MODESTUM EURASIA Journal of Mathematics, Science and Technology Education, 2023, 19(10), em2339
ISSN:1305-8223 (online)
OPEN ACCESS Research Paper https://doi.org/10.29333/ejmste/13631

A phenomenological study of challenges that prospective mathematics teachers face in developing mathematical problems that require higher-order thinking skills

Jailani [1], Heri Retnawati [1*], Ibnu Rafi [1], Ali Mahmudi [1], Elly Arliani [1], Hutkemri Zulnaidi [2], Harris Shah Abd Hamid [3], Harun Joko Prayitno [4]

[1] Faculty of Mathematics and Natural Sciences, Yogyakarta State University, Special Region of Yogyakarta, INDONESIA
[2] Faculty of Education, University of Malaya, Kuala Lumpur, MALAYSIA
[3] Faculty of Management, Education and Humanities, University College MAIWP International, Kuala Lumpur, MALAYSIA
[4] Faculty of Teacher Training and Education, Muhammadiyah University of Surakarta, Central Java, INDONESIA

Received 13 February 2023 ▪ Accepted 28 July 2023

Abstract

Assessment in mathematics education in the 21st century should be more directed at higher-order thinking skills (HOTS). Not only teachers but also prospective mathematics teachers should be supported to design and carry out HOTS assessments. This qualitative study applying a phenomenology approach was mainly conducted to investigate the challenges that prospective mathematics teachers face in developing HOTS questions, including their strategies for dealing with these challenges. Our informants were 20 students of master's program in mathematics education. They were enrolled in the assessment of mathematics learning course, which was designed with a project-based learning model. Considering the learning model used, the purpose of this study then was extended to explore the benefits of implementing a project-based learning model in supporting student competence in developing HOTS questions. Data collection was carried out by (1) administering open-ended questionnaires; (2) observing the end product in the form of mathematics learning achievement tests and test blueprints; and (3) involving two experts who worked independently to judge the questions posed by students based on levels in the cognitive process dimension in the revised Bloom's taxonomy. Data collected from the open-ended questionnaire were analyzed qualitatively, while data from observation and judgment by experts on the end product yielded by students were analyzed descriptively. Our study supports previous studies, which demonstrate that the application of a project-based learning model that involves students to develop HOTS questions deepens students' knowledge of assessment. In developing HOTS questions, students struggled more in matching action verbs, item indicators, and test items to the level of cognitive process they defined, as well as constructing multiple-choice HOTS questions. The strategies that students took in dealing with challenges that arose in developing HOTS questions indicate that students regulate their learning.

Keywords: higher-order thinking skills, HOTS question, mathematics learning assessment, posing mathematics problem, project-based learning, teacher professional development

INTRODUCTION

It cannot be denied that teachers are one of the crucial elements in the implementation of education (Kaur, 2019; Kruszewska, 2021), especially in improving the quality of education, because they are the ones who actualize what the government wants for education that takes place in the country as suggested in the national curriculum. Teachers have a strategic role in implementing education through the implementation of

✉ jailani@uny.ac.id ✉ heri_retnawati@uny.ac.id **(*Correspondence)** ✉ ibnurafi789@gmail.com ✉ alimahmudi@uny.ac.id ✉ arlianielly@uny.ac.id ✉ hutkemri@um.edu.my ✉ drharris@ucmi.edu.my ✉ harun.prayitno@ums.ac.id

Contribution to the literature

- This study has revealed that prospective mathematics teachers still experience challenges in developing HOTS questions even though they have been facilitated to acquire and deepen their knowledge and skills about HOTS in the assessment of mathematics learning course.
- It is still a challenge for prospective mathematics teachers to develop HOTS questions based on the level of cognitive process and action verbs on indicators, especially when an action verb can represent two different levels of cognitive process in the revised Bloom's taxonomy.
- It is more challenging for prospective mathematics teachers to develop multiple-choice HOTS questions than essay or constructed-response types. Project-based learning model can be used in a course that focuses on preparing prospective mathematics teachers to become professional teachers in the future.

the learning process, starting from developing lesson plans, actualizing lesson plans to facilitate learning in the classroom, to conducting assessments that are intended to improve learning or assess student competencies. Teachers have their own responsibility in ensuring that what is planned and actualized through the learning process in class can support students in achieving the standard competencies set out in the curriculum document.

Regarding the strategic role that teachers have in providing good quality education and learning for students, previous studies have attempted to identify the essential knowledge or competencies that teachers must have to become competent teachers. Kyriacou (2007) has identified crucial competencies that teachers need to have, which he divides into three elements, namely knowledge, decision-making, and action. From these three elements, it can be further explained that a teacher

(1) must grasp adequate knowledge about the content he teaches, students, curriculum, models or strategies of learning, assessments, and factors that can influence student learning,

(2) can decide everything from before, during, and after learning that supports students to achieve the competencies that have been set, and

(3) behave well in optimizing student learning (Kyriacou, 2007).

What is mentioned by Kyriacou (2007) regarding essential knowledge of teacher basically refers to content knowledge (CK) (Shulman, 1986) or knowledge base (Shulman, 1987) of teacher, one of which is pedagogical content knowledge (PCK). PCK is related to the teacher's knowledge of facts, concepts, principles, and procedures in the content of the subject and how to organize them so that they can be easily learned by students with various characteristics including knowledge of learning difficulties or misconceptions that students may have and strategies to help students to overcome them (Mishra & Koehler, 2006; Shulman, 1986, 1987). Knowledge of students' misconceptions can be the main basis for teachers to design and carry out learning and conduct assessments (Park & Oliver, 2008). Therefore, PCK can also be associated with the teacher's knowledge of

(1) aspects of student learning that need to be assessed,

(2) appropriate procedures, instruments, techniques, or activities of assessment to use, and

(3) the advantages and disadvantages of using these assessment procedures, instruments, techniques, or activities (Magnusson et al., 1999; Park & Oliver, 2008).

Mathematics learning in the 21st century that is recommended to focus more on promoting higher-order thinking (HOT) (Brookhart, 2010; Richland & Simms, 2015) by implementing several learning strategies or models that have been proven to support this goal such as problem-based learning (Djidu et al., 2021; Jailani et al., 2017) and project-based learning (Holmes & Hwang, 2016; Suherman et al., 2020; Takiddin et al., 2020). Likewise, assessment in mathematics education is also expected to be more directed at the development of students' higher-order thinking skills (HOTS) (Radmehr & Vos, 2020). Based on a review of various literature, Schulz and FitzPatrick (2016) found that the term HOTS is often equated with other thinking skills such as critical thinking, creative thinking, and metacognitive thinking. This view of the similarity of the term HOTS with other thinking skills is used by Brookhart (2010) in defining HOTS. She argues that HOTS is a student's skill in

(1) applying the knowledge and skills he already has through involvement in learning into a context he has never thought of before,

(2) critical thinking, which includes reasoning, reflection, and making wise decisions, and

(3) solving problems whose solving strategies cannot be identified directly, including open-ended problems.

She added that the transfer of knowledge and skills that students have constructed into these new situations requires the skills of analyzing, evaluating, and creating. These three skills are the top three levels in the cognitive process dimension in the revised version of Bloom's taxonomy (Anderson & Krathwohl, 2001; Krathwohl, 2002).

To carry out an assessment that focuses on HOTS, the teacher must not only have sufficient knowledge about

HOTS and contents but must also understand how to measure or assess HOTS, including characteristics of the questions or problems that should be used. Knowledge related to this matter also needs to be owned and mastered by prospective teachers including experience in developing HOTS questions. Singer et al. (2013) suggested that not only teachers, but prospective teachers should also be given more opportunities and support to develop their competencies in developing measurement instruments. Previous studies have been carried out in relation to exploring the teachers' knowledge of HOTS and its assessment (e.g., Retnawati et al., 2018; Schulz & FitzPatrick, 2016; Thompson, 2008) and their experience in developing instrument to measure HOTS (e.g., Dahlan et al., 2020; Saepuzaman et al., 2022; Sinta et al., 2022; Wisrance & Semiun, 2020). Meanwhile, the abilities and challenges that prospective teachers, especially prospective mathematics teachers, face in developing HOTS questions are still understudied.

LITERATURE REVIEW

Higher-Order Thinking Skills

HOTS is still an interesting topic to discuss and study in mathematics education because it is one of the focuses in the implementation of mathematics learning and assessment of mathematics learning worldwide (Brookhart, 2010; Drijvers et al., 2019; Forster, 2004; Liu et al., 2022; Radmehr & Vos, 2020; Richland & Simms, 2015). HOTS has been associated with other types of thinking skills, such as creative thinking, critical thinking, reasoning, decision-making, and metacognitive (Apino & Retnawati, 2017; Brookhart, 2010; Evidiasari et al., 2019; Hmelo & Ferrari, 1997; Ijirana et al., 2021; Lewis & Smith, 1993; Newmann, 1991; Zhou et al., 2023). This association is based on the conception that these thinking skills all require skills in interpreting, analyzing, or manipulating knowledge or information; not just skills in applying prior knowledge (Newmann, 1991). Lewis and Smith (1993) suggested that HOTS is thinking skills used to achieve a goal or to deal with a confusing situation or context by interrelating and expanding existing knowledge and new knowledge acquired. They added that the goals that can be achieved with HOTS can be deriving decisions about what to believe or do, creating new things or ideas, and solving non-routine problems. By referring to the two-dimensional framework in revised Bloom's taxonomy, HOTS is interpreted as skills in analyzing, evaluating, and creating when confronted with three of the four categories of knowledge in the knowledge dimension, namely conceptual, procedural, and metacognitive knowledge (Anderson & Krathwohl, 2001; Retnawati et al., 2018). Skills in analyzing refer to skills in sorting out ideas from an overall idea based on their relevance, identifying coherent interrelations among relevant ideas, and determining values, biases that may exist, conclusions, or points of view from the interrelations of relevant ideas. Meanwhile, skills in evaluating represent skills in judging or assessing a given context based on certain criteria (e.g., quality and consistency) and quantitative or qualitative standards. Finally, skills in creating are related to skills in providing various alternative solutions to problems or deriving hypotheses, developing plans to solve problems, and producing a product with certain specifications (Anderson & Krathwohl, 2001).

