random)	Split-Half (First-Last)	Split-Half (Odd-Even)	S-B Random	S-B First-Last	S-B Odd-Even
Scored items	0.558	1.384	0.356	0.396	0.413	0.525	0.567	0.584

The reliability table is shown at an alpha of 0.558 meaning that the cognitive instrument that has been compiled is reliable with that amount so that it can be categorized as moderate reliability according to the alpha criteria between 0.50 - 0.70 so it can be said that this instrument is reliable. With iteman analysis, a test of the level of difficulty of the items, the distinguishing power and distractor of the multiple choice questions is carried out.

2. Difficulty Level, Discrimination index and Distractor

**Tabel 3. Summary Statistik**

Score	Items	Mean	SD	Min Score	Max Score	Mean P	Mean Rpbis
Scored Items	15	9.559	2.082	2	13	0.637	0.243
Scaled Total	15	0.000	0.000	0.000	0.000	-	-

Based on the statistical summary table, it can be explained that the level of difficulty of the items has a magnitude (Mean P) of 0.637, which means that the questions are included in the medium category because they are in the range of 0.3 - 0.7. The questions are between easy questions and difficult questions so that students can still work on questions with good results. Then at the level of discriminating power is to look at the average Rpbis, which is a positive discriminatory power with a number of 0.243, meaning that the questions included in the category need to be revised on average from the data because they are less than 0.3. The analysis of each item item can be seen in table 4. The following:

**Tabel 4. Results of the Analysis of Cognitive Questions**

No. item	P	Rpbis	Rbis	Pilihan (Prop endorsing)				
				A	B	C	D	E
1	0.484	-0.036	-0.045	0.344	0.484	0.011	0	0.161
2	0.882	0.447	0.729	0.011	0.011	0.043	0.882	0.054
3	0.032	-0.076	-0.186	0	0.022	0.032	0.032	0.914
4	0.29	0.045	0.059	0.108	0.032	0.462	0.108	0.29
5	0.806	0.323	0.465	0.806	0	0.043	0.14	0.011

6	0.935	0.445	0.868	0	0.022	0.935	0.043	0
7	0.559	0.38	0.479	0.151	0.161	0.054	0.559	0.075
8	0.946	0.34	0.701	0.032	0.946	0	0.011	0.011
9	0.505	0.06	0.075	0.505	0.043	0.065	0.344	0.043
10	0.914	0.256	0.457	0.914	0	0.075	0.011	0
11	0.839	0.537	0.807	0.075	0.839	0.032	0.032	0.022
12	0.151	-0.011	-0.016	0.409	0.086	0.151	0.344	0.011
13	0.957	0.5	1	0	0.043	0.957	0	0
14	0.376	-0.026	-0.034	0.376	0.118	0.376	0.118	0.011
15	0.882	0.466	0.759	0.011	0.097	0.882	0.011	0

Based on the table, it can be explained that the level of difficulty of the items for each number can be grouped as follows: 1). For the moderate level of difficulty, namely the range from 0.3 to 0.7, the numbers are 1, 7, 9, and 14; 2). for difficult questions, the range from 0.1 to 0.29 is numbers 4 and 12; 3). For easy questions, the range from 0.7 to 0.90 is numbers 2, 5, 11 and 15; 4). For the very difficult question number 3; 4). While very easy questions > 0.90 are numbers 6, 8, 10, and 13.

