High School Students' Proficiency in Solving Higher-Order Mathematics Problems

Aminah Ekawati, Benny Nawa Trisna, Muhammad Royani
Universitas PGRI Kalimantan, Faculty of Social and Humanities,
Mathematic Education Study Program
Banjarmasin, Indonesia
aminahekawati@upk.ac.id
benny@upk.ac.id
hmroyani@upk.ac.id

Abstract. Higher-order thinking (HOT) skills, such as analysis (C4) and evaluation (C5), are important elements in mathematics learning. This study aims to assess the proficiency of high school students in solving mathematics problems at the C4 and C5 levels, analyse differences in proficiency based on gender, and identify differences in proficiency between students from SMAN A and SMAN B in Banjarmasin. With a quantitative descriptive approach, 129 grade XI students became the study subjects. The HOT essay test was used as an instrument and analysed using descriptive statistics and the Mann-Whitney test. The results showed that the average proficiency of students was in the moderate category (Mean = 60.04), with the average score of female students (60.16) slightly higher than male students (59.87) and a smaller variation in proficiency in females. Based on school, SMAN A students had a higher average score (62.68) than SMAN B (57.73), although this difference was not statistically significant ($p > 0.05$). These findings indicate the importance of developing more structured learning to improve students' high-level thinking skills. These results indicate that high school students' mathematical analysis and evaluation skills still need to be improved through a more targeted and varied educational approach.
Keywords: C4 (analysis), C5 (evaluation), gender, mathematics problems

1. Introduction

The enhancement of Higher-Order Thinking Skills (HOTS) among high school students has become a central focus in educational research. HOTS refers to the proficiency to think critically, creatively, analytically, and evaluatively—skills that are not only valuable in formal education but also in everyday life (N. , Erdiana & Panjaitan, 2023; N.; Tanudjaya & Doorman, 2020; C. P. These skills enable individuals to analyze and evaluate information deeply, going beyond mere memorization and understanding of basic facts, and can be applied to solve more complex problems (Anderson & Krathwohl, 2001). Higher-order thinking skills are important to prepare students for the challenges of the 21st century (Ahmadi, 2021 ; Minarti et al., 2023)but they can also support lifelong learning that allows individuals to continue to acquire and apply knowledge in a variety of evolving contexts.

In the revised Bloom's taxonomy, higher-order thinking consists of three levels, namely C4 (analysis), C5 (evaluation), and C6 (creative) (Anderson & Krathwohl, 2001).This study specifically focuses on analytical (C4) and evaluative (C5) proficiencies, which are critical in mathematics education. Analytical proficiency (C4) requires students to break down complex problems or information into simpler components and understand the relationships among those components. In contrast, evaluative proficiency (C5) demands that students make judgments based on specific criteria and provide clear reasons for their choices (Lewy et al., 2009). These proficiencies are essential for equipping students with the ability to think critically and make sound judgments based on available information, especially when solving more advanced mathematical problems.

Mathematics, as a discipline, inherently supports the development of HOTS. In mathematics, students are trained not only to understand procedures or formulas but also to analyze, evaluate, and solve more complex problems *(Ilyas et al., 2022)*. Through problem-solving in mathematics, students can sharpen their analytical and evaluative proficiencies, which are crucial for enhancing logical and

structured thinking. These proficiencies are not only relevant to mathematics itself but also to mastering other fields such as STEM (Science, Technology, Engineering, and Mathematics), which is increasingly important in addressing global challenges in the digital and globalized era.

Some studies have suggested that gender differences may influence students' proficiency in solving higher-order mathematical problems. For example, (He & Wong, 2021) found that male students tended to perform better on tasks requiring divergent thinking and creative problem-solving. However, Reinhold et al., (2020) showed that gender did not significantly impact students' proficiency in solving complex mathematical problems. On the other hand, research by (W. , Kusumaningsih et al., 2018)indicated that male students tended to have higher logical reasoning skills than female students. This suggests that gender may influence how students approach cognitive challenges and solve problems requiring analytical and evaluative thinking.

However, despite the emphasis on HOTS in mathematics education, few studies have specifically examined students' proficiency levels in analysis and evaluation when facing advanced mathematical problems. Therefore, this study aims to identify and evaluate the proficiency levels in analysis (C4) and evaluation (C5) of high school students in solving higher-order mathematical problems. The findings of this study are expected to provide valuable insights for the development of curricula and more effective teaching strategies, as well as assist mathematics teachers in designing lessons that can enhance students' HOTS, particularly when facing advanced mathematics exams or problems.

