DEVELOPMENT OF ASSESSMENT TOOLS OF CRITICAL THINKING IN MATHEMATICS IN THE CONTEXT OF HOTS

HENRA SAPUTRA TANJUNG¹, SITI AMINAH NABABAN, CHOLIS SADLIJAH, AND SUBANJI

ABSTRACT. The present was purposed to determine whether or not students have possessed critical thinking skills through assessment. In general, the study was aimed at developing tools for assessing mathematical critical thinking skills, specifically in the context of higher-order thinking skills (henceforth, HOTS). To this day, the primary objective of the assessment is to determine students’ competencies before and after learning processes, thus allowing teachers to get feedback. The teachers will benefit from the feedback to improve the instructional media, including learning approaches, activities, and resources. In addition, the feedback also serves to disseminate information to parents regarding the effectiveness of educational activities at schools. This study relied on Dick and Carrey’s research and development model; the stages comprised literature study, planning of the product, development of the product, expert validation, trials, revision, and finalization of the final product. There were two trials in this research: trial I was carried out in class IX-A, and trial II was conducted in class IX-B. According to the results, (1) the designed instructional media were deemed effective. This was evident from the results of individual and classical learning mastery, student activities (within the bounds set by a specific regulation), and students’ positive reception to the learning media. (2) From trial I to trial II, the improvement in students’ critical thinking skills was 0.32 on average, and the classical learning mastery was measured at 21.4%.

1. INTRODUCTION

Education refers to efforts in boosting the competencies of human resources. The success of educational processes is determined by teachers’ assessment skills. Such skills

¹corresponding author
2020 Mathematics Subject Classification. 97D70, 97D60.
Key words and phrases. Assessment, Critical Thinking Skills, Higher-Order Thinking Skills.
are essential to determine whether or not the objectives of learning, stated in the curriculum, are attained. Furthermore, assessment skills are central to actualizing quality teaching and learning processes. This notion also corresponds to The Standard of Academic Qualification and Competencies of Mathematics Teachers in junior high schools and its equivalent [1]. With regard to the issues of critical thinking skills, the present study had carried out a need analysis on selected junior high schools in West Aceh. The result shows that: 1) Cognitive assessment instruments tend to emphasize the students’ memorizing skills, rather than higher-order thinking skills; tests for examining these skills are yet available. 2) According to a report by PISA, the thinking skills of children in Indonesia are considered low. This situation blames several factors, one of which is that the students are not accustomed to solving contextual tests, which demand their reasoning, argument skills, and creativity. 3) Teachers have low competence in developing assessment instruments. 4) The assessment instruments for measuring students’ critical thinking, especially in the context of HOTS, are yet available. Another concern is the fact that Curriculum 2013 focuses on promoting students’ skills in observing, asking, reasoning, and communicating everything they have learned. For this reason, it is suggested that the instruments for assessing mathematical critical thinking in the context of HOTS should enable the students to practice the skills previously mentioned. In Aceh, little is known about the assessment instruments for measuring students’ critical thinking in HOTS context, since the only cognitive assessment instruments tend to emphasize the students’ memorizing skills, rather than higher-order thinking skills. Such tests or assignments are, in fact, essential to help students retain their concentration in the class and, at the same time, discover their potential [2]. The enhancement of HOTS has been one of the main goals of teaching mathematics. In junior high schools, the students are supposed to start practicing thinking skills. Indeed, the skills should fit the capabilities of people in the students’ age [3]. Assessment is simply defined as a comprehensive process to identify students’ performance. This process is central to learning; it is regarded as one of the efforts to improve the quality of education. Recently, there has been a shift in the standard of learning assessment, from everything assessable to everything that must be assessed [4]. Assessment comes from a psychometric/measurement perspective, and is primarily concerned with scores of groups or individuals, rather than examining students’ thinking and communication processes [5]. A psychometric perspective is concerned with reliably measuring the outcome of learning, rather than the learning itself [6]. The development of critical thinking skills assessment instruments is an effort to meet the needs of teachers in assessing critical thinking skills of students. In mathematics, development efforts are also carried out, given the need for critical thinking skills itself [7]. Critical thinking is needed, where every day individuals face
unlimited information, complex problems, rapid technological and social changes [8]. Furthermore, assessment instruments play a strategic role in the decision-making process of teachers and schools regarding student learning outcomes, including the aspect of HOTS. Chief among the components of teaching are the mathematical critical thinking skills. These skills are not granted for granted. In fact, a student should go through intensive practices to master these skills [9]. Before teaching the students, a teacher is demanded to demonstrate his or her mathematical critical thinking skills well. Simply put, critical thinking skills are defined as the highest intellectual activities in human interaction, which allow an individual to get involved in a meaningful decision-making process [10]. HOTS are a set of thinking processes at a high cognitive level developed from various concepts and cognitive processes in the learning taxonomy, such as critical thinking skills, Bloom’s taxonomy, and taxonomy of learning, teaching, and assessment [11]. Through mastering HOTS, students can differentiate ideas and opinions, deliver good arguments, solve problems, construct explanations, formulate hypotheses, and comprehend complex ideas and interpret them into clear, more straightforward ideas [12]. Students are considered to fully master HOTS once they demonstrate the capabilities to correlate new information with the one that they have stored in their memory. On top of that, the students are able to reconstruct and develop everything they have learned to meet a particular goal or to find a solution to a specific, complex problem [13]. The primary objective of teaching HOTS is to enhance students’ thinking skills, bringing the skills to a new level, with an emphasis on critical thinking skills in learning information in various contexts [14]. Today’s learning evaluation is performed through the actualization of quality education in terms of the assessment standards. To this day, the improvement programs are carried out by adapting international assessment models gradually. The designed assessment is expected to assist students in shaping their HOTS better since these skills stimulate the way the students think everything they got from the class comprehensively [15]. HOTS are a part of the revised Bloom’s taxonomy; the skills comprise several action verbs, namely analyze (C4), evaluate (C5), and create (C6). All of which are applicable in designing a test.