The level of differentiating power of each item is for magnitude > 0.3 high category numbers 2, 5, 6, 7, 8, 11, 13 and 15. For the medium category, it ranges between 0.1 - 0, 29 is number 10 while for magnitude < 0.1, including low power of difference, namely at numbers 1, 3, 4, 9, 12, 14. The distractors on each item on average have worked well even though there must be a change in answer choices. The category works well if the P value has a magnitude > 0.05

Analysis of cognitive questions that have been prepared based on item analysis by looking at the level of difficulty, differentiating power and distractors that exist and the percentage of respondents answering can be categorized:

1. Accepted questions

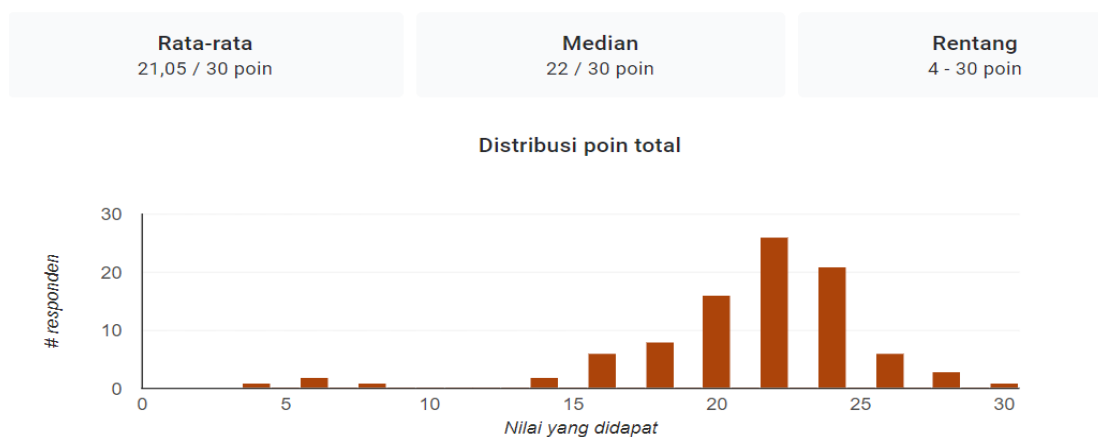
The questions that were well received on the instrument were 8 questions, namely for questions number 2, 5, 6, 7, 8, 11, 13, and 15.

2. Accepted questions with revisions

Questions received with a revision of 1 item, namely question number 10

3. Questions that are not accepted or discarded

Questions that are not accepted or discarded a number of 6 are numbers 1, 3, 4, 9, 12, and 14



Gambar 1. Diagram distribusi poin total responden



Based on the picture above, it can be explained that the average point value is 21.05 and the median is 22 points out of the total points is 30. Respondents who are being tested are 93 and it can be interpreted that most of the students are able to answer cognitive questions in the Social Deviation course, Crime and Law Enforcement.