Students need to have good HOTS because it is critical to grasp in-depth conceptual and disciplinary understanding (Heron & Palfreyman, 2021) and to enable students to use the prior knowledge they have constructed and the understanding they already have to obtain reasonable responses or solutions to the problems or contexts they face (Liu et al., 2022). In addition, Liu et al. (2022) also emphasize that it is important for students to develop their HOTS because these skills are needed so they can be creative. Creativity is one of the keys to success in all fields and in the midst of rapid development (Liu et al., 2022) and trigger someone to be able to find new ideas or make discoveries that might change the world in a better direction (Marczewska et al., 2023). Thus, we argue that it is impossible for a student's creativity to develop optimally if the student is not given the opportunity to engage in mathematical activities and tasks that require HOTS. By using the same idea, we support what Van den Berg (2004) and Radmehr dan Vos (2020) said that in order for students' HOTS to develop, students should be facilitated and guided to engage in activities and tasks that deliberately promote HOTS. Before engaging in such activities and tasks, students need to be facilitated first to have good basic knowledge and understanding in addition to having the ability to recall of information. In addition, efforts to develop student HOTS also need to be accompanied by the creation of a thinking environment that allows students to receive the necessary support, feel confident in their abilities, and address error-making as a learning process and an opportunity to learn (Van den Berg, 2004). The provision of HOTS development facilities needs to be supported by the spirit that all students have the right and are able to access these facilities (Lewis & Smith, 1993; Newmann, 1991; Van den Berg, 2004; Zohar & Dori, 2003) because all students are able to face the challenges that require them to interpret, analyze and manipulate the knowledge they already have (Newmann, 1991).

Because the topic of HOTS is very interesting to discuss, extensive studies have been conducted to explore the topic with various focus of interest. Through bibliometric analysis using metadata of journal articles from Web of Science database from 1984 to 2020, Liu et al. (2022) have found that "higher-order thinking ability, instruction of higher-order thinking, the curriculum and sections of higher-order thinking, and higher-order thinking learning" (p. 635) became the top focus of interest on HOTS topic in recent years. Their findings suggest that many studies on HOTS have explored the skills associated with HOTS; factors that may contribute to the development of HOTS; learning practices that promote HOTS through the application of certain learning models, methods or strategies; and HOTS-oriented learning in various disciplines and levels of education. Thus, issues regarding the competence of teachers or prospective teachers in designing learning, implementing learning, designing learning assessments, and carrying out learning assessments that focus on promoting HOTS still need to be studied more. These issues are very important to be explored because it cannot be denied that the provision of learning facilities for students to develop HOTS requires adequate mastery of knowledge and competence about HOTS by teachers. Several studies have been conducted to investigate teachers' and prospective teachers' understanding of HOTS and its learning strategy (e.g., Barak & Shakhman, 2008; Retnawati et al., 2018; Rianasari & Apriani, 2019; Schulz & FitzPatrick, 2016; Thompson, 2008) and their competency in solving HOTS problems (e.g., Retnawati et al., 2018), developing lesson plan oriented to foster HOTS (e.g., Sa'adah & Anjarwati, 2022), and developing HOTS instruments (e.g., Machromah et al., 2019; Rianasari & Apriani, 2019; Saepuzaman et al., 2022). However, studies that focus on the competence of prospective teachers, especially in the field of mathematics, in designing learning and assessments that are intentionally oriented towards promoting HOTS are still understudied.

Understanding and Competence of Teachers and Prospective Teachers in Developing Tests

The learning process involves planning learning activities, actualizing learning plans in the classroom, and conducting learning assessments. Therefore, teachers at least need to have competencies in these three areas of learning process, as well as prospective teachers who need to be prepared to master these competencies through their teacher education program. By paying attention to the three main activities that teachers need to do in the learning process, facilitating student HOTS development should not only be done through designing learning and actualizing learning that involves students in activities that deliberately promote their HOTS, but should also be done through assessment. Assessment is not only useful to find out the extent of knowledge and competencies that students have mastered (Nitko & Brookhart, 2011; Reynolds et al., 2010), but also as a means for students to evaluate and reflect on their thinking constraints and what they still need to learn more and as a means for teachers to develop students' HOTS (Van den Berg, 2004). The extent to which the assessment that the teacher designs and implements will be able to cultivate HOTS depends on the type, technique, and method of assessment as well as the tasks or problems presented in the assessment. Thus, teachers and prospective teachers need to understand these things, including being competent in developing tests to measure and train students' HOTS, as part of their PCK development (Rafi & Sugiman, 2019).

Test development, especially for summative assessment purposes, is frequently carried out through five main steps, namely conceptualizing the test, constructing the test, conducting test tryout, analyzing data from the test tryout results, and revising based on the results of the analysis (Cohen & Swerdlik, 2018; Retnawati, 2016). Test conceptualization is a step, where teachers as test developers identify the specifications of the tests they want to develop. The test specifications that the teacher can set at least include the objective of the test, the construct variable that is the focus of the test, which is derived based on the results of the theoretical review, the target of the test, the content included in the test, how to administer the test, the format of the test items to be used (selected- or constructed-response), test item indicators, and scoring for each test item and the entire test item based on student responses. The test specifications are presented in a table, called a blueprint, which then becomes a guide for the teacher in constructing the content of the test. In developing tests that focus on measuring and promoting HOTS, it is common to use the action verbs suggested in the revised Bloom's taxonomy (see Anderson & Krathwohl, 2001; Newton et al., 2020; Retnawati, 2016; Sideeg, 2016) in constructing test item indicators. These action verbs reflect the level or complexity of thinking that students need to demonstrate to complete a test item.

Several studies have investigated teachers' or prospective teachers' understanding of HOTS assessment and teacher competence in developing tests to assess HOTS. Retnawati et al.'s (2018) study have demonstrated that mathematics teachers at the junior high school level have understood various methods or techniques, including problem or task characteristics, which they can use to measure and assess HOTS. However, their study did not find teachers' responses that connected HOTS assessments with the revised Bloom's taxonomy. Furthermore, Schulz and FitzPatrick's (2016) study revealed that even though all science and social studies teachers who were subjects of their study had heard of the (revised) Bloom's taxonomy, all teachers admitted that they had not understood and been taught to apply it in conducting assessments. Most teachers did not even understand the terms analyze, evaluate, and create, which are often associated with HOTS. Regarding the competence of teachers or prospective teachers in developing HOTS questions, Rianasari and Apriani (2019) have found that most prospective mathematics teachers have not been able to make HOTS questions correctly, where the questions they have made only require thinking skills at the third level in the domain of cognitive processes in the revised Bloom's taxonomy. What was found by Rianasari and Apriani's (2019) study was also found in Sinta et al.'s (2022) study, where when developing tests that focused on HOTS assessments, many test item indicators that mathematics teachers construct did not reflect and were not appropriate with the levels of cognitive process in the revised Bloom's taxonomy. In addition, most of the test items they constructed required only cognitive process at the bottom three levels of the cognitive process dimension in the revised Bloom's taxonomy (i.e., remember, understand, and apply).

What we have reported based on the aforementioned previous studies indicates that teachers and prospective teachers need to be facilitated and supported to have an in-depth understanding of HOTS, including teaching strategies and assessments that focus on HOTS, and competent at it all. For prospective teachers, such facilities and support can be realized by creating a learning environment that integrates both theory and practice to enable them to master a concept theoretically and can apply it in practice confidently and effectively (Radović et al., 2021; Wrenn & Wrenn, 2009). Such learning environment is imperative, especially in professional degree programs such as the teacher education program, which is expected to facilitate prospective teacher students to master education-related concepts (e.g., pedagogy, learning, and assessment) and content in their area of expertise and later apply it to facilitate their students' learning in the classroom. One way to create a learning environment that integrates theory and practice is to apply certain learning models, one of which is project-based learning (Almulla, 2020; Bell, 2010; Blumenfeld et al., 1991; Miller & Krajcik, 2019; Nikolaeva, 2012). This learning model has been applied to facilitate student learning across educational levels, starting from elementary education to higher education, and in wide range of disciplines (Al-Busaidi & Al-Seyabi, 2021). Project-based learning is a student-centered learning model, where students are engaged in a series of investigative or inquiry activities that requires them to construct, apply, and integrate their knowledge to solve authentic or nontrivial problems (Bell, 2010; Guo et al., 2020). Blumenfeld et al. (1991) explained in more detail that students can solve problems that are authentic or nontrivial by doing the following activities.

> "… asking and refining questions, debating ideas, making predictions, designing plans and/or experiments, collecting and analyzing data, drawing conclusions, communicating their ideas and findings to others, asking new questions, and creating artifacts" (p. 371).

Although previous studies (e.g., Retnawati et al., 2018; Rianasari & Apriani, 2019; Schulz & FitzPatrick, 2016; Sinta et al., 2022) have succeeded in identifying teacher or prospective teacher obstacles in developing HOTS questions, these studies have not explained with certainty whether teachers and prospective teachers have been given the facilities to develop HOTS questions. The challenges that prospective teachers face in developing HOTS questions when they have been facilitated to understand HOTS and develop HOTS assessments, especially through the implementation of project-based learning, have unfortunately not been extensively explored.

Research Questions

The current study mainly seeks to reveal the various challenges experienced by students who were projected to become mathematics teachers in developing HOTS questions from their involvement in the assessment of mathematics learning course designed by applying project-based learning model. Their involvement in the course is important in this study so that students' challenges in developing HOTS questions were not necessarily caused by the unavailability of facilities to develop these competencies. Considering the setting and main objectives to be achieved, there are four research questions (RQ) that we address in this study as follows.

RQ1. How do students' perspectives on the implementation of the project-based learning model in the assessment of mathematics learning course relate to their competence in developing a mathematics learning achievement test that contains HOTS questions?

RQ2. What challenges or difficulties did students experience in developing a mathematics learning achievement test that contains HOTS questions based on their opinions?

RQ3. What strategies have students used to overcome challenges or difficulties in developing a mathematics learning achievement test that contains HOTS questions?

RQ4. What challenges or difficulties did students experience in developing a mathematics learning achievement test that contains HOTS questions based on their end product?

METHODS

Design, Participants, and Context of the Study

This is a qualitative study that applied a phenomenological approach to explore the challenges faced by prospective mathematics teachers when developing mathematical problems that require HOTS and their strategies to address these challenges. This qualitative study enabled us to understand in detail the issues, phenomena, or contexts experienced by the participants in our study that are complex and difficult (even impossible) to measure (Creswell, 2013). We chose to use a phenomenological approach because we thought that it fits our study objectives better than the other four approaches to qualitative study (see Creswell, 2013). The phenomenological approach provides a means of depicting and obtaining the essence of the phenomena that groups of individuals directly experience and how they experience and make sense of these phenomena (Creswell, 2013; Teherani et al., 2015). The phenomenon in this study was that students who took the assessment of mathematics learning course were facing a number of challenges in developing tests that contain HOTS questions. In fact, in the course held by applying the project-based learning model, students have been facilitated to understand HOTS and implement this understanding in developing tests.