2. Method

The present work applied a research and development (R&D) model by Dick and Carrey and was carried out in selected junior high schools in West Aceh regency. All quantitative and qualitative data were retrieved from the results of expert validation and trials, respectively. These data functioned to provide an overview of the developed products. The procedures encompassed (1) the analysis of learning assessment, i.e., determining
the design of the developed assessment based on the HOTS and digital literacy principles. (2) Development of assessment tools of critical thinking in mathematics in the context of dan digital literacy the students. (3) Validation of the developed assessment; its steps encompassed empirical trials, limited trials, and field trials. (4) Evaluation of the data of trials. In this research, the data were retrieved from the validation sheets. The sheets were completed by material experts, instructional media experts, and assessment experts. Another data collection instrument involved questionnaires distributed to the students. Stages in collecting the data were: (1) designing research instruments, e.g., tests, scoring and assessment rubrics, (2) determining the validity, (3) revising the product based on the input from the validators, (4) trials of the research instruments, (5) examining the reliability and determining the difficulty and distinguishing features of the test items, and (6) revision based on the results of the trials.

3. Results and Discussion

The primary goal of the present study is to report the results of the development of assessment tools of critical thinking skills in the context of HOTS. It was also aimed at reporting the student activities when using the developed instructional media, the students’ responses regarding the media, and the critical thinking learning outcomes. Observation of student activities is performed during learning processes for three times in trial I and trial II since the classes were also in three meetings. The result of the analysis of the student activities in trial I is displayed in Table 1, while the following Table 2 provides the data of Trial II.

