Assessment of Higher Order Thinking Skills in Genetic Learning: A Validity, Practicality, and Reliability

D I Marlina, Y L Rahmi*

Biology Department, Faculty of Mathematics and Science, Universitas Negeri Padang, Jl. Prof. Dr. Hamka, Air Tawar, Padang, West Sumatra, Indonesia, 25131

*yosibio@fmipa.unp.ac.id

Abstract. The ability of higher order thinking skills students in West Sumatra is still low relatively with an average value in the range of 17.88-38.45. This is evidenced by the existence of trial tests at schools in several areas in West Sumatra. One way to improve the level of thinking ability of students is to develop assessment of higher order thinking skills. This research aims to develop instruments of higher order thinking skills about genetic material that is valid, practical and reliable. This study uses a 4-D models development. The subjects of this study were two biology lecturers from Faculty of Mathematics and Science, Universitas Negeri Padang, two biology teachers and 30 students of class XII SMAN 1 Nan Sabaris. Data were collected by using questionnaire for validity and practicality furthermore for getting value of reliability by using test. Based on data analysis, the value of validity, practicality, and reliability obtained an average of 87.72%, 88.54% and 0.80 respectively. In a nutshell, the instruments which is designed has been valid, practical and reliable, which can be used by teachers to measure and foster students’ of higher-order thinking skills.

1. Introduction

Higher order thinking skills (HOTS) are the abilities to connect, manipulate, and transform the knowledge and experience to think critically and creatively in an effort to determine decisions and solve problems in new situations. Higher order thinking skills not only require remembering skills, but also require other higher skills [1, 2]. Indicators for measuring higher order thinking skills include analytical skills (C4), evaluating (C5) and creating (C6) [3]. Higher-order thinking skills can be trained in the learning process in the classroom. Learning that is done must provide space for students to find the concept of activity-based and meaningful knowledge. Therefore, in the 2013 curriculum revision emphasizing the need to integrate higher order thinking skills (HOTS) in learning [4].

The higher order thinking skills of students in Indonesia are still low, as evidenced by data obtained from international studies that measure students' higher order thinking skills, namely TIMSS (Trends In Mathematics and Sciences Study) and PISA (Program for International Student Assessment) which shows that Indonesia has participated in the TIMSS study since 1999, but the achievements of students in Indonesia are still categorized as low compared to other countries. In 2015, Indonesia was in the order of 44 out of 49 countries that participated in TIMSS with a score of 397 out of 500 international scores. And PISA data for 2018 shows that the ability of students in the field of science is ranked 69 with a score of 396 from the 78 countries evaluated [5].

We have also tested 30 questions taken from high-order thinking skills in valid, practical and quality functional materials by [6] on October 17, 2019 in Class XI IPA SMAN 1 Nan Sabaris. The
results of the test questions obtained an average class score of 38.45. Based on the International Center for the Assessment of Higher Order Thinking Skills, the scores of 21-40 are in the low category (low) so that the average class score obtained indicates that the higher order thinking skills of students at SMAN 1 Nan Sabaris are still low. This also shows that teachers are not training students to do higher order thinking skills. Moreover problem testing is also carried out by several researchers in various regions in West Sumatra, and the results show that the average ability of students in high-level thinking is still low, therefore an effort is needed to train students in order to develop higher order thinking skills.

Analysis of KD 3.3 daily test questions on genetic material made by teachers in SMAN 1 Nan Sabaris was conducted by experts, showing that of the 15 questions that have been made, the questions are at the C1 cognitive level, C2 and C3. While there are no questions about C4, C5 and C6. Questions that are at the cognitive level of C1 are 6 items, C2 cognitive level is 7 items, C3 cognitive level is 2 items. Whereas the cognitive level in KD 3.3 on Genetic material is at the level of analyzing /C4, so that what is demanded by the curriculum itself has not been done well, and has an impact on the low ability of students to analyze, synthesize and solve a problem.