## Discussion

Educational goals can be classified into three categories, namely thinking skills, behavior (behavior) and physical skills which represent three domains (domains) of education, namely cognitive (cognitive), affective (affective) and psychomotor (psychomotor) (Endrayanto & Harumurti, 2014). This research focuses on the cognitive domain in the aspect of students' knowledge. So that the cognitive assessment in the HOTS oriented Social Deviation, Crime and Law Enforcement course as a Critical Sociology study has the aim of measuring student mastery in the aspect of knowledge. The knowledge/cognitive aspect is in accordance with Anderson's taxonomy, namely remembering, understanding, applying, analyzing, evaluating and creating. The categories of the cognitive process dimension are intended to provide a comprehensive set of classifications for those student cognitive processes that are included in objectives... the categories range from the cognitive processes most commonly found in objectives, those associated with Remember, through Understand and Apply, to those less frequently found, Analyze, Evaluate, and Create. Remember means to retrieve relevant knowledge from long-term memory. Understand is defined as constructing the meaning of instructional messages, including oral, written, and graphic communication. Apply means carrying out or using a procedure in a given situation. Analyze is breaking material into its constituent parts and determining how the parts are related to one another as well as to an overall structure or purpose. Evaluate means making judgments based on criteria and/or standards. Finally, Create is putting elements together to form a novel, coherent whole or to make an original product (Anderson, et.al. 2001). Based on the revised Bloom's Taxonomy that the indicators for measuring HOTS include analyzing (C4), the ability to separate concepts into several components and connect with each other to gain an understanding of the concept as a whole; evaluate (C5) an ability to determine the degree of something based on certain norms, criteria or standards; and create (C6), the ability to fit elements into something new, complete and broad, or make something original (I R N Afifah & H Retnawati, 2019). The importance of HOTS learning for students so that the application of cognitive assessment in Sociology is needed. Learning with the aim of students being able to think at a high level which includes analysis, evaluation and creation because this is needed in the current era. Instruments on cognitive abilities that have been developed are in accordance with the requirements in the steps of compiling high-quality assessment instruments as seen from the validity, reliability, level of difficulty of the items, differentiating power and distractors. McMillan (2018) describes high-quality instruments, including: Clear Learning Targets, Alignment of assessment methods and learning targets, Validity, Reliability/precision, Fairness, Positive Consequences, Alignment, Effective and efficient. Based on the results of the analysis that has been carried out at the development stage, the instrument validity is obtained from experts with medium and high criteria. Validity is important to see the validity of the instruments that have been prepared. Validity is the most important consideration in evaluating the quality of the test as a measuring instrument. The concept of validity refers to the feasibility, significance and usefulness of certain inferences that can be made based on the relevant test scores (Azwar, 2016). This study uses content validity with 4 expert raters with scientific backgrounds relevant to the lecture material, the use of expert raters for Sociology, Education and Law. Haynes, et al. said that the meaning of content validity is the extent to which the elements in a measuring instrument are truly relevant and represent a construct that is in accordance with the measurement objectives (Haynes, Richard & Kubany, 1995; Azwar, 2016). Likewise, the reliability of this instrument development research can be said to be quite in accordance with the results, namely 0.558 by converting it to the reliability coefficient (Gulford, 1956; Istiyono, 2020).

**Table 5. Reliability coefficient**

Reliability Coefficient	Reliability level
0,80 - 1,00	Very high
0,60 - 0,80	Tall
0,40 - 0,60	Enough
0,20 - 0,40	Low
0,00 - 0,20	Very low

The validity and reliability that have been fulfilled in the cognitive instrument indicate that the instrument can be used in the respondent's test. In terms of analyzing the results of the instrument, it can be seen by the level of difficulty of the items, differentiating power and distractors. These three things complement the development of the instrument with the stages of the development research model on Development. The difficulty of the items is needed to see how difficult the questions that have been prepared by looking at the degree of difficulty are, namely easy, medium and difficult. The number of students who can answer questions properly and correctly on one number can be seen in the difficulty of the question. The level of difficulty of the questions tested is in the medium category. The ideal question is a question that is in accordance with the ability of the test taker. Thus, a question that is too difficult is not a good question because it can only be done by a few participants, especially the upper group or no one is even able to solve it. Likewise, questions that are too easy can be done correctly by all participants so that they are unable to distinguish between those who are smart and those who are not. In conclusion, good questions are those at a moderate level (Zaini Rohmad, 2016). Regarding the differentiating power of the questions in the cognitive instrument developed, it is included in the medium category according to the discrimination index criteria (Ebel & Fresbie, 1991; Endrayanto & Harumurti, 2014) as follows :

**Table 6. Discrimination Index Criteria**

Discrimination Index	Criteria
0,40 atau lebih	Very good
0,30 - 0,39	Good enough but the items can be improved
0,20 - 0,29	Moderate but the problem can be fixed
Di bawah 0,19	Bad items can be replaced or repaired

