

PAPER • OPEN ACCESS

## Design of Computer-Based Testing for Higher-order Thinking Skills on Static Fluid Material

To cite this article: Diah Eka Pratiwi *et al* 2020 *J. Phys.: Conf. Ser.* **1467** 012059

View the [article online](#) for updates and enhancements.



**IOP | ebooks™**

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection—download the first chapter of every title for free.

# Design of Computer-Based Testing for Higher-order Thinking Skills on Static Fluid Material

Diah Eka Pratiwi<sup>1</sup>, Agus Suyatna<sup>\*1</sup>, V Viyanti<sup>1</sup>

<sup>1</sup>Physics Education of Graduate Program, Lampung University, Jl. Sumantri Brojonegoro No. 01, Lampung, Indonesia.

\*Correspondence author: [asuyatna@yahoo.com](mailto:asuyatna@yahoo.com)

**Abstract:** The purpose of this study is to develop a Computer Based Test (CBT) design model that is suitable for stimulating students' higher-order thinking skills (HOTS) on static fluid material. The development of CBT-HOTS is used as a practice (drill) which is expected to be able to familiarize students in working on computer-based questions. The design of this study is guided by the ADDIE instructional development model with 5 stages of development, namely: the analysis phase, the initial product design stage, the development stage, the product implementation stage, and the product evaluation stage (evaluation). This article only reported up to the design analysis and design stages. Data collection used was a questionnaire collected from 12 high school physics teachers in Lampung Province. The data analysis used was descriptive qualitative analysis. The results of this study were the design of the CBT model that is suitable for stimulating students' HOTS according to the needs in the field. The designs that were developed include the design of the illustration problems, the design of the quiz settings, the design of the questioning settings, and the design of the stimulus for several types of questions namely fill in the blank, matching, and sequence.

## 1. Introduction

One of the spearheads of the progress of a nation is education. Lately, education in Indonesia has implemented the 2013 curriculum to prepare students who can face the 21st century by adjusting the education of the industrial revolution 4.0. The impact of this adjustment is the development of technology in the world of education, including the transition of paper-based test based on paper or paper-based test (PBT) to computer-based test or computer-based test (CBT) [1]. The use of computer-assisted assessments makes CBT an attractive proposition for many higher education institutions [2].

Even though the exam has been switched to CBT, students still find deviations in the form of cheating on multiple-choice questions even though they sit far apart [3]. The implementation of CBT that has been implemented is still not perfect because the CBT developed is mostly in the form of multiple-choice and seems monotonous. This was stated by one of the teachers in Lampung Province who stated that there were still cheating and lack of stimulus to the questions which limited the students' real experiences. Besides, CBT is still considered ineffective in evaluating creativity, problem-solving skills, and critical thinking [4]. These three components constitute the category of



higher-order thinking or higher-order thinking skills (HOTS) [5]. So it can be said that CBT has not been able to measure students' HOTS.

Based on the analysis of needs distributed to 12 teachers in Lampung Province, data obtained that 100% of teachers agreed that CBT could be applied in evaluating learning to minimize student fraud (cheating) and optimize time in correcting and monitoring students' understanding of a Theory. However, only 14% of schools have used computer-based tests as drills. Respondents claimed that the use of CBT was only for the UN or as an assessment of learning so that many students were unfamiliar with tests using computers. Then, not a few teachers also stated that HOTS questions had not been implemented in their schools because of the lack of teacher understanding of HOTS questions. The type of questions that teachers often apply is multiple-choice questions and multiple response types.

This was reinforced by the findings of the Directorate of High School Development in assisting USBN in 2018/2019 of 136 referral high schools spread across Indonesia. The findings obtained the fact that the questions made are mostly at Level-1 and Level-2. Of the 136 schools, only 27 schools compiled HOTS questions, 84 schools compiled HOTS questions below 20%, and 25 other schools said they did not know whether the questions compiled were HOTS questions or not. This is not following the demands of the 2013 curriculum which further enhances the implementation of HOTS assessment models.