### Table 1. Description of the table

<table>
<thead>
<tr>
<th>Activity</th>
<th>Meeting 1</th>
<th>Meeting 2</th>
<th>Meeting 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency</td>
<td>Percentage (%)</td>
<td>Frequency</td>
</tr>
<tr>
<td>a</td>
<td>57</td>
<td>26.39</td>
<td>58</td>
</tr>
<tr>
<td>b</td>
<td>30</td>
<td>13.89</td>
<td>28</td>
</tr>
<tr>
<td>c</td>
<td>67</td>
<td>31.02</td>
<td>80</td>
</tr>
<tr>
<td>d</td>
<td>43</td>
<td>19.91</td>
<td>35</td>
</tr>
<tr>
<td>e</td>
<td>18</td>
<td>8.33</td>
<td>12</td>
</tr>
<tr>
<td>f</td>
<td>1</td>
<td>0.46</td>
<td>3</td>
</tr>
<tr>
<td>Sum Total</td>
<td>216</td>
<td>100</td>
<td>216</td>
</tr>
</tbody>
</table>
Description:

(a) Paying attention to the teacher’s/friend’s explanation.
(b) Reading/comprehending contextual problems in the textbooks.
(c) Solving problems/formulating solutions.
(d) Discussing/asking others and/or teachers.
(e) Drawing conclusions of a particular procedure or concept, and presenting results.
(f) Irrelevant behaviors during the class.

The student activities, meeting 1, of each category is measured at 26.39%, 13.89%, 31.02%, 19.91%, 8.33%, and 0.46%. The percentage of student activities in the aspect of paying attention to teacher’s or friend’s explanations in 90 minutes scores 26.39%. This is based on the result of dividing the score of the frequency of activity of 12 students’ activities in category (a), or 57, and 216. Further, the result of the division is multiplied by 100%. The value 216 refers to the result of the division of the total duration of the learning process in meeting 1 (90 minutes) with the duration of observation per class (5 minutes); the result is multiplied by 12, or the number of the students observed. These calculations are also used to determine the percentage of other activities in each meeting since the duration and number of students in one observation session always vary (90 minutes or 135 minutes; there were 12 students in meeting 1 and 2, while in meeting three, there were 13 students). For this reason, the percentage of each category is subject to change, depending on the duration and number of students who participated in one observation session.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Meeting 1 Frequency</th>
<th>Meeting 1 Percentage (%)</th>
<th>Meeting 2 Frequency</th>
<th>Meeting 2 Percentage (%)</th>
<th>Meeting 3 Frequency</th>
<th>Meeting 3 Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>68</td>
<td>26.98</td>
<td>64</td>
<td>27.35</td>
<td>90</td>
<td>27.8</td>
</tr>
<tr>
<td>b</td>
<td>33</td>
<td>13.1</td>
<td>26</td>
<td>11.11</td>
<td>46</td>
<td>14.2</td>
</tr>
<tr>
<td>c</td>
<td>98</td>
<td>38.89</td>
<td>90</td>
<td>38.46</td>
<td>115</td>
<td>35.5</td>
</tr>
<tr>
<td>d</td>
<td>37</td>
<td>14.68</td>
<td>34</td>
<td>14.53</td>
<td>41</td>
<td>12.7</td>
</tr>
<tr>
<td>e</td>
<td>14</td>
<td>5.56</td>
<td>19</td>
<td>8.12</td>
<td>25</td>
<td>7.72</td>
</tr>
<tr>
<td>f</td>
<td>2</td>
<td>0.79</td>
<td>1</td>
<td>0.43</td>
<td>7</td>
<td>2.16</td>
</tr>
<tr>
<td>Sum Total</td>
<td>252</td>
<td>100</td>
<td>234</td>
<td>100</td>
<td>324</td>
<td>100</td>
</tr>
</tbody>
</table>
The student activities, meeting 2, of each category is measured at 26.98%, 13.1%, 38.89%, 14.68%, 5.56%, and 0.79%. The percentage of student activities in the aspect of paying attention to teacher's or friend's explanations in 90 minutes scores 26.98%. This is based on the result of dividing the score of the frequency of activity of 14 students’ activities in category (a), or 68, and 252. Further, the result of the division is multiplied by 100%. The value 252 refers to the result of the division of the total duration of the learning process in meeting 2 (90 minutes) with the duration of observation per class (5 minutes); the result is multiplied by 14, or the number of the students observed. These calculations are also used to determine the percentage of other activities in each meeting as the duration and number of students in one observation session always vary (90 minutes or 135 minutes; there were 14 students in meeting 1, 13 students in meeting 2, and 12 students in meeting 3). By that, the percentage of each category is subject to change, depending on the duration and number of students participated in one observation session. Provided in Table 3 is the percentage of each category in all three meetings, in trial I and trial II. On average, the percentage of each category of the student activities in trial I is measured at 27.37%, 12.8%, 37.61%, 13.96%, 7.13%, and 1.13%. This is based on the division of the percentage of each category and the total meetings (three meetings). For instance, the percentage of category (a) of the student activities in trial I gets 26.39%, 26.85%, and 28.49%. Each percentage is divided by 3. This method also applies during calculating the percentage of other activity categories.