The school environment is also one-factor influencing students' proficiency. SMAN A and SMAN B are two favourite schools in Banjarmasin, with national UTBK rankings of 760 and 624, respectively. Although both schools use the same curriculum and have the same accreditation, the differences in UTBK rankings indicate variations in learning outcomes that are worth exploring. Thus, this study aims to: (i) assess the level of high school students' proficiency in solving mathematical problems at the C4 and C5 levels, (ii) analyze the differences in proficiency between male and female students, and (iii) identify differences in proficiency between students from SMAN A and SMAN B.

2. Method

Research Design

This study adopts a quantitative descriptive approach to evaluate high school students' high-level thinking skills in solving mathematical problems at levels C4 (analysis) and C5 (evaluation). It also compares students' proficiency based on gender and school. The study involved two senior high schools: SMAN A and SMAN B.

Research Subjects

The study subjects were grade XI students from two schools, namely SMAN A and SMAN B. The following are the details of the research subjects.

Table 1 Distribustion of Subject Research

No.	School	Class	Number of Students		Total
			Male	Female	
1	SMAN A	X1.1	11	16	27
2		X1.10	13	20	33
3	SMAN B	X1.2	11	21	32
4		X1.3	22	15	37
		Total			129

The consideration of selecting a sample in class X1 is because the topic used to measure students' mathematical proficiency has been studied in class X, and students in that class are willing to participate fully in all stages of the study.

Instrument and Collect Data

The main instrument in this study is a high-level thinking proficiency test comprising five essay questions categorized at levels C4 and C5 based on Bloom's Taxonomy. The instrument was validated by three experts and deemed feasible. Empirical testing confirmed that all questions are valid and meet reliability standards.

To prevent question leakage, the test was administered on the same day for all classes within each school, with only timing variations. Students were given 80 minutes to complete the test. Afterward, their answer sheets were collected, graded according to a predefined rubric, and analyzed to produce a descriptive profile of students' proficiency.

Data Analysis

The data obtained were analyzed using descriptive statistics to determine the average scores of students' proficiency in solving C4 and C5 level questions at both schools. Additionally, an inferential statistical analysis was conducted to compare the means and examine whether there were significant differences in high-level thinking skills between students at SMAN A and SMAN B, considering both school and gender factors, with the aid of SPSS 25.

3. Result and Discussion

The following presents the descriptive results of high school students' mathematical proficiency.

Table 2 Descriptive results of high school students' mathematical proficiency

	N	Minimum	Maximum	Mean	Std. Deviation
Male	57	24,41	100,00	59,87	19,78
Female	72	30,29	95,82	60,16	16,39
SMAN B	69	24,41	100,00	57,73	19,60
SMAN A	60	28,18	95,82	62,68	15,45
All Students	129	24,41	100,00	60,04	17,90

Based on table 2. it is obtained that the average mathematics proficiency at the C4 and C5 levels of female students (60.16) is slightly higher than male students (59.87). In addition, the variation in proficiency in the male group is greater (Standard Deviation: 19.78) than female (Standard Deviation: 16.39), this shows that the proficiency of female students is more uniform. In terms of school, SMAN A students have a higher average proficiency (62.68) than SMAN B students (57.74), meaning that in general, SMAN A students have better performance in mathematics analysis and evaluation proficiency than SMAN B students. The proficiency of SMAN A students tend to be more uniform than those of SMAN B, this is indicated by the value of the variation in proficiency at SMAN B being greater (Standard Deviation: 19.60) than SMAN A (Standard Deviation: 15.45). Overall, female students and students from SMAN A showed higher and more uniform academic performance compared to male students and students from SMAN B. The higher variation in proficiency in the male group and SMAN B students reflects a more diverse distribution of proficiency.

Next, a test was conducted to see the significance of the differences based on school and gender. However, a prerequisite test was conducted first. The prerequisite tests carried out were the normality test and the homogeneity test. The normality test used the Kolmogorov-Smirnov test, here are the results of the normality test.