There are not only obstacles regarding the application of the HOTS problem. Material from physics itself is still often considered abstract by students because of the limitations of students' real experiences. Also, the presentation of questions with long words makes students often have difficulty understanding and answering physics problems. As a result, students often experience misconceptions about physical material. As for one of the materials which still often occurs misconceptions in class XI students are on static fluid material. That is because static fluid material is material that is often found by students in daily life, so indirectly students already have the initial knowledge gained from their own life experiences. Furthermore, not all of this knowledge is practiced directly in class as a confirmation of knowing whether the knowledge he constructed himself is correct or not. The inaccuracy of initial knowledge possessed by students can then develop into misconceptions. This is supported by the statement from Warton et al. [6] which states that there are still many students who experience misconceptions on static fluid material. Then, Goszweski et al. [7] and Wagner et al. [8] also proved the existence of misconceptions through the presentation of questions about the static fluid that was allegedly the occurrence of misconceptions.

To perfect the demands of the 2013 curriculum, it is necessary to have a CBT model that can measure and even train students' HOTS by optimizing the role of stimulus, choosing the type of questions, feedback, quiz settings, and question settings on static fluid material.

## 2. Method

The design of this study was guided by the ADDIE instructional development model with 5 stages of development, namely: the analysis phase, the initial product design stage, the product development stage, the product implementation phase, and the product evaluation stage [9]. In the first stage, analysis, using a data source in the form of a questionnaire from 12 teachers spread in Lampung Province to determine the criteria for the CBT model needed to stimulate students' HOTS on static fluid material. Furthermore, the questionnaire obtained from respondents was used as a reference for designing CBT-HOTS products.

In the second stage, the design of the CBT-HOTS model is based on the questionnaire answers obtained from respondents. However, this article does not cover the product development stage, the product implementation phase, and the product evaluation phase.

Data were collected through a questionnaire instrument using a Likert scale with a scale of 1-5. Furthermore, a needs analysis technique regarding CBT is obtained from converting respondents' questionnaire statements into percent using the following formula.

$$\frac{\sum \text{respondent}}{\sum \text{every respondent}} \times 100\%$$

Then, the CBT-HOTS display valuation analysis technique is obtained from the conversion of the respondent's questionnaire statement to a score using the following formula.

$$\frac{\sum \text{respondent}}{\sum \text{every respondent}} \times \text{maximal score}$$

The scores obtained are then used as a reference for quantitative descriptive analysis as in Table 1

<b>Table 1.</b> Score Conversion	
Average Score	Decision
4,20 – 5,00	CBT-HOTS design is very suitable
3,40 – 4,19	CBT-HOTS design is appropriate
2,60 – 3,39	CBT-HOTS design is quite appropriate
1,80 – 2,59	CBT-HOTS design is not appropriate
1,00 – 1,79	CBT-HOTS design is not suitable

### 3. Result and Discussion

The following are the results of the analysis of CBT needs from 12 teachers from various regions in Lampung province. The data can be seen in Table 2.

**Table 2.** Results of needs analysis regarding Computer Based Test (CBT)

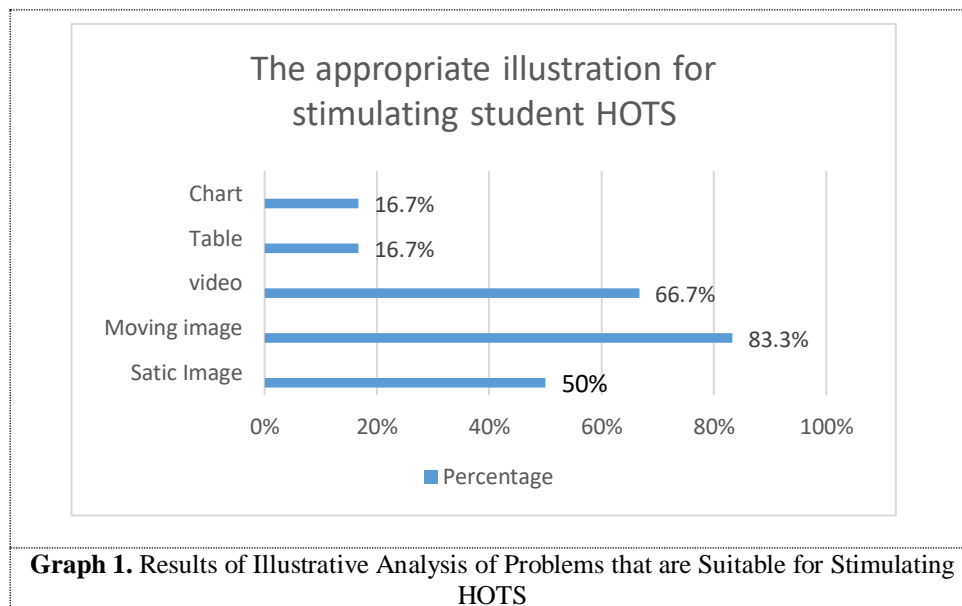
No	Teacher's Statement	%
1	Provision of computer facilities to support the learning process	100
2	Application of computer-based learning media	100
3	The method applied by the teacher can assist students in understanding static fluid material	100
4	Knowing about computer-based exams or computer-based tests (CBT)	100
5	Agree if the test or test given to measure a student's ability is converted to a computer-based test (CBT)	100
6	Agree if the application of computer-based tests (CBT) can reduce cheating by students such as cheating	100
7	Agree if the application of computer-based tests (CBT) can facilitate assessment rather than paper-based tests (PBT)	100
8	Application of computer-based exams (CBT) in schools	100
9	The application of computer-based tests has been used as a daily practice drill (drill) in addition to the National Examination	16.67