Table 3. The Average Percentage of Student Activity

<table>
<thead>
<tr>
<th>Activity</th>
<th>Percentage of Activity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trial 1</td>
</tr>
<tr>
<td>a</td>
<td>27.37</td>
</tr>
<tr>
<td>b</td>
<td>12.80</td>
</tr>
<tr>
<td>c</td>
<td>37.61</td>
</tr>
<tr>
<td>d</td>
<td>13.96</td>
</tr>
<tr>
<td>e</td>
<td>7.13</td>
</tr>
<tr>
<td>f</td>
<td>1.13</td>
</tr>
<tr>
<td><strong>Sum Total</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

On average, the percentage of the duration of each activity in the class is displayed in the following Figure 1. The activity that the students spent the most is category (c), i.e., formulating solutions to a specific problem. The percentage of this category is 37.61%
and 33.42% in trial I and trial II, respectively. Such results signify that, for the students, solving mathematical problems in the worksheets is time-consuming compared to other activities.

![Graph 1: Average Percentage of the Duration of Student Activities](image1)

**Figure 1.** The Average Percentage of Student Activity

![Graph 2: Classical Mastery Standard](image2)

**Figure 2.** The Outcomes of the Test of Students’ Critical Thinking Skills
Figure 2 highlights the significant increase in the number of students who pass the standard in cycle II, thus implying the decline in the number of students who score below the minimum mastery. The students’ critical thinking skills, from trial I to trial II, are improved by 21.4%.

4. CONCLUSION

From the above discussion, this research concludes that: (1) The designed assessment instruments for assessing HOTS are deemed effective. This was evident from the results of individual and classical learning mastery, student activities (within the bounds set by a specific regulation), and students’ positive reception to the learning media. (2) From trial I to trial II, the improvement in students’ critical thinking skills was 0.32 on average, and the classical learning mastery was measured at 21.4%

ACKNOWLEDGMENT

We would like to express our gratitude to the Ministry of Research and Technology/National Research and Innovation Agency for their financial support to this research through a research grant Hibah Penelitian Kerja Sama Perguruan Tinggi (PKPT) 2020, No. Decree: No. 8/E1/KPT/2020. Also, we thank the research partners from Universitas Negeri Malang for their guidance and cooperation throughout the course of this present study.

REFERENCES


DEPARTMENT OF MATHEMATICS EDUCATION
STKIP BINA BANGSA MEULABOH
ACEH BARAT, INDONESIA
Email address: hmr.saputra@gmail.com

DEPARTMENT OF MATHEMATICS EDUCATION
STKIP BINA BANGSA MEULABOH
ACEH BARAT, INDONESIA
Email address: sitinababan28@gmail.com

DEPARTMENT OF MATHEMATICS EDUCATION
UNIVERSITAS NEGERI MALANG
MALANG, INDONESIA
Email address: cholis.sadijah.fmipa@um.ac.id

DEPARTMENT OF MATHEMATICS EDUCATION
UNIVERSITAS NEGERI MALANG
MALANG, INDONESIA
Email address: subanji.fmipa@um.ac.id