Table 3. Kolmogorov-Smirnov normality test results

		SMAN B	SMAN A	Male	Female
N		69	60	57	72
Normal Parameters[a,b]	Mean	57,7380	62,6800	59,8746	60,1649
	Std. Deviation	19,60	15,45	19,78	16,39
Most Extreme Differences	Absolute	0,127	0,168	0,110	0,111
	Positive	0,127	0,125	0,106	0,111
	Negative	-0,093	-0,168	-0,110	-0,103
Test Statistic		0,127	0,168	0,110	0,111
Asymp. Sig. (2-tailed)		.007[c]	.000[c]	.082[c]	.027[c]

Based on Table 3. it was found that the data from the SMAN B, SMAN A, and female student groups were not normally distributed, with p values of 0.007, 0.000, and 0.027 ($p < 0.05$), respectively. In contrast, the data from the male student group showed a normal distribution, with a p-value of 0.082 ($p > 0.05$). This indicates that the null hypothesis, namely that the data comes from a normal distribution, is rejected for the SMAN B, SMAN A, and female groups but accepted for the male group. Because only 1 group meets the normality test, the homogeneity test was not carried out. Further testing uses a non-parametric statistical test, the Mann-Whitney test. The following are the results of the Mann-Whitney test to see the differences in the proficiency of male and female students.

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Nilai is the same across categories of Sex.	Independent-Samples Mann-Whitney U Test	.783	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Figure 1. Hypothesis test Summary for sex categories

Based on Figure 1, the significance value (Sig.) is 0.783, which is more than 0.05. Therefore, the null hypothesis (H_0) is accepted or there is no significant difference in the proficiency of male and female students. Furthermore, based on the school, the following is obtained.

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Nilai is the same across categories of Sekolah.	Independent-Samples Mann-Whitney U Test	.072	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Figure 2. Hypothesis test Summary for school categories

Based on Figure 2, a significance value (Sig.) of 0.072 was obtained, which is more than 0.05. Therefore, the null hypothesis (H_0) is accepted, or there is no significant difference in students' proficiency at SMAN A and SMAN B. The school environment can influence students' proficiency in higher-order thinking skills. SMAN A, despite having a UTBK ranking of 760, shows better performance variability in the proficiency of analytical and evaluative skills (C4 and C5) compared to SMAN B, which has a UTBK ranking of 624. Although the average proficiency of students at SMAN A is higher, this difference is not strong enough to be statistically significant at the 95% confidence level. Thus, while there is a difference in the average mathematics proficiency at C4 and C5 levels between students at SMAN A and SMAN B, statistically, both schools exhibit relatively comparable performance in these proficiencies. The observed differences may be attributed to random factors or variations within the sample.

In general, the average mathematics proficiency of high school students at the C4 and C5 levels is 60.036, which falls within the moderate category. This indicates that, overall, students' abilities in mathematical analysis and evaluation require further enhancement. Despite some variation between individual students and schools, the results suggest that high school students' proficiency in these higher-order thinking skills is still developing and needs improvement to reach higher levels of mastery. According to study by Widana (2018) higher-order thinking skills are key to developing a deep understanding of mathematics subjects and helping students to be more critical in solving complex problems. A study by Azid et al., (2022) also found that developing these higher-order cognitive skills can increase students' engagement in learning and contribute to better academic achievement. Therefore, more structured efforts are needed in mathematics learning to improve high school students'

analytical and evaluative skills and better prepare them to face academic and real-life challenges that require deep understanding and problem-solving.

From an educational perspective, these results indicate that the learning methods used in both schools are likely to produce balanced cognitive achievements among students, especially in mathematical, analytical, and evaluation skills. In the context of educational development, these results emphasize the importance of maintaining consistent teaching quality and standards to ensure that the achievement of higher-order thinking skills is evenly distributed across institutions. However, further research may be needed to identify other factors that contribute to students' achievement in this cognitive aspect, for example, differences in assessment methods, intensity of problem exercises, and access to additional learning resources. The similarity of results in both groups can also be a basis for educators to explore more varied methods in improving students' critical thinking skills, especially in more complex questions at the C5 level.

4. Conclusion

The results showed that the average mathematics proficiency of high school students at levels C4 and C5 was 60.04, and they were included in the moderate category. Descriptively, female students had a slightly higher average proficiency (60.16) than male students (59.87), with smaller variations in proficiency in females. Based on school, the average proficiency of students at SMAN A (62.68) was higher than that of SMAN B (57.73), and the variation in the proficiency of students at SMAN A was smaller than that of SMAN B. However, the Mann-Whitney statistical test showed that differences in proficiency based on gender and school were not statistically significant ($p>0.05$). So, the differences seen descriptively cannot be concluded as real differences.