A total of 20 (i.e., seven males and 13 females) prospective mathematics teachers participated in this study. They were second semester students of the master's program in mathematics education at a public university who enrolled in the assessment of mathematics learning course (two credits) in the same class. The public university is one of the educational personnel education institutes (*lembaga pendidikan tenaga kependidikan* [LPTK]) in Indonesia. The assessment of mathematics learning course is a compulsory course that lasted 16 meetings (one meeting per week). There are no prerequisites for the course. This course is focused on facilitating students to learn about learning assessment and its application in learning, especially mathematics learning, including utilizing the results of assessment or evaluation. Topics that students learn in this course include tests, measurements, assessments, evaluations, validity evidence and reliability estimation, qualitative and quantitative item analysis, as well as the development of instruments for assessment/evaluation of learning, improvement of mathematics learning, or mathematics education research. In addition to being held classically, in that course, students were also facilitated to deepen their understanding of assessment of mathematics learning, especially in developing measuring instruments in the form of tests, through a project-based learning model. Therefore, by considering the learning model that was applied in the course, in addition to exploring the challenges and strategies to deal with these challenges in relation to developing measurement instruments, we also explored students' views on the implementation of the project-based learning model and the project they worked on.

Students of the master's program in mathematics education at the university, where this study was conducted are students who have completed undergraduate programs in mathematics, mathematics, or statistics. In the case of this study, most of the students in the mathematics education master's program were graduates of the undergraduate program in mathematics education. In this study, however, the background of the participant's major in undergraduate education was not identified. Students who are graduates from an undergraduate program in mathematics or statistics are required to take matriculation courses in the first two semesters. Matriculation courses in the first semester consist of mathematics curriculum and learning (two credits) and mathematics learning strategies (two credits). The matriculation courses in the second semester consist of mathematics learning planning (two credits) and development and production of mathematics learning media (two credits). Graduates of the master's program in mathematics education are expected to master competence in terms of attitude, knowledge, specific skills, and general skills. This competence can be obtained through learning that focuses on four main study materials: PCK, CK, technological pedagogical knowledge (TPK), and research in the field of TPK.

Data Collection

Data on students' views on their engagement in the project-based learning model in the assessment of mathematics learning course, the challenges that students faced in developing mathematical problems that require HOTS to solve the problems, and strategies to overcome the challenges that they faced were collected using document analysis and an open-ended questionnaire. The document referred to here is in the form of student work resulting from their involvement in a project designed on the assessment of mathematics learning course, which was also used as an end product in the course. The project given to students was to develop an instrument in the form of a test for measuring the achievement of junior high school (grade 7 to 8) or senior high school (grade 10 to 11) students in a pair of basic competencies on the knowledge and skills dimensions in mathematics (see Ministry of Education and Culture, 2018), which also reflect the content or topic. Topics include perimeter and area of rectangles, surface areas and volumes of polyhedrons, circles, triangles and quadrilaterals, social arithmetic, composition and inverse functions, angles relationship formed by transversal and parallel lines, trigonometry ratios, laws of sines and cosines, graphs of trigonometric functions, derivatives of algebraic functions, statistics, and vectors. Each student was asked to determine basic competencies at either the junior high school or senior high school level with the condition that no one chooses the same basic competencies and that the number did not differ too much between those choosing the junior high school level and those choosing the senior high school level.

The test that students have to develop must consist of 15 items in total, which include 12 multiple-choice and three constructed-response items. Considering that constructed-response items and essay items refer to the same question type, in this paper we use the two terms interchangeably. As an integral part of a test development, students were also required to develop a test blueprint, which is expected to contain information related to a pair of basic competencies on the dimensions of knowledge and skills selected according to the core and basic competencies in the national curriculum document (Ministry of Education and Culture, 2018), competency achievement indicators, question indicators, type of test item, level of thinking skills on the dimension of cognitive process in the revised Bloom's taxonomy, item number, and scoring rubric. Although students were allowed to develop test items with levels of cognitive process dimension that range from the lowest (C1, remember) to the highest (C6, create), they were strongly encouraged to develop tests with most of the items requiring cognitive processes at the C3 (apply) level or higher.

The open-ended questionnaire used in this study contains three questions. The first question is focused to explore students' views on the impact of implementing project-based learning models in supporting them to develop learning achievement tests. The first question is "what experience did you get from your involvement in the assessment of mathematics learning course, which required you to work on a project in the form of developing a test that contained HOTS questions for junior or senior high school levels?" The second question is focused to explore the challenges students experienced in developing tests that are also address the challenges of developing HOTS questions. The second question is "what obstacles did you experience while carrying out the project of developing a test that contained HOTS questions for junior high or high school level?" The third is focused to explore the strategies that students used to overcome challenges in developing tests and HOTS questions. The third question is "what strategies did you put in place to overcome the obstacles you faced so you could complete the project?".

The questionnaire was administered to students at the end of the 16th week or meeting. We have informed students that all the responses they provided to the questionnaire are only intended for study in the context of developing the quality of lectures, would not affect the grade or score that students would obtain in the assessment of mathematics learning course, and would not affect their future life; and students gave their consent. In addition, to maintain the privacy of students, we also coded their identities (student 1 [S1], student 2 [S2], …, student 20 [S20]) when presenting the results of our study.

Data Analysis

Data analyses performed in this study consisted of analyzes of students' responses to the open-ended questionnaire and the results of students' work in developing learning achievement tests. Because the data collected from the open-ended questionnaire is qualitative in nature, the data was analyzed qualitatively starting from data reduction to obtain codes, sub-themes or interrelations of codes that demonstrate similarities, and themes or interrelations of sub-themes (Bogdan & Biklen, 2007). The results of this analysis are presented in three tables, which have been adjusted to the number of main issues explored through the open-ended questionnaire.

The analysis of student work in developing learning achievement tests was carried out by first identifying the test items that based on the test blueprint document that students made these test items require cognitive processes at the level of C4 (analyze), C5 (evaluate), or C6 (create). The results of this identification were contained in a spreadsheet document along with information on the name of the student who made the test item and the test item number. This spreadsheet document was then equipped with columns to fill in the main information used to investigate the challenges faced by students in developing test items that require HOTS. The main information was provided by two experts or raters who worked independently, where the two experts have earned their master's degree in mathematics education. To fill in the main information in the spreadsheet document, the two raters were given access to the work results document of all the students participating in this study in the form of a test blueprint, which already contained the developed test items.

The main information that raters need to provide includes two aspects of the assessment, namely the match of the indicators with the level of cognitive processes needed to solve the test items obtained from the results of previous identification and the match of the test items with the test item indicators. For the first aspect of the assessment, the raters were asked to assign code 1 for each test item that there is a match between the indicator and level of cognitive processes and to assign code 0 when such a condition is not reached (percent agreement=67%). Such an assessment method was also applied to the second aspect of the assessment (percent agreement=71%). In addition to assigning a code of 1 or 0, raters were also required to provide additional comments for each case they assigned a code of 0, which means the desired match between indicators and levels of cognitive process or between test items and indicators is not satisfied. Additional comments on the first aspect of the assessment were focused on identifying which level of cognitive process is best represented by the existing test item indicator. For additional comments on the second aspect of the assessment, it focused more on identifying possible conceptual errors contained in the test items, insufficient information available to solve the test items, and the inappropriateness or illogicality of the context used.

Ensuring the Quality of the Study

Several strategies were undertaken to ensure the quality of this study in relation to the results obtained and interpretations based on one of the four trustworthiness criteria in qualitative studies (Lincoln & Guba, 1985; Stahl & King, 2020), namely credibility. Ensuring the credibility of the results of this study and their interpretation were provided by means of prolonged engagement and method triangulation techniques. Because two of the authors of this article were lecturers in the assessment of mathematics learning course that dealt directly with the participants of this study, the results from the analysis of the data gathered from the open-ended questionnaire could be confirmed through observation by lecturers during learning activities. Nevertheless, to avoid the possibility of bias that arises in student responses to the open-ended questionnaire, we informed students that all the responses they gave to the questionnaire did not affect their grade point on the assessment of mathematics learning course. Apart from the two authors who were lecturers in the course, the other authors of the present article were also involved in ensuring and confirming data reduction, themes, interrelations of themes, conclusions, and interpretations.

The credibility of this study was also ensured through the technique of method triangulation, where the data used to investigate student challenges in developing HOTS questions were collected from open-ended questionnaires and investigations on learning achievement measurement instruments in the form of tests that students produced. The later was used to support the earlier and vice versa. In addition, to ensure the credibility of the results related to students' challenges in developing HOTS questions based on the products produced by students, as we mentioned earlier, we recruited two experts who worked independently to provide their judgment on the student's work. Furthermore, we also strengthen the credibility of the results of this study by providing direct quotations from the responses of several participants to the three open-ended questions in the questionnaire. The use of direct quotations allows readers to assess the accuracy of our data analysis, interpretation of results, and drawing conclusions based on the original data, thereby strengthening what our study reports (Corden & Sainsbury, 2006).

RESULTS

This study strives to delve into the challenges experienced by students who enrolled in assessment of mathematics learning course in developing HOTS questions included in the mathematics learning achievement test for secondary education level. To anticipate challenges that arose more due to insufficient facilities received by students to develop HOTS questions, learning activities in the assessment of mathematics learning course were designed by following a project-based learning model. Therefore, in addition to presenting results related to these challenges, in this section we also present results related to the benefits felt by students from their involvement in lectures that were designed with a project-based learning model so that they succeed in producing instruments in the form of tests to measure learning achievement in mathematics and the strategies students took to overcome the challenges that arose.

Benefits that Students Experience When Engaged in Project-Based Learning in the Assessment of Mathematics Learning Course

Because the assessment of mathematics learning course was designed according to a project-based learning model, in practice, of course, at the end of this course students would be successful in producing an artifact or end product. The end product is an instrument in the form of a test to measure learning achievement in mathematics in which some of test items are expected to be categorized as HOTS questions, and the test blueprint for secondary education (i.e., junior high or senior high school level). Nonetheless, the lectures were still centered on the process of producing the end product, namely providing support to students to gain an in-depth understanding of central topics in mathematics learning assessment and build skills related to these central topics. Following are some of the responses that students gave regarding the experiences they gained from attending the assessment of mathematics learning course.

> "…, this course project also trained me to develop HOTS questions. Despite experiencing various obstacles, everything that has been passed through this course project would shape my competence in the field of assessment" (S9).

> "It is not easy to make HOTS questions. More references are needed to make the question. However, by engaging in project-based learning, it adds to my experience as well as training for me in the future as a teacher in making good quality assessments" (S11).