The different power level in the instrument must be improved because it is in the range of 0.20 - 0.29. In this case the level of distinguishing power is the ability of the items to distinguish students who have high learning achievement or the upper group and students whose learning achievement is low or the lower group. A positive discrimination index means that many students in the upper group are able to answer correctly on an item than students in the lower group (Endiyanto & Harumurti, 2014). Distractors function well on instruments that have been developed so that they are able to distinguish high and low capable students. Distractor analysis is classified as an incorrect answer in multiple choice questions. Based on Instructional Assessment Resources (IAR), student achievement in an exam item is very much influenced by the quality of the distractor given. Therefore, it is necessary to determine the effectiveness of each item given a distractor as an addition to the analysis of the difficulty of the item. The analysis was performed only for the items in part A of the comprehensive test because the distractor analysis was only related to the multiple choice format (Shafizan Sabri, 2013). In every objective test, alternative answers are always used which contain 2 elements at once, namely the right answer and the wrong answer as a distractor. The purpose of using this distractor is to deceive those who are less able or do not know to be distinguished from those who can. Therefore, a good distractor is one that smart children can avoid and is chosen by less intelligent children (Chabib Toha, 1994; Laela Umi Fatimah & Kahiruddin Alfath, 2019)

## Conclusion



HOTS-oriented cognitive instrument in the form of multiple choice in Sociology, namely the Social Deviation, Crime and Law Enforcement courses using a 4D development model namely Define, Design, Development and Disseminate which has been implemented to the 3D stage. The instrument has been tested on 93 students with the results of the analysis that the questions have met reliability. Regarding the validity, it was carried out on expert validity with 4 raters and the Aikens count was carried out with moderate and high validity results. The level of difficulty of the questions that are owned as a whole is included at the medium level, for the different power is included in the moderate category which requires revision. The distractors in the question have functioned well, although there are still questions that must be revised from the 15 questions that have been prepared.

## References

- Afifah, I. R. N., & Retnawati, H. (2019, October). Is it difficult to teach higher order thinking skills?. In *Journal of Physics: Conference Series* (Vol. 1320, No. 1, p. 012098). IOP Publishing.
- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman,.
- Arifin, Z. (2017). *Konsep dan model pengembangan kurikulum: konsep, teori, prinsip, prosedur, komponen, pendekatan, model, evaluasi dan inovasi*. Remaja Rosdakarya.
- Azwar, S. (2016). *Reliabilitas dan validitas*. Pustaka Pelajar.
- Endrayanto, H. Y. S., & Harumurti, Y. W. (2014). *Penilaian belajar siswa di sekolah*. Kanisius.
- Fatimah, L. U., & Alfath, K. (2019). Analisis kesukaran soal, daya pembeda dan fungsi distraktor. *AL-MANAR: Jurnal Komunikasi dan Pendidikan Islam*, 8(2), 37-64.
- Hamalik, O. (2017). *Kurikulum dan pembelajaran*. Bumi Aksara.
- Istiyono, E. (2020). *Pengembangan instrumen penilaian dan analisis hasil belajar fisika dengan teori tes klasik dan modern*. UNY Press.
- Kemendikbud. (2020). *Buku panduan penyusunan kurikulum pendidikan tinggi di era industri 4.0 untuk mendukung Merdeka Belajar-Kampus Merdeka*
- Mardapi, D. (2008). *Teknik penyusunan instrumen tes dan nontes*. Mitra Cendekia Press.
- McMillan, J. H. (2017). *Classroom assessment: Principles and practice that enhance student learning and motivation*. Pearson.
- Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia No. 3 Tahun 2020 tentang Standar Nasional Pendidikan Tinggi.
- Pratama, G. S., & Retnawati, H. (2018, September). Urgency of higher order thinking skills (HOTS) content analysis in mathematics textbook. In *Journal of Physics: Conference Series* (Vol. 1097, No. 1, p. 012147). IOP Publishing.
- Pratiwi, P. H., Hidayah, N., & Martiana, A. (2017). Pengembangan modul mata kuliah penilaian pembelajaran sosiologi berorientasi HOTS. *Cakrawala Pendidikan*, (2), 85339.
- Rohmad, Z. (2016). *Asesmen & evaluasi pembelajaran*. UNS Press.
- Sabri, S. (2013). Item analysis of student comprehensive test for research in teaching beginner string ensemble using model based teaching among music students in public universities. *International Journal of Education and Research*, 1(12), 1-14.
- Sugiyono. (2013). *Metode penelitian kuantitatif, kualitatif dan R&D*. Alfabeta.