10	Program ownership can stimulate HOTS capabilities	0
11	Never gave HOTS questions	25
12	Agree if the illustrations in the form of animation/video help enrich the form/type of questions and the level of ability to be measured	100

Based on the results of the needs analysis in Table 2, all schools have provided computer facilities that can be used by teachers as learning media and are expected to be able to help understand physical material, especially static fluid material. Besides being used to deliver material, the computer can also be used as a test or often known as a computer-based test (CBT). Unfortunately, this computer-based exam is only used in assessments of learning such as the National Examination (UN). Based on the respondent's information, very few schools use computer-based exams on assessment as learning such as for daily tests or drills. During this time, teachers only use paper-based exams that have the potential for cheating by students such as cheating. When viewed from the perspective of students, the application of CBT can increase competition between students who are fair if done correctly and if viewed from the teacher's point of view, the application of CBT can facilitate the assessment process because the assessment should spend about 20% -30% of teacher professional time [10]. Another convenience obtained from the use of CBT is reducing greenhouse gases because it saves tons of paper [11] and allows more innovative and authentic assessments because of the capacity of more sophisticated technology [12]. Also, the 2013 curriculum requirements for HOTS questions have not yet reached the desired results. Only 25% of teachers have given HOTS questions and one of them stated that they did not know whether or not they were given HOTS questions. Most teachers think that HOTS questions are questions that students find difficult. For teachers who have given HOTS type questions, they say that only giving questions on a paper without having any program can stimulate students' HOTS abilities. All teachers agree that giving illustrations to questions in the form of videos / moving images/static images can help enrich the shape/type of questions and the level of ability to be measured.

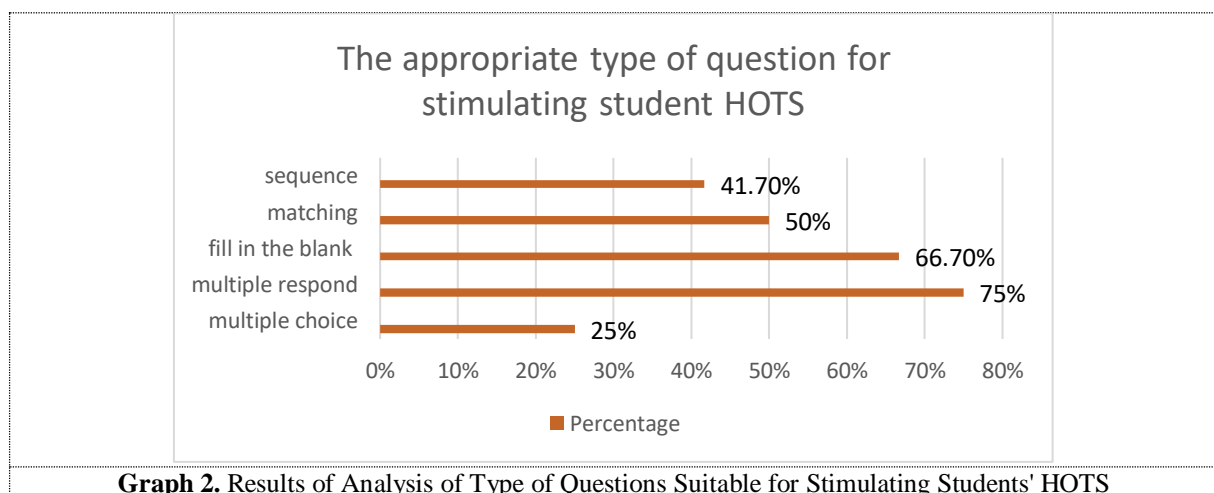
Computer-based exams have several important advantages compared to traditional paper-based exams (PBT) such as efficiency, assessment and fast feedback [12]. Nonetheless, CBT is still considered ineffective in evaluating creativity, problem-solving abilities, and critical thinking that are included in the HOTS assessment category [4]. Therefore, it is necessary to develop a CBT model that can measure HOTS abilities while at the same time can be used as a question exercise ( drill) for students.