> "Through this project-based learning … I better understand the assessment process, starting from determining the item indicators based on the selected basic competencies, constructing test items, ensuring the suitability of the item with the cognitive level, to the item validation and revision of the items that have been made. I also learn to construct multiple choice items by paying attention to the quality of the distractor, while paying attention to the possibilities of student responses when constructing essay items … Based

Table 1. Benefits of project-based learning model

Codes	Sub-themes	Theme
Becoming familiar with different types of assessment and have experience developing mathematical problems that require HOTS	Project-based learning provides learning opportunities for students to acquire knowledge about assessment, including developing measurement instruments and analyzing their characteristics	Project-based learning not only supports students to gain knowledge about assessment, but also provides opportunities for them to apply the knowledge they have acquired, thereby gaining experience in developing various forms of assessment including developing test items that require HOTS
Becoming familiar with strategies in developing HOTS questions based on the level of cognitive processes		
Acquiring new knowledge about the development of HOTS questions		
Acquiring knowledge about test development to analysis of test characteristics and applying them		
Acquiring knowledge about assessment		
Acquiring knowledge about different types of assessment and test item		
Gaining more experience and opportunities to practice in developing good test item for assessment purposes	Project-based learning provides learning opportunities by practicing directly to develop HOTS questions and analyze their characteristics	
Gaining experience in developing instruments for assessment and administering them		
Gaining experience in developing instruments for assessment and analyzing their characteristics		
Gaining experience in developing multiple-choice items that require HOTS to solve them		
Gaining experience in developing constructed response items that require HOTS to solve them		

on the learning process and understanding this, it can be my provision to develop competence in developing good HOTS questions" (S16).

The results of the analysis of student responses to the open-ended questionnaire suggest that the learning activities facilitated in the assessment of mathematics learning course have reflected the characteristics of the learning model used, i.e., the project-based learning model, as well as demonstrating the benefits gained from implementing the learning model (see **Table 1**). According to student responses, the implementation of the project-based learning model has not only promoted their understanding of central topics in learning assessments in general and specifically in mathematics learning but also increased their skills to develop instruments as a means to deepen their understanding of learning assessments. The understanding and skills they acquired support them in developing test blueprints and test items with certain specifications as the end product of the assessment of mathematics learning course. Even though the end product of the course was in the form of a test blueprint and test items, the understanding and skills that students gained during their involvement in the course did not only extend to the topic of instrument development but also to guaranteeing the quality of the instrument. If it is related to the focus of the study on students' ability to make questions that require HOTS, the implementation of lectures using a project-based learning model was considered by students to provide their own benefits. Students found that existing lectures with the applied learning model have supported their acquisition of understanding of the characteristics of HOTS questions based on the level of cognitive process according to the revised version of Bloom's taxonomy and strategies for developing HOTS questions in the form of multiple-choice items.

Difficulties Experienced by Students in Developing Learning Achievement Tests that Contain Items That Require HOTS to Solve Them

This study has revealed that students gained benefits from their involvement in lectures that were designed by applying a project-based learning model, where they succeed in producing a set of learning achievement tests accompanied by the acquisition of skills and understanding of central topics in the assessment of mathematics learning. However, in the process of producing this end product students encountered several challenges regarding test development in general and specifically in the development of HOTS questions (see **Table 2**). As an integral part of developing a test instrument, students required to develop a test blueprint as a guide in constructing test items. The test blueprint at least contains information about the question indicators and the level of cognitive process in the revised Bloom's taxonomy needed to solve the problem. The item indicators were required to use action verbs suggested in the revised Bloom's taxonomy. The following are some of the responses that students expressed regarding their obstacles in developing HOTS questions.

> "… I have difficulties in making multiple choice questions that require cognitive process at the C6 level" (S8).

> "I find it difficult to make questions with levels C1 to C6 because there are action verbs that are not only at one level of cognitive process" (S13).

> "Constructing test items based on indicators listed is still the main obstacle for me (in developing HOTS questions). Sometimes the item that I

Table 2. Difficulties that students have in developing a test that contains HOTS questions

Codes	Sub-themes	Theme
Experiencing difficulties in developing HOTS questions Experiencing difficulties in developing multiple-choice test items that require HOTS to solve them Experiencing difficulties in developing test items based on the revised Bloom's taxonomy action verbs used in the indicators Do not understand how to develop test and its items Experiencing difficulties in developing HOTS questions Experiencing difficulties in developing HOTS questions that match the selected content and level of cognitive process according to the revised version of Bloom's taxonomy Experiencing difficulties in developing HOTS multiple-choice questions that require cognitive process at level of C5 (evaluate) or C6 (create) Experiencing difficulties in constructing multiple-choice item distractor	Students experience difficulties in developing HOTS questions based on the level of cognitive process, revised version of Bloom's taxonomy action verbs, and writing effective distractors for the case of developing multiple-choice items	Difficulties experienced by students in developing tests with test items requiring HOTS include difficulties in developing questions that match levels of cognitive process or action verbs of revised Bloom's taxonomy, writing effective distractors for multiple choice item, & distinguishing & classifying cognitive process levels of revised Bloom's taxonomy. As a result of these difficulties, students became doubtful about HOTS questions they developed, whether questions matched level of cognitive process that had been specified, especially for HOTS questions
Experiencing difficulties in classifying level of cognitive process of the test item Experiencing difficulties in distinguishing between the level of C4 (analyze) and C5 (evaluate) Experiencing difficulties in determining the level of cognitive process that best fit to represent the related test item	Students experience difficulties in distinguishing & classifying levels of cognitive processes in revised Bloom's taxonomy	
Experiencing doubts with the levels of cognitive process of the test items developed Experiencing doubts with the test items developed	Students experience doubts about the HOTS questions developed because there are doubts about whether the questions match at the level of cognitive processes specified	

> construct does not match the indicator. Because of that, I sometimes make items based on existing indicators and sometimes I construct items first and then the indicators" (S14).

> "In making HOTS questions, I experienced problems in constructing test questions that fit cognitive processes, from C1 to C6 levels" (S18).

From the results of the analysis of students' responses to the open-ended questionnaire, it was found that matching the test items with the appropriate item indicators was challenging for students, especially when the test items had to match the action verbs used in the item indicators and the level of cognitive process that has been determined. In addition, it was challenging for students when constructing test items that require cognitive process at the levels of C4 (analyze), C5 (evaluate), and C6 (create) due to insufficient understanding of the essential differences between the three levels of cognitive process in the revised Bloom's taxonomy. As a consequence, when they have successfully constructed a test item, they felt unsure about the test item, whether it was in accordance with the level of cognitive process that they have set, or it was better for them to match the level of cognitive process to the test item they have constructed. Furthermore, when dealing with the construction of multiple-choice test items, students experienced difficulties in constructing effective distractors and constructing questions that require HOTS.

Strategies That Students Did to Overcome Difficulties Encountered in Developing Learning Achievement Tests That Contain HOTS Questions

It has been revealed that there were a number of challenges faced by students when they were involved in lectures that apply a project-based learning model, where the end product produced from this lecture is a learning achievement test in mathematics, which is expected to be dominated by HOTS questions. The responses of some students when asked to describe the strategies they have used to overcome obstacles in developing HOTS questions are, as follows.

Table 3. Strategies to overcome challenges in developing a learning achievement test that contains HOTS questions

Codes	Sub-themes	Theme
Sharing developed test with the peers to inquire and receive any feedback or suggestions for improvement Sharing developed test with the peers and receive feedback from them and suggestions for improvement from lecturers or supervisors	Sharing draft of end product with peers & lecturers or supervisors to collect feedback or suggestions for improvement or making some revisions	Revisiting knowledge about instrument development and other relevant knowledge obtained during lectures, sharing problems in instrument development with peers and lecturers, and paying attention to related references
Utilizing the knowledge and experience gained during lecture activities Following and fulfilling test and test items development procedures as well as providing validity evidence	Revisiting knowledge or understanding of test development procedures & related skills acquired during involvement in lecture activities	
Reading reference books and reviewing examples of (good) HOTS questions	Reviewing related reference books or literature including examples of (good) HOTS questions	

Table 4. Characteristics of test items investigated

Item type	C4 item, n (%)	C5 item, n (%)	C6 item, n (%)
Multiple-choice (selected-response)	48 (32.00)	29 (19.33)	18 (12.00)
Essay (constructed-response)	13 (8.67)	20 (13.33)	22 (14.67)
Total	61 (40.67)	49 (32.67)	40 (26.67)

> "I studied more examples of HOTS questions with different levels of cognitive (process) to understand in depth the characteristics of each level (of cognitive process)" (S2).

> "I ask for suggestions or recommendations from validators, peers, and lecturers. In addition, I looked at HOTS questions available on the internet to get more insights on them" (S14).

> "I was greatly assisted by my classmates who helped me validate the test items that I developed so that if there were errors I could fix them immediately. My classmates were also willing to provide suggestions or recommendations, which I have found very useful for me to produce better test items" (S19).

Table 3 presents the strategies that students have used to overcome the challenges they experienced when developing an instrument for measuring learning achievement in mathematics.

Even though the project given to students is individual in nature, during the instrument development process students were allowed to consult with peers and lecturers in the assessment of mathematics learning course. This was used as one of the strategies carried out by students when they faced challenges during the process of completing the end product of the course. In addition, when facing obstacles in developing the required product, they reviewed the knowledge and skills acquired during learning activities in lectures, read related literature or reference books, and paid attention to examples of HOTS questions that are considered to be of good quality.

Challenges Experienced by Students in Constructing Test Items That Require HOTS

This section focuses on presenting the challenges that students have in developing HOTS questions based on each student's end product. These challenges were identified based on the judgment given by two experts working independently. Because each student was required to construct 15 test items consisting of 12 multiple-choice items and three essay items, with 20 students participating in this study, there were a total of 300 items (i.e., 240 multiple-choice items and 60 essay items). Of the total 300 items, 150 items were further investigated by the experts because these items were determined by students as items that required cognitive processing at the level of C4, C5, or C6 (so-called HOTS questions). One hundred and fifty items were obtained from each student who constructed five to 11 items (M=7.50, SD=1.76). These items consisted of 95 (63.33%) multiple-choice items and 55 (36.67%) essay items. The characteristics of the items investigated further in this study are presented in **Table 4**. **Table 4** demonstrates that the level of cognitive process is inversely proportional to the number of test items proposed by students regardless of the type of test items. Furthermore, at levels C4 and C5, students proposed

Table 5. Distribution of expert judgment on items proposed by students

Expert judgment	Suitability of item indicator with level of cognitive process			Suitability of item with its indicator		
	C4 item, n (%)	C5 item, n (%)	C6 item, n (%)	C4 item, n (%)	C5 item, n (%)	C6 item, n (%)
Both experts assign code 1	10 (16.39)	21 (42.86)	4 (10.00)	41 (67.21)	33 (67.35)	24 (60.00)
Either expert 1 or expert 2 assigns code 0	20 (32.79)	13 (26.53)	16 (40.00)	17 (27.87)	11 (22.45)	15 (37.50)
Both experts assign code 0	31 (50.82)	15 (30.61)	20 (50.00)	3 (4.92)	5 (10.20)	1 (2.50)

Table 6. Number of HOTS items constructed by students based on expert judgement

Student	Expert 1, n (%)	Expert 2, n (%)
Student 1 (S1)	1 (12.50)	0 (0.00)
Student 2 (S2)	0 (0.00)	1 (20.00)
Student 3 (S3)	2 (28.57)	4 (57.14)
Student 4 (S4)	7 (77.78)	7 (77.78)
Student 5 (S5)	3 (50.00)	1 (16.67)
Student 6 (S6)	3 (33.33)	3 (33.33)
Student 7 (S7)	3 (50.00)	2 (33.33)
Student 8 (S8)	6 (66.67)	5 (55.56)
Student 9 (S9)	4 (44.44)	4 (44.44)
Student 10 (S10)	3 (37.50)	2 (37.50)
Student 11 (S11)	6 (100)	2 (33.33)
Student 12 (S12)	8 (72.73)	4 (36.36)
Student 13 (S13)	3 (33.33)	6 (66.67)
Student 14 (S14)	2 (40.00)	1 (20.00)
Student 15 (S15)	3 (50.00)	0 (0.00)
Student 16 (S16)	1 (11.11)	6 (66.67)
Student 17 (S17)	3 (37.50)	4 (50.00)
Student 18 (S18)	0 (0.00)	0 (0.00)
Student 19 (S19)	2 (22.22)	1 (11.11)
Student 20 (S20)	4 (66.67)	2 (33.33)

more multiple-choice test items, while at level C6 they proposed more essay type test items.