The results of the analysis of items conducted by the Directorate of High School Development in assisting USBN in 2018/2019 academic year for 26 subjects in 136 referral high schools spread across 34 provinces, showed that of the 1,779 items analyzed were mostly at level-1 and level 2 Out of 136 Referral High Schools, only 27 schools compiled HOTS questions as much as 20% of all USBN questions made, 84 schools compiled HOTS questions below 20%, and 25 schools said they did not know whether HOTS questions were prepared. This is not in accordance with the demands of the 2013 curriculum assessment which further enhance the implementation of HOTS assessment models. So that when viewed on a large scale, there are still many teachers in Indonesia who have not yet carried out the HOTS assessment, because there are many who are less able to make instruments of higher order thinking abilities.

Based on the results of observations made at SMAN 1 Nan Sabaris with a biology teacher, it was revealed that the teacher had not fully conducted an instrument of higher order thinking skills. Teachers are not well-trained in making high-level assessment instruments, and it is difficult to change questions that initially have cognitive levels C1, C2, C3 to levels C4, C5, even C6. In addition, teachers also find it difficult to implement higher-order thinking assessment instruments because students' thinking abilities are still low, making it impossible for teachers to give HOTS questions to students.

The low ability of the teacher in making higher order thinking assessment instruments, because the teacher lacks practice in making high-order thinking questions, while the difficulty of students answering high-order thinking questions, because students are less trained in answering high-order questions in the learning process. This is in line with the opinion of [8], which states that one of the efforts to develop higher-order thinking skills is to use appropriate assessments.

Currently schools have started to launch HOTS learning in every subject. Therefore, it is necessary to develop higher order thinking skills assessment instruments that are valid, practical and reliable that can be used during learning evaluations and can be used as a reference for teachers to improve teacher skills in making assessment instruments. Higher order thinking skills and can activate the higher order thinking skills of students.

Assessment of higher order thinking skills can use the form of subjective tests and objective tests. Subjective test is a form of essay test. Essay tests are a form of description using one's own language. In essay-form tests students are required to think about and use what is known regarding the questions that must be answered. Objective test is a form of test consisting of true-false answers, multiple choice, completion, and matching [9]. The characteristics of assessment instrument of the higher order thinking skills (HOTS) that is, must be able to measure higher-order thinking skills which must have the ability to solve unfamiliar problems, the ability to evaluate strategies used to solve problems from different points of view and be able to find solutions that are different from the previous ways. In addition, the characteristics of higher-order thinking (HOTS) are divergent (allowing students to give
different answers), using multi-representation, based on contextual problems and using various question forms.\[10\]

The unavailability of instruments for evaluating higher order thinking skills about genetic material in SMAN 1 Nan Sabaris is shown from the results of KD 3.3 daily test questions about genetic material made by teachers. Based on this, the problem of this research is how to produce an instrument of assessment of higher order thinking skills ability about genetic material for students of class XII SMA/MA that is valid, practical and reliable. Genetic material is material that is more involved in the process, is contextual, and is at the cognitive level of analyzing/C4. So that genetic material supports the development of an instrument for assessing the ability to think highly. The purpose of this study is to produce instruments that assess the ability to think at a high level of genetic material for Class XII SMA/MA that is valid, practical, and reliable.

2. Methods
This type of research is developmental research with a 4-D (Four-D) development model which consist of the stages of define, design, develop, and disseminate. The disseminate stage was not carried out due to time and cost limitations. This research was conducted at the Faculty of Mathematics and Natural Sciences (FMIPA) Universitas Negeri Padang (UNP) and at SMAN 1 Nan Sabaris. The research subjects consisted of experts consisting of two lecturers from the UNP biology department and one biology teacher at SMAN 1 Nan Sabaris, practicality testing by two biology teachers at SMAN 1 Nan Sabaris and subject for test questions by 30 students of class XII at SMAN 1 Nan Sabaris. The object of this research is an instrument for assessing higher order thinking skills on Genetics material for class XII SMA/MA students. The resulting product was a research instrument for high-level thinking skills about valid, practical and reliable genetic material that was tested at SMAN 1 Nan Sabaris on March 18, 2020. The data collection instruments were quantitative in the form of a validation questionnaire and a practicality questionnaire.