The following are the results of the illustrative analysis of questions that are suitable for stimulating students' HOTS.



The need for the development of CBT to overcome the problem of Paper-Based Test (PBT) such as the form of the questions used when the test is difficult to be varied and the appearance of limited questions that is only two dimensions [13] can be overcome through the selection of illustrated questions. Based on the results of the illustrative analysis of the questions that have been obtained from respondents' answers, it is said that moving images occupy the highest position of 83.3%, followed by videos of 66.7%, static images by 50%, while graphs and tables by 16.7%. From the graph, it can be seen that 50% and above respondents agree to use moving images, videos, and static images to illustrate the questions that are suitable for stimulating students' HOTS. But for tables and graphs, the data is below 50%. From this graph, it can be concluded that the illustration of questions that are suitable and have the potential to be developed in stimulating students' HOTS are videos, moving pictures and static images.

The following is an analysis of the type of questions suitable for stimulating students' HOTS.



According to King, Goodson, & Rohani [14], there are three task items/formats that are useful in measuring students' HOTS: (a) choosing, including plural choices, matching, and sorting; (b) short

answers, essays, and performance or assignments; and (c) explain, including giving reasons for the selected answers. HOTS measurement through tests can use multiple-choice questions, essay questions, fill in the blank, matching, sequence, and other questions [15]. During this time, teachers usually provide questions with multiple-choice types that tend to often be found cheating actions between students. According to Setiawati, et al. [16] there are several alternative types of questions that are suitable for HOTS, namely complex multiple-choice (true/false, or yes/no) and the form of the problem description. However, the form of complex multiple-choice questions often results in inaccurate answers because students only guess the answers, while the type of problem description is quite difficult to combine with CBT because the answer key recorded on the computer will be more than one estimate. Based on Graph 2, the possible types of HOTS questions that are suitable to be developed for CBT are fill in the blank, matching, and sequence. This is because the three of them are in the middle, which is around 40% to 70% and has the potential to be developed.

In addition to the need for illustrative design questions for CBT, it should also be noted about the appearance of the CBT itself. Item and answer items should appear randomly. This can reduce cheating on students, both sitting close together and sitting far apart. Even so, there are still respondents who disagree with the reasons that it will complicate students. The computer screen can display the time that has been used in working on the test or the remaining time available for working on the test. However, in this case, respondents more agree to display the remaining time available for the test. Also, assigning values to each item can be distinguished based on the level of thinking (C1, C2, C3, C4, C5, C6), the difficulty level of the questions, and the type of questions (short answers, matching, sorting, reasonable choices, etc.). 100% of respondents agree if the marking is distinguished by the level of thinking ability (C1, C2, C3, C4, C5, C6) and the level of difficulty of the questions. However, there were 8.3% of respondents who disagreed if the marking was differentiated based on the type of question. The feedback given can be done on students' answers that are right or wrong, but respondents tend to choose feedback given to students' answers that are wrong after all items are done rather than feedback given each item that has been done.

In this CBT, respondents agreed if students were given another chance to repeat the test even though there were 16.7% of respondents disagreed with the statement. Also, according to 91.6% of respondents more agree if the collection of answers (click submit) should be done after all items have been completed (or after time runs out) with risk, students can rework the questions on it.