We have mentioned earlier that the students' challenges in developing tests that contain HOTS questions investigated from the end products that students produce were judged based on two aspects. The expert assigns code 1, which represents the fulfillment of that aspect or code 0, which indicates the aspect is not fulfilled. **Table 5** demonstrates the results regarding the judgment of experts on these two aspects. In the aspect of suitability between the item indicators and the level of cognitive process, it was found that students experienced difficulties in that aspect, where the number of items developed by students who were coded 0 by one or both experts was more than the number of items coded 1 by the two experts. This finding indicates that even though students have used action verbs according to each level of the revised Bloom's taxonomy, because some action verbs correspond to two or even three levels, students have difficulty differentiating the use of such action verbs appropriately in the test item indicators, which ultimately also have an impact on the test items they construct. Furthermore, in terms of the second aspect, it was revealed that the challenges students faced in constructing test items that matched the indicators were not as big as the challenges they faced in constructing indicators that matched the level of cognitive process, where more than 60% of the items have been justified to satisfy the second aspect by the two experts (see **Table 5**).

This study has demonstrated that the big challenge faced by students when constructing test items that require HOTS was making the indicators and levels in the revised Bloom's taxonomy compatible with each other. This challenge hinders students from constructing test items according to the level of cognitive process they want, especially at level C4 or higher. **Table 6** presents the number of test items that each expert could categorize as HOTS questions from each student participating in this study and the percentage of the number of items compared to the total items that each student determined as items that require cognitive process at level C4 or higher. From a total of 150 items that were supposed to be HOTS questions, based on the judgment of expert 1, each student was able to create zero to eight HOTS questions (M=3.20, SD=2.17) with a total of 64 (42.67%) items. Meanwhile, based on the judgment of expert 2, it was found that each student was able to create zero to seven HOTS questions (M=2.75, SD=2.15) with a total of only 55 (36.67%) items. Thus, less than half of the total items proposed by students, which according to them these items require cognitive process at the level of C4, C5, or C6, can be categorized as HOTS questions by experts.

Table 7 presents the characteristics of the test items, which are categorized as HOTS questions, which are specified by type of item and level of cognitive process. From **Table 7**, it was found that among the 64 test items, which were categorized as HOTS questions by expert 1, 36 (56.25%) items were multiple choice, and 28 (43.75%) items were essays. Meanwhile, of the 55 items categorized as HOTS questions by expert 2, 31 (56.36%) items were in the form of multiple choice and the remaining 24 (43.64%) items were in the form of essays. If it is associated with the characteristics of the test items that students proposed (see **Table 4**), the decrease in the number of items categorized as HOTS questions by expert 1 and expert 2, respectively is 62.11% and 67.37% for multiple-choice items and 49.09% and 56.36% for essay type items. Although in terms of the quantity of multiple choice items constructed by students there were more than essay items as a consequence of the rule that each student must construct 12 multiple-choice items

Table 7. Number of HOTS question developed by students in terms of type of item & level of cognitive process based on expert judgement

Item type	Expert 1, n (%)			Expert 2, n (%)		
	C4 item	C5 item	C6 item	C4 item	C5 item	C6 item
Multiple-choice (selected-response)	16 (25.00)	16 (25.00)	4 (6.25)	12 (21.82)	14 (25.45)	5(9.09)
Essay (constructed-response)	6 (9.38)	13 (20.31)	9 (14.06)	7 (12.73)	11 (20.00)	6 (10.91)
Total	22 (34.38)	29 (45.31)	13 (20.31)	19 (34.55)	25 (45.45)	11 (20.00)

and three essay items to obtain 95 multiple-choice items and 55 essay items, which are considered HOTS questions by students, the percentage decrease in the number of test items that can be categorized as HOTS questions indicates that students experience more challenges in developing multiple-choice compared to essay HOTS questions.

The main issue that became the concern of the experts until they finally decided that an item could not be categorized as a HOTS question was that the action verbs used in the question indicators for cognitive process of level C4, C5, or C6 more reflected one or even four levels below the level that students determined. Besides the issue of the accuracy of the use of action verbs, which affects the questions students develop, the issue that gains the attention of experts was the use of domain knowledge, where it was detected that several items only require factual knowledge or basic information.

Table 8 summarizes the findings of some of the test items proposed by students, which according to the two experts, these items cannot be categorized as HOTS questions. Furthermore, the two experts pointed out that the test items only required cognitive process at level C2 or C3, or those test items were said to be questions that only required lower order thinking skills (LOTS).

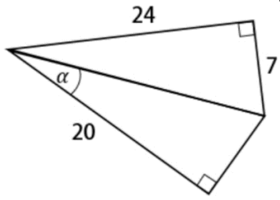
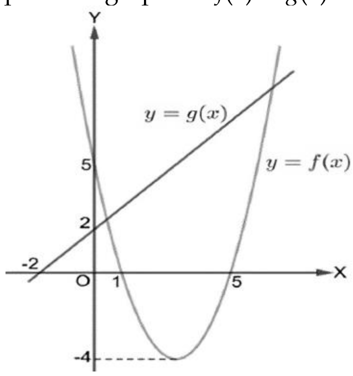
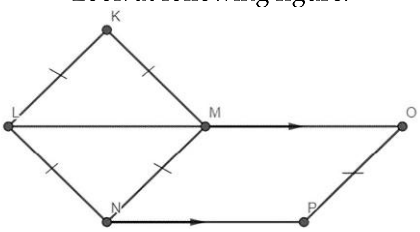
Table 8 shows the five test items each proposed by five students, which are more suitable to be categorized as LOTS questions than HOTS questions. The action verb chosen and used by the student S1 in the test item indicator is considered more suitable to reflect cognitive process at the level of C3 (apply). If the question posed by the student S1 was examined further, it is found that it only takes an understanding of the definition and properties of a pyramid and the Pythagorean theorem or Heron's formula to answer the item correctly. In other words, this test item does not require the ability to evaluate, criticize, or check the correctness, consistency, or effectiveness of a process, a procedure, or an idea. In the next case, the action verb used by student S2 in the test item indicator also do not match the recommended action verbs at the level C5, where the action verb "analyze" reflects the level of C4. After investigating the question posed by the student S2, the two experts agreed that the ability to apply basic knowledge or understanding of the Pythagorean theorem and trigonometric ratios in right triangles that had been acquired from learning in class was sufficient to answer the question correctly. In other words, the thinking skills needed to answer this question are still at the level of C3 in the revised Bloom's taxonomy.

We continue the discussion on the test items posed by students S8, S15, and S18. For the case of the test item posed by student S8, the experts argued that the action verb used, i.e., determine, better reflect cognitive process at the level of C2 or C3. If we look at the multiple-choice question posed by the student S8, it can be found that in order to solve the problem correctly, it is enough to know the functions that represent the given parabola and straight line and the concept of composite functions. Once we understand that the function of a parabola that intersects the x-axis at $(x_1, 0)$ and $(x_2,0)$ and intersects the y-axis at $(0, y)$ can be determined through the equation $y=a(x-x_1)(x-x_2)$ and the function of the line that intersects the x-axis and y-axis at $(x_1, 0)$ and $(0, y_2)$ respectively can be determined through the equation $y_2x+x_1y=x_1y_2$, the composition of the functions $f(x)$ and $g(x)$ can be determined including its value for $x=4$. Accordingly, the two experts agreed to categorize the test items posed by the student S8 as a LOTS question even though it involved the ability to make connections among the concepts of parabola, straight line, and composition of functions.

Afterwards, the action verb used in the test item indicator proposed by the student S15 as provided in **Table 8**, i.e., "conclude", is in accordance with the cognitive process level chosen, namely level C5. However, when this action verb is viewed in the test item indicator as a whole along with the test item that student S15 put forward, it can be concluded that the test item is not a HOTS question. Even expert 1 and expert 2 suggested that the test item can be solved simply by using thinking skills or cognitive processes at level C1 and C2 respectively. The suggestions from the two experts seem to be derived from the fact that the understanding of the period of the trigonometric functions $f(x)=\sin x$, $f(x)=\cos x$, and $f(x)=\tan x$ must have been learned by students in classroom either through information that the teacher provides directly or through activities that facilitated by the teacher through observing the graphical representation of the three trigonometric functions. Lastly, even though the action verb used by the student S18 in the test item indicator she proposed (see **Table 8**) is in accordance with the determined level of cognitive process, based on experts the test item is still categorized as the test item that requires cognitive process at the level of C3 (apply) and hence it is still categorized as a LOTS question.

Table 8. Items proposed by students which cannot be categorized as HOTS questions

<table>
<tr><th>Student</th><th>Item indicator (cognitive process level proposed by student)</th><th>Item/question</th></tr>
<tr><td>S1</td><td>Students can solve problems related to area of a triangle correctly (C5)</td><td>Ahmad made a miniature pyramid from 50,000 IDR bill as shown in figure below. It is known that length of base is 3.5 cm & length of its edges is 4.5 cm. What is surface area of miniature pyramid wall?

</td></tr>
<tr><td>S2</td><td>Presented two pictures of right-angled triangles whose information is incomplete; students are asked to analyze secant & cotangent values (C5)</td><td>Determine sec a and cot a from the figure below.

</td></tr>
<tr><td>S8</td><td>Given a parabola that represents a quadratic function f(x) & a graph of a linear function g(x), students are asked to determine value of (f○g)(x) for a certain x (C4)</td><td>Following figure represents graphs of f(x) & g(x). Value of (f○g)(4) is …

</td></tr>
<tr><td>S15</td><td>Students can conclude a period of a tangent function (C5)</td><td>Period of graph of trigonometric function f(x)=tan x is …
A. 45° B. 90° C. 180° D. 270° E. 360°</td></tr>
<tr><td>S18</td><td>Students are asked to associate formula for area of a parallelogram & a rhombus to find area of a flat shape in figure presented (C4)</td><td>Look at following figure.