3. Results and Discussion
3.1 Define Phase
3.1.1 Problem analysis. At this stage, it was carried out by finding data about the assessment instruments used in schools in biology learning activities as targets carried out in this study and conducting interviews with biology teachers of SMAN 1 Nan Sabaris. Based on the results of researchers' analysis of the assessment instruments used by the teacher, it is known that the cognitive level of the instrument used has not yet reached the cognitive level of high-level thinking ability (C4-C6). Based on the results of the analysis of the assessment instruments used by teachers when daily tests are still at the cognitive level of remembering (C1), understanding (C2), and applying (C3). Furthermore, based on the results of an interview with one of the biology teachers at SMAN 1 Nan Sabaris, it is known that there are several factors of the low level of thinking ability of high school students of SMAN 1 Nan Sabaris namely the teacher is still having difficulty making assessment instruments of higher order thinking skills especially in changing a problem that is at the C1-C3 level to C4-C6. Therefore, researchers developed an instrument of high-order thinking skills ability about genetic material for class XIISMA/MA students.

3.1.2 Student Analysis. The results of the analysis of learners are obtained through valid and practical test questions, namely the instrument of assessment of high-level thinking abilities on fungi material by\[10\], the average class results are 38.45 and are classified in the low category. This happens because so far the assessment instruments used by students when evaluating learning only focus on the ability to memorize only and not accustom students to think at a higher level.

3.1.3 Curriculum Analysis. Curriculum analysis is carried out so that the resulting high-order thinking skills ability assessment instrument refers to the Core Competencies (KI) and Basic Competencies (KD) in the applicable curriculum. The analysis is more focused on the breakdown of Core
Competencies (KI) and Basic Competencies (KD) for genetic material which are translated into several indicators. The analysis was conducted with reference to the 2016 revised 2016 curriculum biology syllabus.

3.1.4 Concept Analysis. Concept analysis is done by identifying the main concepts about genetic material. The main concept of genetic material is the relationship between genes, alleles, DNA, chromosomes (chromatin), structure and function of chromosomes, DNA, RNA and protein synthesis. The main concepts that must be mastered by students to be able to work on assessment instruments of higher order thinking skills. Reference to this analysis is the subject matter in the 2016 Revised 2013 Curriculum syllabus.

3.1.5 Formulation of Learning Objectives. Reference to the formulation of this learning objective is an indicator of achievement of competence that refers to the 2013 curriculum.

3.2 Design Phase
The results of the design of the instrument of high-order thinking skills ability about genetic material are the problem lattice by selecting interesting and contextual stimuli, writing items that correspond to the problem lattice, and making the answer key. \cite{11} states that the question grids are well made so that questions produced from quality grids. The design of the instrument lattice of high-order thinking skills ability is made based on indicators of achievement of competencies that have been developed, and using stimuli, tables, discourse, diagrams and pictures.

![Image](image.jpg)

**Figure 1.** Cover of Assessment of High-order Thinking Skills About Genetic Material

The initial design in the development of a higher order thinking skills ability assessment instrument developed was 50 items. The question lattice consists of six indicators of competency achievement, each competency indicator is broken down into 2-18 question indicators. Each question indicator consists of one question. The types of questions that will be developed are 30 multiple choice questions, there are 10 problem causal questions, and 10 multiple choice association questions. The higher order thinking skills ability assessment instrument developed was at the C4-C6 cognitive level (analyzing, evaluating, and creating). The number of questions at the C4 cognitive level amounts to 37 items, the number of questions at the C5 cognitive level is 7 items and at the C6 level there are 5 items. This instrument is equipped with a cover (cover), questions consist of giving discourse questions, presented pictures and tables, answers to questions consists of 5 choices of multiple choice
questions, cause and effect questions, multiple choice associations, and also equipped with instrument instructions, which consists of general instructions and specific instructions.

3.3 Development Phase

3.3.1 Logical Validity of Higher Order Thinking Skills Assessment. The logical validity of the higher order thinking skills assessment instrument aims to prove the validity of the assessment instrument developed. Validation was carried out by two Biology Department lecturers and one biology teacher at SMAN 1 Nan Sabaris. The logical validity analysis of the high-order thinking skills assessment instrument is briefly shown in table 1.

| No | Analyzed aspects                   | Validity | Criteria |
|----|-----------------------------------|----------|----------|
| 1  | Theori                            | 87.96%   | Valid    |
| 2  | Construction                      | 85.18%   | Valid    |
| 3  | Language                          | 86.11%   | Valid    |
| 4  | Higher Order Thinking Skills      | 91.66%   | Very     |
|    | Average                           | 87.72%   | Valid    |

The results of the validation of the higher order thinking skills of assessment instrument obtained a value of 87.72% with valid criteria. This means that the instrument for evaluating the ability to think high level about genetic material for Class XII SMA/MA students can already be used.