**Table 3.** Results of the Analysis of the CBT-HOTS Design Display Assessment

<b>CBT-HOTS Design</b>	<b>Score</b>	<b>Remarks</b>
<b>Content</b>		
Random item appearance	4.08	CBT-HOTS design is appropriate
The appearance of options (answer choices) randomly	4.00	CBT-HOTS design is appropriate
The appearance of time already spent working on test items	4.33	CBT-HOTS design is very suitable
The appearance of the remaining time available for work on the test	4.83	CBT-HOTS design is very suitable
Assessment weights differ according to the level of thinking (C1, C2, C3, C4, C5, C6)	4.67	CBT-HOTS design is very suitable
The weight of the assessment is divided according to the level of difficulty of the questions	4.75	CBT-HOTS design is very suitable
The assessment weights differ according to the type of Problem	4.33	CBT-HOTS design is very suitable

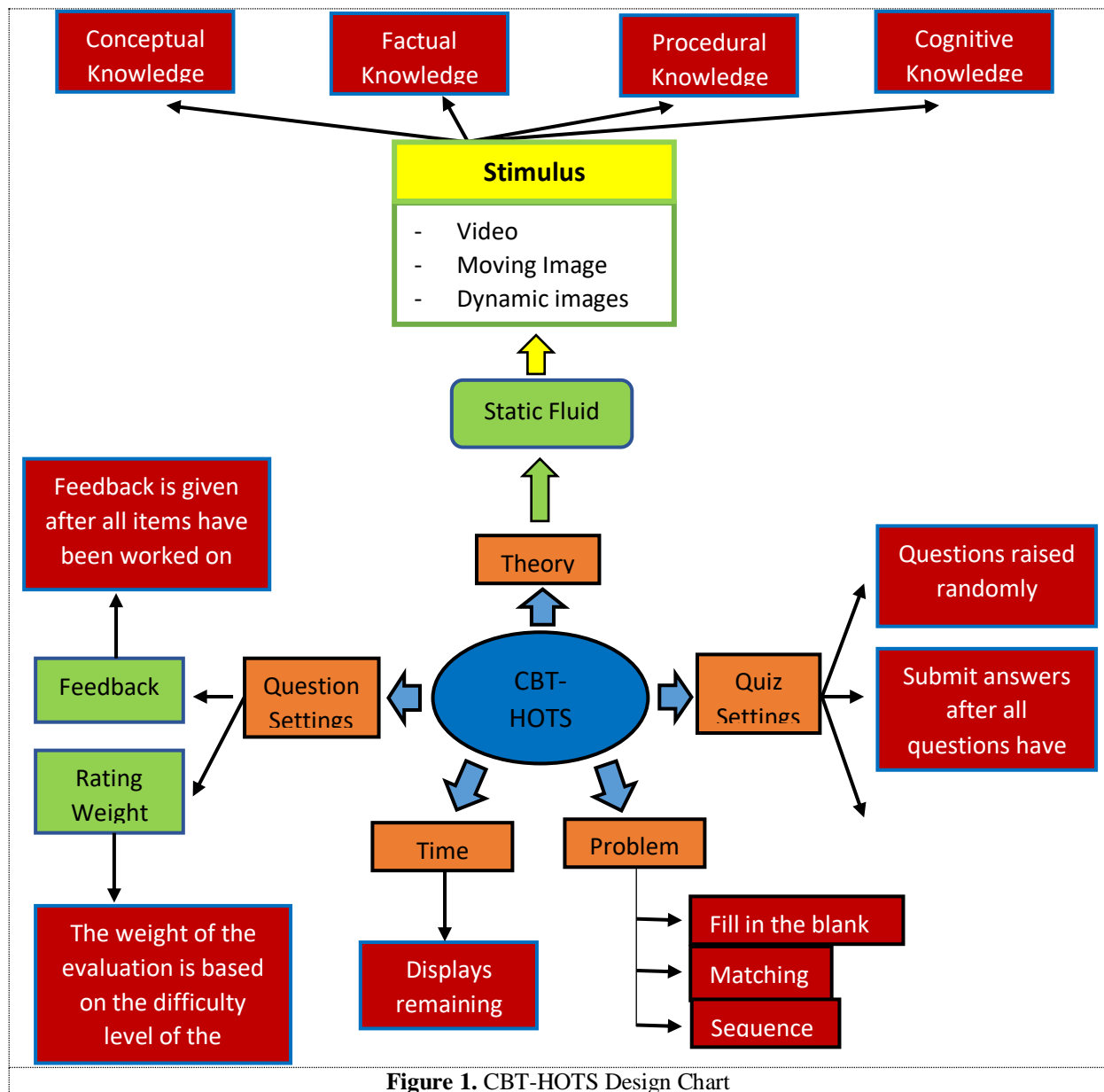
<b>CBT-HOTS Design</b>	<b>Score</b>	<b>Remarks</b>
<b>Content</b>		
Feedback is given to students who can answer the questions correctly	4.33	CBT-HOTS design is very suitable
Feedback is given to students who cannot answer the questions correctly	4.67	CBT-HOTS design is very suitable
Feedback is given for each item that has been worked on	3.67	CBT-HOTS design is appropriate
Feedback is given after all items have been worked on	4.58	CBT-HOTS design is very suitable
CBT-shaped tests are allowed to repeat the test	3.92	CBT-HOTS design is appropriate
The collection of answers (click submit) is done after each item is completed, with the risk that students cannot rework the questions on it	3.33	CBT-HOTS design is quite appropriate
The collection of answers (click submit) is done after all items have been completed, (or after time runs out) with the risk, students can rework the questions on it	4.42	CBT-HOTS design is very suitable