In figure above, NP//MO, NP=18 cm, LO=30 cm, & KN=16 cm. Area of LNPOMK is …
A. 192 cm² B. 240 cm² C. 288 cm² D. 336 cm²</td></tr>
</table>

As the required information to solve the problem has been provided, we only need to apply our understanding of the formula to determine the area of a trapezoid and a triangle or the area of a parallelogram and a rhombus.

DISCUSSION

Apart from primarily investigating the challenges that prospective mathematics teacher faced in developing test items that require HOTS both based on their opinions and the end product, this study also explores students' views of project-based learning as a learning model that was implemented in the assessment of mathematics learning course. This study has shown benefits felt by students from the implementation of this learning model. From their engagement in lectures facilitated by a project-based learning model, students not only acquired knowledge or understanding on topics related to mathematics learning assessment but also have the opportunity to directly practice their knowledge or understanding so that they succeed in producing an end product in the form of a mathematics learning achievement test and its test rubric.

The development of the end product provides an opportunity for students to deepen their knowledge because it does not only focus on the final result but also on planning, the development process, and evaluation or reflection for improvement. This kind of thing is certainly important for students considering they are prospective mathematics educators who are expected to be not only competent at the theoretical level but also must be competent at the practical level or implementation of the knowledge they have mastered. The results of our study thus have provided additional support to previous studies, which demonstrated the benefits obtained from implementing a project-based learning model both from a general (e.g., Almulla, 2020; Chen et al., 2022; Crespí et al., 2022; Frank et al., 2003; Stefanou et al., 2013) and specific perspective on a teacher or prospective teacher professional development (e.g., Al-Busaidi & Al-Seyabi, 2021; Gumartifa et al., 2023; Guo & Yang, 2012; Herawati, 2018; King & Smith, 2020; Tsybulsky & Muchnik-Rozanov, 2021).

From a general perspective on the potential and advantages of implementing a project-based learning model, Stefanou et al. (2013) have provided evidence that this learning model provides more support to students in terms of elaborating their understanding, developing critical thinking skills, and controlling their thoughts and actions to achieve optimal learning. The benefits of implementing the project-based learning model mentioned by Stefanou et al. (2013) also indirectly obtained by students in this study, which were shown by the strategies they took in addressing challenges in developing achievement tests including HOTS questions. In the end product development process, students were given the opportunity to present the progress and obstacles they faced during the process. This opportunity allows students to elaborate on the understanding they have as well as opens opportunities for lecturers to provide appropriate support, either in the form of giving reinforcement or correcting misunderstandings exhibited by students. In addition, our study has demonstrated that when students encountered problems in developing HOTS questions, apart from consulting with lecturers, they took the initiative to discuss with colleagues or peers so that they received feedback from each other and reviewed various forms of references to strengthen their understanding. The actions that students take indicate that project-based learning applied in assessment of mathematics learning lectures promotes communication and collaboration skills as suggested by Crespí et al. (2022) and student self-regulated learning to achieve optimal learning.

We emphasize that through the application of the project-based learning model, students directly apply the knowledge and understanding they have to develop mathematics tests that contain HOTS questions with multiple-choice and essay items. The practice of developing this test certainly requires critical and creative thinking in considering the content and context used in the constructed questions as well as in constructing distractors. This indicates that project-based learning contribute to the development of critical and creative thinking as suggested by previous studies (e.g., Chen et al., 2022; Saimon et al., 2022). Based on the advantages provided by project-based learning facilitating students to acquire knowledge of learning assessment and deepen competence in developing mathematics tests, this learning model is considered as an alternative approach to preparing students to become teachers who can promote their student achievement as well as become professional teachers in the future (Guo & Yang, 2012). Considering one way to acquire knowledge more easily is by experiencing (Al-Busaidi & Al-Seyabi, 2021), essential knowledge about the assessment of mathematics learning including the development of measurement instruments in the form of tests would be easier for students to grasp because knowledge acquisition is accompanied by experience in developing mathematics tests that contain HOTS questions.

Although a number of advantages have been demonstrated from the project-based learning model applied in the assessment of mathematics learning course, it turns out that it does not necessarily make students develop mathematics tests that contain HOTS questions without any obstacle. Our study has identified the challenges students faced in developing mathematics achievement tests, which are then specifically focused on developing HOTS questions based on the revised version of Bloom’s taxonomy. The main challenge for students in developing the test is to match the level of cognitive process that they have set in the test rubric with the test item and its indicator, including using an action verb in the test item indicator appropriately. This result is in line with one of the findings from a study conducted by Winarti et al. (2021), namely, even though the teachers have attended a teacher’s training in developing HOTS questions, it was difficult for them to distinguish the level of cognitive processes in the revised version of Bloom’s taxonomy, especially distinguishing between C2 (understand) and C4 (analyze). Not only teachers, from the results of exploration by Purwasih (2020) on the perceptions of prospective teacher students enrolled in the Learning Evaluation course on the experience of constructing HOTS questions, it was revealed that most of them asserted to be confused in choosing an action verb that matched the cognitive level they chose, especially when that action verb represents two different cognitive levels. Confusion in selecting action verbs has an impact on the difficulty of formulating test item indicators to measure HOTS (Dahlan et al., 2020) and the emergence of doubts in students about the questions they have constructed, whether these questions already require cognitive processes at level C4 or are actually sufficient with cognitive processes at the C2 level to solve those problems (Purwasih, 2020). We found a similar phenomenon in our study, where students doubted the questions that they had developed, whether the questions were in accordance with the action verbs, test item indicators, and cognitive level they intended.

Saepuzaman et al. (2022) in their study found that it turned out to be more difficult for teachers to develop multiple-choice HOTS questions than essay types. Through an investigation using an open-ended questionnaire and the results of expert judgments on the end products that students have produced, our study also disclosed that students struggled more in developing multiple-choice HOTS questions and their options (a keyed option and distractors). This may be due to an understanding of HOTS questions that are more associated with constructed-response or essay type questions, where through these types of questions students would be more able to demonstrate their HOTS; while multiple-choice questions are perceived as only requiring cognitive processes at a lower level (Scouller, 1998; Scully, 2017; Simkin & Kuechler, 2005; Stanger-Hall, 2012). This kind of understanding needs to be set aside because it can hinder them from being able to properly develop multiple-choice HOTS questions. Another reason for the need to set aside this understanding is that in practice when multiple-choice items are well developed it can also be used to assess HOTS (Brookhart, 2010; Retnawati et al., 2018; Scully, 2017), which focuses on skills in analyzing and evaluating (Brookhart, 2010; Retnawati et al., 2018) and does not even rule out the possibility of arriving at skills in creating (Scully, 2017). More challenges that arise in developing HOTS questions can also be caused by the features of multiple choice items, where when the focus is on assessing HOTS it is highly recommended to contain introductory or stimulus material, stem, and options (Brookhart, 2010). Previous studies (e.g., Purwasih, 2020; Saepuzaman et al., 2022) revealed that one of the difficulties in developing HOTS questions is constructing introductory or stimulus material. Furthermore, the challenges experienced by students in developing distractors could be caused by the nature of the distractors themselves, which must be plausible for those who have misunderstandings or misconceptions or make errors. Lack of sensitivity to misunderstandings, misconceptions, or errors that students demonstrate when learning or solving problems can hinder developing distractors that can function as they should. In her article, Scully (2017) offers several strategies that can be taken to properly develop multiple-choice items to assess HOTS.

Implications for Practice

On the one hand, this study has demonstrated a number of benefits from the implementation of a project-based learning model in the assessment of mathematics learning course in supporting the competency development of students who are prospective mathematics teachers in conducting assessments. Accordingly, we argue that the implementation of this learning model could also be applied to other courses (e.g., mathematics learning design) that focus on providing support for prospective mathematics teacher students to master and develop the competencies or knowledge needed to become a professional mathematics teacher in the future. On the other hand, students who participated in this study still experienced challenges in developing a mathematics learning achievement test in which the test items were constructed according to levels on the cognitive process dimension in the revised version of Bloom's taxonomy, including in developing HOTS questions. The main challenges faced by students in developing HOTS questions particularly for multiple-choice items are related to two things, namely the construction of indicators according to the set cognitive process level and the use of action verbs. As a consequence, these two issues need more attention. There needs to be more emphasis on the meaning of each level or category in the cognitive process domain and its subcategories (see Anderson & Krathwohl, 2001; Krathwohl, 2002). In addition, given that there is an action verb that can be used to construct test item indicators and learning objectives that represent more than one level of cognitive process, it is important to present examples of questions whose indicators use the same action verb but represent different levels of cognitive process. Attention to these issues can also be expanded not only in higher education to prepare prospective teachers to become professional teachers in conducting good quality learning assessments in the future but can also be adapted for teacher professional development programs organized in the form of workshops or community service by university lecturers.

Limitations and Future Directions

The study that we conducted to reveal the phenomenon of the challenges that prospective mathematics teacher students face in developing HOTS was limited to data collected from administering open-ended questionnaires and analysis conducted by experts on the end products that students yielded. We suggest future studies to explore this phenomenon by conducting interviews with students. It is suggested that this interview should be focused on verifying students' understanding and satisfaction of HOTS questions that students developed in the mathematics learning achievement test. Although in practice the use of this interview technique has challenges when it involves many participants in terms of time and effort, it can be a means for lecturers to provide direct constructive feedback on student work and provide material for consideration for experts in judging the suitability of test item indicators with the level of cognitive process and test items with indicators.

Through this study we have exhibited that one of the challenges experienced by students in developing mathematics learning achievement tests was constructing distractors for a multiple-choice test item. Unfortunately, in this study we have not investigated further the extent of students' difficulties in constructing good distractors, especially for HOTS questions with multiple choice item types based on the end product that students have produced. Further investigation of this is interesting to do in future studies considering that several studies (e.g., DiBattista & Kurzawa, 2011; Rafi et al., 2023) have shown that through distractor analysis we can detect possible difficulties, misunderstandings, or misconceptions that students may have and mistakes or errors that they may do. The advantage of this distractor analysis can be obtained, of course, when the distractors constructed are of good quality in the sense that they are derived from an understanding of the mistakes that students frequently make when solving problems or misconceptions demonstrated by students when learning in class (Gierl et al., 2017; Shin et al., 2019). Investigation of how students construct distractors from HOTS questions on the other hand is also a means of uncovering the extent of PCK of prospective mathematics teacher students, which includes knowledge of misconceptions experienced by students (Gudmundsdottir & Shulman, 1987; Hill et al., 2008) and their future development strategies considering that a good PCK is important for teachers to have.