Based on the results obtained, the mathematics proficiency of high school students at levels C4 and C5 needs to be improved because this proficiency is important to face global challenges. This improvement can be done through student-centred learning, innovative learning methods, models, or strategies, and evaluations designed to develop analytical and evaluative thinking skills.

Reference

Ahmadi, A. (2021). Teachers as Ethnographers: Narrative Study of Inquiry of Indonesian Teachers Assigned to Teach in Remote Areas. . *European Journal of Educational Research*, *10*(1), 115–126.

Anderson, L. W. , & Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching, and Assessing. A Revision of Bloom's Taxonomy of Educational Objectives (Complete Edition*. Addison Wesley Longman, Inc.

Azid, N., Ali, R. M., El Khuluqo, I., Purwanto, S. E., & Susanti, E. N. (2022). Higher order thinking skills, school-based assessment and students' mathematics achievement: Understanding teachers' thoughts. *International Journal of Evaluation and Research in Education*, *11*(1), 290–302. https://doi.org/10.11591/ijere.v11i1.22030

Erdiana, N. , & Panjaitan, S. (2023). How is HOTS Integrated into the Indonesian High School English Textbook? *Studies in English Language and Education*, *10*(1), 60–77.

Erdiana, N., & Panjaitan, S. (2023). How is HOTS Integrated into the Indonesian High School English Textbook? *Studies in English Language and Education*, *10*(1), 60–77. https://doi.org/10.24815/siele.v10i1.26052

He, W. J. , & Wong, W. C. (2021). Gender Differences in the Distribution of Creativity Scores: Domain-Specific Patterns in Divergent Thinking and Creative Problem Solving. *Frontiers in Psychology*.

Ilyas, M. , Meiyani, E. , Ma'rufi, M. , & Kaewhanam, P. (2022). Improving students' ability in learning mathematics by using the science, technology, engineering, and mathematics (STEM) approach. *Frontiers in Education*.

Kusumaningsih, W. , Darhim, Herman, T. , & Turmudi. (2018). Gender differences in algebraic thinking ability to solve mathematics problems. *Journal of Physics*, 1013.

Kusumaningsih, W., Darhim, Herman, T., & Turmudi. (2018). Gender differences in algebraic thinking ability to solve mathematics problems. *Journal of Physics: Conference Series*, *1013*, 012143. https://doi.org/10.1088/1742-6596/1013/1/012143

Lewy, L. , Zulkardi, & Aisyah, N. (2009). Pengembangan soal untuk mengukur kemampuan berpikir tingkat tinggi pokok bahasan barisan dan deret bilangan di kelas IX akselerasi SMP Xaverius Maria Palembang. *Jurnal Pendidikan Matematika*, *3*(2), 15–28.

Minarti, I. B. , Dzakiy, M. A. , & Nilautama, D. (2023). The Effect of STEM (Science, Technology, Engineering, and Mathematics) Based Learning Approach on Critical Thinking Skills and Cognitive Learning Outcomes of Class X SMA Negeri 1. *At-Tasyrih: Jurnal Pendidikan Dan Hukum Islam*, *8*(2), 126–136.

Reinhold, F., Hofer, S., Berkowitz, M., Strohmaier, A., Scheuerer, S., Loch, F., Vogel-Heuser, B., & Reiss, K. (2020). The role of spatial, verbal, numerical, and general reasoning abilities in complex word problem solving for young female and male adults. *Mathematics Education Research Journal*, *32*(2), 189–211. https://doi.org/10.1007/s13394-020-00331-0

Tanudjaya, C. P. , & Doorman, M. (2020). Examining higher order thinking in Indonesian lower secondary mathematics classrooms. *Journal on Mathematics Education, 11*(2), 277–300.

Tanudjaya, C. P., & Doorman, M. (2020). Examining higher order thinking in Indonesian lower secondary mathematics classrooms. In *Journal on Mathematics Education* (Vol. 11, Issue 2, pp. 277–300). Sriwijaya University. https://doi.org/10.22342/jme.11.2.11000.277-300

Widana, I. W. (2018). Higher Order Thinking Skills Assessment towards Critical Thinking on Mathematics Lesson. *International Journal of Social Sciences and Humanities (IJSSH)*. https://doi.org/10.29332/ijssh.v2n1.74