Based on Table 4, it can be said that all content is deemed appropriate for the development of CBT designs. But for which selection is better, we can review the average score obtained from respondents. The appearance of items and answer choices should be raised randomly using the remaining time available in the execution of the test [17]. The weight of the assessment can be distinguished based on the level of thinking, the level of difficulty of the questions, and the type of questions, but it is better if the weight of the assessment is distinguished by the level of difficulty of the questions. The feedback can be given to students who cannot answer the questions correctly after all the items are worked on. Feedback is one of the most powerful tools to improve student learning [18]. From the student side, this online feedback will help students improve the quality of their learning [19-22] and from the teacher's side, will provide students with timely information about their academic progress and can diagnose and analyze weaknesses. and their strength [23]. In general, students prefer to receive general feedback to the class regarding 'explanation of common mistakes and how to improve future performance [24].

Also, students are allowed to repeat the test as training material in developing student knowledge. Then, the collection of answers (click submit) is done after all items have been completed, (or after time runs out) with risk, students can rework the questions on it.

After going through the stages of collecting and analyzing data obtained from physics teachers in Lampung Province, the CBT-HOTS design was designed on static fluid material. The CBT-HOTS design chart can be seen in Figure 1.





The CBT-HOTS design in Figure 1 consists of quiz settings, question settings, time design, and design of several types of questions that are used to stimulate students' HOTS on static fluid material. In the quiz setting, we can set the appearance of questions and submit answers questions. Questions will be arranged so that they appear randomly and submit answers allowed after all the questions are worked on. Next, design the question settings to set the feedback and determine the weighting for each question. Feedback will be given after all items are worked on and the weight of the assessment is set based on the level of difficulty of the questions. Then, the time design display shows the remaining time and type of questions used in the CBT-HOTS design are fill in the blank, matching, and sequence in the static fluid items. The stimulus used for static fluid items is in the form of videos, moving pictures and static images arranged based on the dimensions of knowledge such as conceptual knowledge, factual knowledge, procedural knowledge, and metacognitive knowledge.

#### 4. Conclusion

Based on the results of the analysis of CBT-HOTS, the design of CBT-HOTS needed to stimulate students' HOTS in accordance with the needs in the field include the design of the quiz settings, design question settings, design time, and stimulus design that is found in several types of questions namely fill in the blank, matching, and sequence. The stimulus design created is adjusted to measure the dimensions of students' conceptual, factual, procedural and metacognitive knowledge.