CONCLUSIONS

As mathematics education in the 21st century should be more focused on providing support to students to develop their HOTS, learning activities and assessment that teachers facilitate should also be oriented towards HOTS. Accordingly, teachers need to be provided with supports to design and implement learning activities and carry out assessments according to what is desired, as well as prospective mathematics teachers. The challenges students face when developing HOTS questions included in learning achievement tests in mathematics as part of the course activities that they participate in are still understudied. In our study, besides focusing on the main goal of identifying the challenges in developing HOTS questions, we also investigated the benefits of implementing a project-based learning model in the courses that students take on their competence in developing mathematics learning achievement tests and HOTS questions and strategies that students take to overcome obstacles in this development.

Our study has contributed to providing additional support to the existing literature regarding the benefits of implementing project-based learning in higher education for prospective teacher. Project-based learning has shown its benefits in supporting prospective mathematics teacher students to deepen their understanding and knowledge of mathematics learning assessment, test development, and HOTS question development while at the same time putting their knowledge into practice directly to develop mathematics learning achievement tests. From the responses that students put forward regarding the strategies they used in overcoming obstacles in developing HOTS questions, it indicates that project-based learning also promotes self-regulated learning. Returning to the main focus of this study, we have identified challenges faced by prospective mathematics teachers in developing HOTS questions. These challenges include distinguishing cognitive process levels in revised Bloom's taxonomy, selecting action verbs that match the level of cognitive process specified, integrating action verbs into indicators of test items, and developing multiple-choice HOTS questions including the distractors.

Author contributions: All authors have sufficiently contributed to the study and agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Ethical statement: The authors stated that the study did not require an ethics committee approval. The study was carried out when the course was taking place.

Declaration of interest: No conflict of interest is declared by authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Al-Busaidi, S., & Al-Seyabi, F. (2021). Project-based learning as a tool for student-teachers' professional development: A study in an Omani EFL teacher education program. *International Journal of Learning, Teaching and Educational Research, 20*(4), 116-136. https://doi.org/10.26803/ijlter.20.4.7

Almulla, M. A. (2020). The effectiveness of the project-based learning (PBL) approach as a way to engage students in learning. *SAGE Open, 10*(3), 1-15. https://doi.org/10.1177/2158244020938702

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Addison-Wesley Longman.

Apino, E., & Retnawati, H. (2017). Developing instructional design to improve mathematical higher order thinking skills of students. *Journal of Physics: Conference Series, 812*(1), 1-7. https://doi.org/10.1088/1742-6596/812/1/012100

Barak, M., & Shakhman, L. (2008). Reform-based science teaching: Teachers' instructional practices and conceptions. *EURASIA Journal of Mathematics,* *Science and Technology Education, 4*(1), 11–20. https://doi.org/10.12973/ejmste/75301

Bell, S. (2010). Project-based learning for the 21st century: Skills for the future. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 83*(2), 39-43. https://doi.org/10.1080/00098650903505415

Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist, 26*(3-4), 369-398. https://doi.org/10.1080/00461520.1991.9653139

Bogdan, R. C., & Biklen, S. K. (2007). *Qualitative research for education: An introduction to theory and methods* (5th edn.). Pearson Education.

Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. ASCD.

Chen, S.-Y., Lai, C.-F., Lai, Y.-H., & Su, Y.-S. (2022). Effect of project-based learning on development of students' creative thinking. *The International Journal of Electrical Engineering & Education, 59*(3), 232-250. https://doi.org/10.1177/0020720919846808

Cohen, R. J., & Swerdlik, M. E. (2018). *Psychological testing and assessment: An introduction to tests and measurement*. McGraw-Hill Education.

Corden, A., & Sainsbury, R. (2006). Exploring 'quality': Research participants' perspectives on verbatim quotations. *International Journal of Social Research Methodology, 9*(2), 97-110. https://doi.org/10.1080/13645570600595264

Crespí, P., García-Ramos, J. M., & Queiruga-Dios, M. (2022). Project-based learning (PBL) and its impact on the development of interpersonal competences in higher education. *Journal of New Approaches in Educational Research, 11*(2), 259-276. https://doi.org/10.7821/naer.2022.7.993

Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches*. SAGE.

Dahlan, D., Permana, L., & Oktariani, M. (2020). Teacher's competence and difficulties in constructing HOTS instruments in economics subject. *Jurnal Cakrawala Pendidikan* [*Journal of Educational Horizons*], *39*(1), 111-119. https://doi.org/10.21831/cp.v39i1.28869

DiBattista, D., & Kurzawa, L. (2011). Examination of the quality of multiple-choice items on classroom tests. *The Canadian Journal for the Scholarship of Teaching and Learning, 2*(2), 1-23. https://doi.org/10.5206/cjsotl-rcacea.2011.2.4

Djidu, H., Jailani, J., & Retnawati, H. (2021). Higher-order thinking skills among male and female students: An experimental study of the problem-based calculus learning model in secondary schools. *Beta: Jurnal Tadris Matematika* [*Beta: Journal of Mathematics Education*], *14*(2), 107-125. https://doi.org/10.20414/betajtm.v14i2.432

Drijvers, P., Kodde-Buitenhuis, H., & Doorman, M. (2019). Assessing mathematical thinking as part of curriculum reform in the Netherlands. *Educational Studies in Mathematics, 102*(3), 435-456. https://doi.org/10.1007/s10649-019-09905-7

Evidiasari, S., Subanji, S., & Irawati, S. (2019). Students' spatial reasoning in solving geometrical transformation problems. *Indonesian Journal on Learning and Advanced Education*, *1*(2), 38-51. https://doi.org/10.23917/ijolae.v1i2.8703

Forster, M. (2004). Higher order thinking skills. *Research Developments, 11*(11), 10-15.

Frank, M., Lavy, I., & Elata, D. (2003). Implementing the project-based learning approach in an academic engineering course. *International Journal of Technology and Design Education, 13*(3), 273-288. https://doi.org/10.1023/A:1026192113732

Gierl, M. J., Bulut, O., Guo, Q., & Zhang, X. (2017). Developing, analyzing, and using distractors for multiple-choice tests in education: A comprehensive review. *Review of Educational Research, 87*(6), 1082-1116. https://doi.org/10.3102/0034654317726529

Gudmundsdottir, S., & Shulman, L. (1987). Pedagogical content knowledge in social studies. *Scandinavian Journal of Educational Research, 31*(2), 59-70. https://doi.org/10.1080/0031383870310201

Gumartifa, A., Syahri, I., Siroj, R. A., Nurrahmi, M., & Yusof, N. (2023). Perception of teachers regarding problem-based learning and traditional method in the classroom learning innovation process. *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, *5*(2), 151-166. https://doi.org/10.23917/ijolae.v5i2.20714

Guo, P., Saab, N., Post, L. S., & Admiraal, W. (2020). A review of project-based learning in higher education: Student outcomes and measures. *International Journal of Educational Research, 102*, 1-13. https://doi.org/10.1016/j.ijer.2020.101586

Guo, S., & Yang, Y. (2012). Project-based learning: An effective approach to link teacher professional development and students learning. *Journal of Educational Technology Development and Exchange, 5*(2). https://doi.org/10.18785/jetde.0502.04

Herawati, A. (2018). The application of project-based learning in teachers' professional development program. In *Proceedings of the 6th International Conference on Information and Education Technology* (pp. 119-123). https://doi.org/10.1145/3178158.3178182

Heron, M., & Palfreyman, D. M. (2021). Exploring higher-order thinking in higher education seminar talk. *College Teaching*. https://doi.org/10.1080/87567555.2021.2018397

Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education, 39*(4), 372-400. https://doi.org/10.5951/jresematheduc.39.4.0372

Hmelo, C. E., & Ferrari, M. (1997). The problem-based learning tutorial: Cultivating higher order thinking skills. *Journal for the Education of the Gifted, 20*(4), 401-422. https://doi.org/10.1177/016235329702000405

Holmes, V.-L., & Hwang, Y. (2016). Exploring the effects of project-based learning in secondary mathematics education. *The Journal of Educational Research, 109*(5), 449-463. https://doi.org/10.1080/00220671.2014.979911

Ijirana, I., Mansyur, J., Rizal, M., & Aminah, S. (2021). Longitudinal study of metacognitive skills and external representation of students in the context of problem-solving. *Indonesian Journal on Learning and Advanced Education*, *3*(3), 194-206. https://doi.org/10.23917/ijolae.v3i3.13563

Jailani, J., Sugiman, S., & Apino, E. (2017). Implementing the problem-based learning in order to improve the students' HOTS and characters. *Jurnal Riset Pendidikan Matematika* [*Journal of Mathematics Education Research*], *4*(2), 247-259. https://doi.org/10.21831/jrpm.v4i2.17674

Kaur, S. (2019). Role of a teacher in student learning process. *International Journal of Business and Management Invention, 8*(12), 41-45.

King, B., & Smith, C. (2020). Using project-based learning to develop teachers for leadership. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 93*(3), 158-164. https://doi.org/10.1080/00098655.2020.1735289

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice, 41*(4), 212-218. https://doi.org/10.1207/s15430421tip4104_2

Kruszewska, A. (2021). Teachers' understanding of the role of key competences in working with children: A study of the perceptions of early years educators in Southern Poland. *Education 3-13: International Journal of Primary, Elementary and Early Years* Education, *49*(6), 650-660. https://doi.org/10.1080/03004279.2020.1762691

Kyriacou, C. (2007). *Essential teaching skills*. Nelson Thornes.

Lewis, A., & Smith, D. (1993). Defining higher order thinking. *Theory into Practice, 32*(3), 131-137. https://doi.org/10.1080/00405849309543588

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE. https://doi.org/10.1016/0147-1767(85)90062-8

Liu, J., Ma, Y., Sun, X., Zhu, Z., & Xu, Y. (2022). A systematic review of higher-order thinking by visualizing its structure through HistCite and CiteSpace software. *The Asia-Pacific Education Researcher, 31*(6), 635-645. https://doi.org/10.1007/s40299-021-00614-5

Machromah, I., Setyaningsih, R., Sari, C. K., Swastika, A., & Khoirunnisa, M. (2019). HOTS problem posing on mathematics pre-service teachers. In *Proceedings of the 4th Progressive and Fun Education International Conference* (pp. 1-11). https://doi.org/10.4108/eai.7-8-2019.2288426

Magnusson, S., Krajcik, J., & Borko, H. (1999). Nature, sources, and development of pedagogical content knowledge for science teaching. In J. Gess-Newsome, & N. G. Lederman (Eds.), *Examining pedagogical content knowledge* (pp. 95-132). Kluwer Academic Publishers. https://doi.org/10.1007/0-306-47217-1_4

Marczewska, M., Weresa, M. A., & Lachowicz, M. (2023). Towards creativity and innovation in universities: Study on Central and Eastern Europe. *Journal of the Knowledge Economy*. https://doi.org/10.1007/s13132-023-01139-6

Miller, E. C., & Krajcik, J. S. (2019). Promoting deep learning through project-based learning: A design problem. *Disciplinary and Interdisciplinary Science Education Research, 1*(1), 7. https://doi.org/10.1186/s43031-019-0009-6

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record, 108*(6), 1017-1054.