Analysis of data from the validity questionnaire of the higher order thinking skills assessment instrument by lecturers and teachers is based on four components, namely material, construction, language, and higher order thinking skills. The validation results show that the mean percentage value of all aspects of the three experts is 87.72% with valid criteria. The percentage value indicates that the instrument for assessing higher order thinking skills developed is valid both in terms of material aspects, construction, language and higher order thinking skills assessment instruments.

Based on the results of the data analysis carried out, the experts test results of the higher order thinking skills assessment instrument by the experts in terms of the material developed were classified as valid with a value of 87.96%. This material aspect is related to science and the level of thinking on the items that have been made [12]. The percentage value obtained means that the instrument developed is in accordance with the core competence (KI) and basic competence (KD). This is in accordance with the Ministry of National Education (2017) which states that the material developed must be in accordance with Basic Competencies and Indicators.

In the construction aspect of the instrument for the assessment of high-order thinking skills, the system material is valid with a value of 85.18%. This assessed construction aspect is useful for improving the quality of questions from writing questions and answers [13]. Based on the results of validity, this shows that the assessment instrument used is using question construction with a clear formula. According to [14] items with construction are good if the subject matter is formulated clearly which includes only the statements needed.

In a review of the language aspect, the instrument developed was declared valid with a value of 86.11%. The language aspect of the assessment component relates to the communicative and clarity of the items in question [12]. The percentage of scores obtained shows that the assessment instrument has used language in accordance with the rules of Indonesian, using communicative language [15].

The validation results in the aspect of higher order thinking skills in the assessment instrument developed were stated to be very valid with a value of 91.66%. The overall results obtained by the experts in terms of high-order thinking skills are valid. The instrument for assessing higher order thinking skills that has been developed can be used as an instrument that can activate students' higher-order thinking skills.
3.3.2 Empirical Validity of Higher Order Thinking Skills Assessment. The value of the empirical validity of the high-order thinking skills of assessment instrument was obtained by analyzing the items. Item analysis was carried out after testing the higher order thinking skills assessment instrument that had been developed for students. The item analysis data was obtained after the assessment instrument was tested on 30 students of class XII IPA 2 SMAN 1 Nan Sabaris with a total of 50 multiple choice items. Analysis of the items was carried out using the ANATES program Multiple Choice Version 4.0.9.

Based on the results of the item analysis, 41 items were valid with a percentage of 82% and 9 items were invalid with a percentage of 18%. The logical validity analysis of the high-order thinking skills assessment instrument is briefly shown in table 2.

Table 2. Results of the Empirical Validity Analysis of Higher Order Thinking Skills Assessment

| Validity | Form of questions | Multiple choice | Multiple choice causality | Multiple choice of association |
|----------|-------------------|-----------------|---------------------------|-----------------------------|
| Valid    | Multiple choice   | 26 Question (0.29-0.67) | 9 Question (0.32-0.61) | 6 Question (0.29-0.54) |
| Invalid  | Multiple choice   | 4 Question (0.47-0.24) | 1 Question (0.518) | 4 Question (0.16-0.17) |

Empirical validity aims to determine the reliability level of the assessment instrument items. The items of the assessment instrument are said to have empirical validity if the results are consistent with experience. The result of higher order thinking skills assessment is valid based on empirical validity. According to [16] an item is said to be valid if it has great support for the total score.

The predictive or predictive value of empirical validity indicators shows that, only about 41 questions can be used to test students’ higher order thinking skills in the future after being tested at SMAN 1 Nan Sabaris. Meanwhile, 9 more questions were not maintained, because they had a correlation <0.20. According to [16], if the value of r lies between 0.00 to 0.20 then this correlation coefficient is interpreted as having very low reliability so that it does not show significant or invalid.

3.3.3 Practicality of Higher Order Thinking Skills Assessment

The practicality test of a high order thinking skills assessment instrument was carried out by giving a questionnaire to two teachers of SMAN 1 Nan Sabaris. A brief analysis of the practicality test data can be seen in table 3.