#### Reference

- [1] Yan, C. 2012. Replacing paper-based testing with computer-based testing in assessment : Are we doing wrong ? 64, 655–664. <https://doi.org/10.1016/j.sbspro.2012.11.077>
- [2] Jimoh, R. G., Abduljaleel, K. S., & Kawu, Y. K. 2012. Students' Perception of Computer Based Test (CBT) for Examining Undergraduate Chemistry Courses. *Journal of Emerging Trends in Computing and Information Sciences*, **3** 2 125–134.
- [3] Novrianti. 2014. Pengembangan *Computer Based Testing* (CBT) sebagai Alternatif Teknik Penilaian Hasil Belajar. *Lentera Pendidikan*, **17** 1 34-42.
- [4] Fluck, A., Pullen, D., & Harper, C. 2009. Case study of a computer based examination system. *Australasian Journal of Educational Technology*, **25** 4 509-523.
- [5] Conklin, W. 2012. *Higher-order Thinking Skills To Develop 21<sup>st</sup> Century Learners*. Huntington Beach: Shell Educational Publishing, Inc.
- [6] Wartono, Saifullah, A.M., Sugiyanto. 2016. Identifikasi Miskonsepsi Siswa Kelas X pada Materi Fluida Statis dengan Instrumen Diagnostik *Three-Tier*. *Jurnal Pendidikan Pembelajaran (JPP)*, **23** 1
- [7] Goszewski, M., Moyer, A., Bazan, Z., dan Wagner, DJ. 2012. *Exploring Student Difficulties with Pressure in a Fluid*. PERC Proceedings, Published by the American Association of Physics Teachers under a Creative Commons Attribution.
- [8] Wagner, D. J., Carbone, E. & Lindow, A. 2013. *Exploring Student Difficulties with Buoyancy*. PERC Proceedings, Published by the American Association of Physics Teachers under a Creative Commons Attribution.
- [9] Branch, R. M. (2009). *Instructional Design-The ADDIE Approach*. (New York: Springer)
- [10] Santrock, J. 2008. *Psikologi Pendidikan Edisi Kedua*. (Jakarta: Kencana)
- [11] Chua, Y. P. 2012. Computers in Human Behavior Effects of computer-based testing on test performance and testing motivation. *Computers in Human Behavior*, **28** 5 1580–1586. <https://doi.org/10.1016/j.chb.2012.03.020>
- [12] Boevé, A. J., Meijer, R. R., Albers, C. J., Beetsma, Y., & Bosker, R. J. 2015. Introducing Computer-Based Testing in High- Stakes Exams in Higher Education : Results of a Field Experiment. *Plos One*, **10** 12 1–13. <https://doi.org/10.1371/journal.pone.0143616>
- [13] Puspendik. (2015). Artikel Mengenal PUSPENDIK. <http://www.litbang.kemdikbud.go.id>. Diakses pada tanggal 21 Oktober 2015 pukul 14.12 WIB
- [14] King, F.J., Goodson, L., & Rohani, F. 2015. *Higher-order Thinking Skill, Definition, Teaching Strategies, Assessment*. Center for Advancement of Learning and Assessment. Diakses 7 Oktober pada <http://www.cala.fsu.edu>
- [15] Retnawati, Heri 2018 *Desain Pembelajaran Matematika untuk Melatihkan Higher-order Thinking Skill*. (Yogyakarta: UNY PRESS)
- [16] Setiawati, W., Asmira, O., Ariyana, Y., Bestary, R., & Pudjiastuti, A. 2018 *Buku Penilaian Berorientasi Higher-order Thinking Skills*. (Jakarta: Direktorat Jenderal Guru dan Tenaga Kependidikan)
- [17] Dewi, C.N.P, & Rafi'udin, R. 2018. Perancangan Tampilan Aplikasi Ujian Berbasis Komputer untuk Ujian Harian Sekolah Menengah Atas. *ILKOM Jurnal Ilmiah*, **10** 3
- [18] Van der Kleij, F., Feskens, R., & Eggen, T. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: A meta-analysis. *Review of Educational Research*. **85** 4 475–511.

- [19] Barbera, E. 2009 Mutual feedback in e-portfolio assessment: An approach to the netfolio system. *British Journal of Educational Technology*, **40** 2 342-357.
- [20] Helsper, E., & Eynon, R. 2010 Digital natives: Where is the evidence?. *British Educational Research Journal*, **36** 3 503–520.
- [21] Nguyen, T., Charity, I., & Robson, A. 2016 Students' perceptions of computer-based learning environments, their attitude towards business statistics, and their academic achievement: Implications from a UK university. *Studies in Higher Education*, **41** 4 734-755.
- [22] Timmis, S., Broadfoot, P., Sutherland, R., & Oldfield, A. 2016 Rethinking assessment in a digital age: Opportunities, challenges and risks. *British Educational Research Journal*. **42** 3 454–476.
- [23] Evans, C. 2013 Making sense of assessment feedback in higher education. *Review of Educational Research*. **83** 1 70–120.
- [24] Helfaya, A. 2018 Assessing the use of computer-based assessment-feedback in teaching digital accountants. *Accounting Education*, 1-31. doi:10.1080/09639284.2018.1501716

### Acknowledgments

Thank you to the University of Lampung Research and Community Service Institute for funding this research through a professor's grant.