Newmann, F. M. (1991). Organizational features and classroom thoughtfulness in secondary school social studies departments. *Theory & Research in Social Education, 19*(4), 391-409. https://doi.org/10.1080/00933104.1991.10505648

Newton, P. M., Da Silva, A., & Peters, L. G. (2020). A pragmatic master list of action verbs for Bloom's taxonomy. *Frontiers in Education, 5*, 1-6. https://doi.org/10.3389/feduc.2020.00107

Nikolaeva, S. (2012). Improving initial teacher education by using the project-based approach. *Educational Research, 1*(1), 51-60. https://doi.org/10.5838/erej.2012.11.04

Nitko, A. J., & Brookhart, S. M. (2011). *Educational assessment of students*. Pearson Education.

Park, S., & Oliver, J. S. (2008). Revisiting the conceptualization of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Research in*

Science Education, 38(3), 261-284. https://doi.org/10.1007/s11165-007-9049-6

Purwasih, J. H. G. (2020). Kendala calon pendidik dalam membuat soal pilihan ganda higher order thinking (HOT) [Obstacles of prospective educators in constructing higher order thinking (HOT) multiple-choice questions]. *Jurnal Sosial Humaniora* [*Journal of Social Humanities*], *13*(1), 12-22. https://doi.org/10.12962/j24433527.v13i1.6746

Radmehr, F., & Vos, P. (2020). Issues and challenges of 21st century assessment in mathematics education. In L. Leite, E. Oldham, A. S. Afonso, F. Viseu, L. Dourado, & H. Martinho (Eds.), *Science and mathematics education for 21st century citizens: Challenges and ways forward* (pp. 437-462). Nova Science Publishers.

Radović, S., Firssova, O., Hummel, H. G. K., & Vermeulen, M. (2021). Strengthening the ties between theory and practice in higher education: An investigation into different levels of authenticity and processes of re- and de-contextualization. *Studies in Higher Education, 46*(12), 2710-2725. https://doi.org/10.1080/03075079.2020.1767053

Rafi, I., & Sugiman, S. (2019). Problem posing ability among prospective mathematics teachers. *Pythagoras: Jurnal Pendidikan Matematika* [*Pythagoras: Journal of Mathematics Education*], *14*(2), 188-198. https://doi.org/10.21831/pg.v14i2.29976

Rafi, I., Retnawati, H., Apino, E., Hadiana, D., Lydiati, I., & Rosyada, M. N. (2023). What might be frequently overlooked is actually still beneficial: Learning from post national-standardized school examination. *Pedagogical Research, 8*(1), 1-15. https://doi.org/10.29333/pr/12657

Retnawati, H. (2016). *Analisis kuantitatif instrumen penelitian* [*Quantitative analysis of research instrument*]. Parama Publishing.

Retnawati, H., Djidu, H., Kartianom, K., Apino, E., & Anazifa, R. D. (2018). Teachers' knowledge about higher-order thinking skills and its learning strategy. *Problems of Education in the 21st Century, 76*(2), 215-230. https://doi.org/10.33225/pec/18.76.215

Reynolds, C. R., Livingston, R. B., & Willson, V. (2010). *Measurement and assessment in education*. Pearson.

Rianasari, V. F., & Apriani, M. S. (2019). Pre-service mathematics teachers' knowledge about higher-order thinking skills. *Journal of Physics: Conference Series, 1366*(1), 1-9. https://doi.org/10.1088/1742-6596/1366/1/012083

Richland, L. E., & Simms, N. (2015). Analogy, higher order thinking, and education. *Wiley Interdisciplinary Reviews: Cognitive Science, 6*(2), 177-192. https://doi.org/10.1002/wcs.1336

Sa'adah, L., & Anjarwati, R. (2022). High order thinking skill in pre-service teachers' lesson plan. *International Journal of Education, Language, and Religion, 4*(1), 37-46. https://doi.org/10.35308/ijelr.v4i1.5457

Saepuzaman, D., Retnawati, H., Istiyono, E., & Haryanto, H. (2022). Developing HOTS instruments: Is it difficult for physics teachers? *Jurnal Pendidikan MIPA* [*MIPA Education Journal*], *23*(4), 1802-1826. https://doi.org/10.23960/jpmipa/v23i4.pp1802-1826

Saimon, M., Lavicza, Z., & Dana-Picard, T. (2022). Enhancing the 4Cs among college students of a communication skills course in Tanzania through a project-based learning model. *Education and Information Technologies, 28*, 6269-6285. https://doi.org/10.1007/s10639-022-11406-9

Schulz, H., & FitzPatrick, B. (2016). Teachers' understandings of critical and higher order thinking and what this means for their teaching and assessments. *Alberta Journal of Educational Research, 62*(1), 61-86.

Scouller, K. (1998). The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. *Higher Education, 35*(1), 453-472. https://doi.org/10.1023/A:1003196224280

Scully, D. (2017). Constructing multiple-choice items to measure higher-order thinking. *Practical Assessment ,Research, and Evaluation, 22*(4), 1-13. https://doi.org/10.7275/swgt-rj52

Shin, J., Guo, Q., & Gierl, M. J. (2019). Multiple-choice item distractor development using topic modeling approaches. *Frontiers in Psychology, 10*(1), 1-14. https://doi.org/10.3389/fpsyg.2019.00825

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14. https://doi.org/10.3102/0013189X015002004

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1-23. https://doi.org/10.17763/haer.57.1.j463w79r56455411

Sideeg, A. (2016). Bloom's taxonomy, backward design, and Vygotsky's zone of proximal development in crafting learning outcomes. *International Journal of Linguistics, 8*(2), 158-186. https://doi.org/10.5296/ijl.v8i2.9252

Simkin, M. G., & Kuechler, W. L. (2005). Multiple-choice tests and student understanding: What is the connection? *Decision Sciences Journal of Innovative Education, 3*(1), 73-98. https://doi.org/10.1111/j.1540-4609.2005.00053.x

Singer, F. M., Ellerton, N., & Cai, J. (2013). Problem-posing research in mathematics education: New questions and directions. *Educational Studies in Mathematics, 83*(1), 1-7. https://doi.org/10.1007/s10649-013-9478-2

Sinta, U. A., Roebyanto, G., & Nuraini, N. L. S. (2022). Analisis kesulitan guru dalam menyusun soal evaluasi berbasis HOTS pada pembelajaran matematika di SDN Torongrejo 2 [Analysis of teacher difficulties in compiling HOTS-based evaluation questions in mathematics learning at SDN Torongrejo 2]. *Jurnal Pembelajaran, Bimbingan, dan Pengelolaan Pendidikan* [*Journal of Learning, Guidance, and Education Management*], *2*(1), 45-53. https://doi.org/10.17977/um065v2i12022p45-53

Stahl, N. A., & King, J. R. (2020). Expanding approaches for research: Understanding and using trustworthiness in qualitative research. *Journal of Developmental Education, 44*(1), 26-28.

Stanger-Hall, K. F. (2012). Multiple-choice exams: An obstacle for higher-level thinking in introductory science classes. *CBE-Life Sciences Education, 11*(3), 294-306. https://doi.org/10.1187/cbe.11-11-0100

Stefanou, C., Stolk, J. D., Prince, M., Chen, J. C., & Lord, S. M. (2013). Self-regulation and autonomy in problem- and project-based learning environments. *Active Learning in Higher Education, 14*(2), 109-122. https://doi.org/10.1177/1469787413481132

Suherman, S., Prananda, M. R., Proboningrum, D. I., Pratama, E. R., Laksono, P., & Amiruddin, A. (2020). Improving higher order thinking skills (HOTS) with project based learning (PjBL) model assisted by GeoGebra. *Journal of Physics: Conference Series, 1467*, 012027. https://doi.org/10.1088/1742-6596/1467/1/012027

Takiddin, T., Jalal, F., & Neolaka, A. (2020). Improving higher order thinking skills through project-based learning in primary schools. *Tarbiya: Journal of Education in Muslim Society, 7*(1), 16-28. https://doi.org/10.15408/tjems.v7i1.14052

Teherani, A., Martimianakis, T., Stenfors-Hayes, T., Wadhwa, A., & Varpio, L. (2015). Choosing a qualitative research approach. *Journal of Graduate Medical Education, 7*(4), 669-670. https://doi.org/10.4300/JGME-D-15-00414.1

Thompson, T. (2008). Mathematics teachers' interpretation of higher-order thinking in Bloom's taxonomy. *International Electronic Journal of Mathematics Education, 3*(2), 96-109. https://doi.org/10.29333/iejme/221

Tsybulsky, D., & Muchnik-Rozanov, Y. (2021). Project-based learning in science-teacher pedagogical practicum: The role of emotional experiences in building preservice teachers' competencies. *Disciplinary and Interdisciplinary Science Education Research, 3*(9), 9. https://doi.org/10.1186/s43031-021-00037-8

Van den Berg, G. (2004). The use of assessment in the development of higher-order thinking skills. *Africa Education Review, 1*(2), 279-294. https://doi.org/10.1080/18146620408566285

Winarti, W., Hairida, H., & Lestari, I. (2021). Deskripsi kemampuan guru membuat soal berdasarkan pada kurikulum 2013 di sekolah menengah atas Kabupaten Landak [A description of the teacher's ability to create questions based on the curriculum 2013 in high schools in Landak District]. *Jurnal Ilmiah Wahana Pendidikan* [*Scientific Journal of Educational Forum*], *7*(2), 108-115. https://doi.org/10.5281/zenodo.4659018

Wisrance, M. W., & Semiun, T. T. (2020). LOTS and HOTS of teacher-made test in junior high school level in Kefamenanu. *Journal of English Education, 6*(2), 62-76. https://doi.org/10.30606/jee.v6i2.574

Wrenn, J., & Wrenn, B. (2009). Enhancing learning by integrating theory and practice. *International Journal of Teaching and Learning in Higher Education, 21*(2), 258-265.

Zhou, Y., Gan, L., Chen, J., Wijaya, T. T., & Li, Y. (2023). Development and validation of a higher-order thinking skills assessment scale for pre-service teachers. *Thinking Skills and Creativity, 48*(2), 1-20. https://doi.org/10.1016/j.tsc.2023.101272

Zohar, A., & Dori, Y. J. (2003). Higher order thinking skills and low-achieving students: Are they mutually exclusive? *Journal of the Learning Sciences, 12*(2), 145-181. https://doi.org/10.1207/S15327809JLS1202_1