Table 3. The Results of the Practicality Assessment of Higher Order Thinking Skills

| No | Analyzed Aspect | Validity | Criterium |
|----|----------------|----------|-----------|
| 1  | Examination    | 90.62%   | Very Pract |
thinking skills is categorized as practical with a value of 88.54. In terms of implementation aspects, the assessment instrument is very practical according to the teacher with a score of 90.62. This shows that the instrument developed is time efficient, does not require students to use a lot of equipment in working on the questions, the questions are done in the question answer sheet and gives freedom to students to work on questions that are considered easy by students. This is in line with [16], which states that a test is said to be practical if it is easy to implement, for example, it does not require a lot of equipment and gives students the freedom to do the parts that are considered easy by students first.

In terms of examination, the assessment instrument developed has a value of 87.50% with a practical category, this shows that the assessment instrument is easy to carry out examinations because the questions are equipped with an answer key, in line with [16] which states that, easy to check is a test that is equipped with both an answer key and a scoring guide. For objective form questions, the examination is easier to do if it is done by students in the answer sheet.

In terms of the aspect of question instructions, the teacher's practical assessment instrument was with a value of 87.50%. This shows that the instrument developed is equipped with clear instructions, so that it can be tried out by others, in line with [16] which states that practical assessment instruments are equipped with clear instructions so that people can give / initiate them. other.

Overall, the instrument for assessing higher order thinking skills developed has practical criteria. This means that the assessment instrument is easy to use both in terms of aspects of implementation, examination and question instructions.

3.3.4 Reliability of Higher Order Thinking Skills Assessment. The reliability of the higher order thinking skills instrument was obtained from the results of the item analysis using the ANATES program of Multiple Choice Version 4.0.9. Based on the results of the item analysis, 41 reliable items were produced with a reliability value of 0.80.

The resulting high-order thinking skills assessment instrument is reliable, which has an r11 of 0.80. From these results, it can be said that the resulting assessment instrument has high reliability. According to [17] a person is said to have high reliability if it has r11 equal to or greater than 0.70. Meanwhile, smaller than 0.70, it does not have high reliability.

Reliability relates to the trust problem of a test. According to [16] a test can be said to have a high level of confidence if the test can provide permanent results. In line with [18] opinion reliability is an instrument that shows reliable (consistent) measurement results, a reliable assessment instrument will show relatively the same results if tested on the same group at different times.

Reliability on the other hand is defined as ‘the extent to which test scores are free from measurement error’ [19]. It is a measure of stability or internal consistency of an instrument in measuring certain concepts [20]. Test-retest reliability is a form of reliability achieved when the same instrument is administered to the same group of respondents on two different occasions and yet look at the correlation between the two sets of scores [21]. The higher the correlation value meaning that the instrument is more reliable. There is a relationship between validity and reliability. Any instrument can be reliable but not valid however, it cannot be valid if it is not reliable [20]. In other words, if an instrument is valid, it must reliable. And, in general, checking for validity of an instrument is more difficult than checking for reliability because validity is measuring data related to knowledge whereas reliability only concerns with the consistency of scores [22].

4. Conclusion
The process of developing an instrument for assessing higher order thinking skills about genetic material for class XII SMA/MA students was developed through a 4D model development process which consists of the stages of define, design, develop, and disseminate. The disseminate stage was not carried out due to time and cost limitations.

The instrument for assessing high-order thinking skills resulted in a logical validity average value of 87.72% with valid criteria, the value of empirical validity is 41 valid questions with a percentage of 82%, an average practical value of 88.54% so that it can be declared practical and has reliability value.
of 0.80 so that it can be declared reliable. So, it can be concluded that an assessment instrument for high-level thinking about genetic material has been produced for class XII SMA/MA students that is valid, practical and reliable that can be used by teachers to measure and activate students' higher order thinking skills.

**Acknowledgment**
The final manuscript of this publication was presented at Webinar SEMIRATA 2020 on The 4th International Conference on Mathematics, Science, Education and Technology (ICOMSET) in conjunction with The 2nd International Conference on Biology, Science and Education (ICoBioSE) virtually at Universitas Negeri Padang, Padang City, Indonesia on September 19th, 2020.

**References**

[1] Rofiah, Aminah E N S and Ekawati E Y 2013 Penyusunan Instrumen Tes Kemampuan Berpikir Tingkat Tinggi Fisika pada Siswa SMP *Pendidikan Fisika* 1

[2] Rahmi Y L, Alberida H and Astuti M Y 2019 Enhancing students’ critical thinking skills through inquiry-based learning model *J. Phys. Conf. Ser* 1317 012208

[3] Anderson L W and Krathwohl D R 2001 *A Taxonomy of Learning, Teaching, and Assessing: A Revision of Bloom’s Taxonomy of Educational Objectives* (New York: Longman)

[4] Supardi 2015 *Penilian Autentik* (Jakarta: Raja Grafindo Persada)

[5] OECD 2018 Contry Result Form PISA 2018 Online http://www.bps.go.id/linktabledinamis/view/id/1054

[6] Haryanto 2018 *Pengembangan Instrumen Penilaian Kemampuan Berpikir Tingkat Tinggi pada Materi Fungi untuk Peserta Didik SMA/MA Kelas X Skripsi* (Padang: Universitas Negeri Padang)

[7] Kementerian Pendidikan dan Kebudayaan Republik Indonesia 2015 *Tentang Kebijakan Perubahan Ujian Nasional p-38.*

[8] Rahmi Y L and Alberida H 2017 Improving Students’ Higher Order Thinking Skills Trough Portofolio Assessment on Biology Curriculum and Textbook Analysis *Bioeducation Journal* 1 (1) pp 22-23

[9] Suwandi S 2009 *Model Asesmen dalam Pembelajaran* (Surakarta: Mata Padi Presindo)

[10] Isbandiyah S and Sanusi A 2019 *Modal Penyusunan Soal Berpikir Tingkat Tinggi (Higher Order Thinking Skills) Biologi* (Jakarta: Direktorat Pembinaan Sekolah Menengah Atas)

[11] Susiatin 2019 Meningkatkan Kemampuan Guru dalam Menyusun Kisi-Kisi Soal dengan Metode Pendampingan Pola OCF *J. Dinamika Manajemen Pendidikan* 4 pp. 17-24

[12] Munadi S 2011 Analisis Validasi Kualitas Soal Tes Hasil Belajar pada Pelaksanaan Program Pembelajaran *Cakrawala Pendidikan* 1 pp 145-159

[13] Ariningrum W 2016 *Aspek Materi, Konstruksi, dan Bahasa pada Soal Sastra Ujian Nasional Tingkat SMK Mata Pelajaran Bahasa Indonesia Tahun 2014/2015* Naskah Publikasi (Surakarta: Universitas Muhammadiyah Surakarta) pp 5-13

[14] Kadir A 2015 *Menyusun dan Menganalisis Tes Hasil Belajar Jurnal Al-Ta’dib* 8 2 pp. 70-81

[15] Depdiknas 2017 *Panduan penulisan oleh Pendidik dan Satuan Pendidikan untuk Sekolah Menengah Atas* (Jakarta: Direktorat Pembinaan SMA Dirjen Pendidikan Dasar dan Menengah)

[16] Arikunto S 2012 *Dasar-Dasar Evaluasi Pendidikan* (Jakarta: Bina Aksara)

[17] Sari P I 2020 *Pengembangan Instrumen Penilaian kemampuan berpikir Tingkat Tinggi tentang Materi Sistem Koordinasi untuk Peserta Didik Kelas XI SMA/MA Skripsi* (Padang: Universitas Negeri Padang)

[18] Dewi N P, Rahmi Y L, Alberida H and Darusyamsu R 2020 Validitas dan Reliabilitas Instrumen Penilaian Kemampuan Berpikir Tingkat Tinggi tentang Materi Hereditas untuk Peserta Didik SMA/MA *Jurnal Eksakta Pendidikan* 4 2 p 145

[19] Muijjs D 2011 *Doing Quantitative Research in Education with SPSS* (London: SAGE Publications)

[20] Jackson S L 2003 *Research Methods and Statistics, A Critical Thinking Approach* (USA)
[21] Pallant J 2007 *SPSS Survival Manual: A step by step guide to data analysis using SPSS for Windows 3ed* (New South Wales, Allen and Unwin)

[22] Ghazali N H Md 2016 A Reliability and Validity of an Instrument to Evaluate the School-Based Assessment System: A Pilot Study *International Journal of Evaluation and Research in Education (IJERE)* 5 2 p149