Developing Higher-Order Thinking Skills in Mathematics Learning through Challenging Mathematical Task: A Single Subject Study on Gifted Students

Kadir Kadir[1*], Dedek Kustiawati[2] ,and Hania Rahmah[3]

[1, 2] Department of Mathematics Education, UIN Syarif Hidayatullah Jakarta, Indonesia
[3] Master of Mathematics Education Study Program, Universitas Pendidikan Bandung, Indonesia
[*] kadir@uinjkt.ac.id

Abstract. This research aims to analyze the higher-order mathematical thinking skills of gifted students before, during, and after being given challenging mathematical tasks. This research was conducted at a State Madrasah Tsanawiah in Depok City. This study used a single-subject research design (A-B-A). The subjects of this study were three students who were included in the gifted student category. The research instrument used was a test. The results of this study revealed that the highest- order mathematical thinking skills of all subjects were obtained during the intervention condition (B). The higher-order mathematical thinking skills of gifted 1 and 3 students at baseline 2 (A2) were higher than baseline 1 (A1), while higher-order mathematical thinking skills of gifted 2 students at baseline 2 (A2) were lower than baseline 1 (A1). In general, the conclusion of this study is that the giving of challenging mathematical tasks consisting of stages of relate, investigate, communicate, evaluate, and create can improve higher order mathematical thinking skills of gifted students, which includes analyzing, evaluating, and creating. This study suggests that integrating challenging mathematical tasks into planning, learning interventions, and assessment supports further efforts to improve the mathematical thinking ability of gifted students.

Keywords: higher-order thinking skills, single-subject research A-B-A, challenging mathematical tasks, gifted students.

1 Introduction

Higher-order thinking skills are essential for applying mathematical knowledge to tasks such as reasoning, problem-solving, communication, inquiry, and conceptual understanding. These skills are crucial not only in learning mathematics but also as core competencies for navigating life, work, and problem-solving across various fields. Many challenges individuals face, both within and beyond mathematics, stem from deficiencies in these skills [1]. Therefore, it is vital for students to develop and apply higher-order thinking skills in both education and daily life. Brookhart (2010) categorizes higher-order thinking skills into three key areas: transfer, critical thinking, and

T. A. Aziz et al. (eds.), *Proceedings of the 4th Science and Mathematics International Conference (SMIC 2024)*, Advances in Physics Research 11,
https://doi.org/10.2991/978-94-6463-624-6_15

problem-solving. As transfer, these skills enable students not only to recall learned information but also to interpret and apply it in new contexts. In terms of critical thinking, higher-order thinking encompasses reasoning, questioning, investigating, observing, describing, comparing, connecting ideas, identifying complexity, and exploring perspectives. Finally, as problem-solving, it involves tackling challenges through non-routine procedures and innovative thinking [2].

Mathematical thinking is categorized by complexity into two types: low-level thinking, which involves memorization, comprehension, and application, and high-level thinking, which encompasses deep understanding, forming conjectures, drawing analogies and generalizations, logical reasoning, problem-solving, representation, abstraction, mathematical creativity, proof, and non-procedural communication and connections. These higher-level skills are essential components of advanced mathematical thinking [3]. Krathwohl categorizes higher-order thinking indicators into levels: (1) analysis, involving cognitive processes of distinguishing, organizing, and connecting; (2) evaluation, through examining and critiquing; and (3) creation, which includes generalizing, planning, and producing [4].

The development of tasks aimed at enhancing higher-order thinking skills is a defining feature of learning for gifted students. This aligns with Davis (2012), who notes that the use of higher-order thinking is a core characteristic of highly intelligent students [5]. More specifically, a student with exceptional mathematical ability is identified as mathematically gifted. Such students are poised to contribute significantly to various aspects of human advancement. This assumption is empirically supported by decades of research from the Study of Mathematically Precocious Youth (SMPY), which shows that individuals with high-level mathematical thinking skills are pivotal in 21st-century life, particularly in the economic sector [6].

Gifted students engage in creative thinking when solving higher-order problems, demonstrating strategies to connect and transform ideas, and generate unique solutions to novel challenges—approaches not typically seen among average students [7]. Given the unique strengths of gifted students, it is crucial to support their development. However, the National Council of Teachers of Mathematics (NCTM) in Sheffield has noted that gifted students often remain overlooked in efforts to fully cultivate their potential [8].

In Indonesia, the Association of Special Education Organizers, Developers, and Supporters for Gifted Students (CI+BI Association) reports that 1.3 million school-age children are potentially gifted, yet only 9,500 receive specialized services such as accelerated programs [9]. Some gifted individuals may experience underachievement, a phenomenon often emerging in high school. Addressing this issue involves nurturing their motivation and abilities in ways that fully realize their potential [10]. Gifted students are valuable assets for the future advancement of society; therefore, their potential must be nurtured through appropriate educational programs. However, suitable programs remain limited and have yet to fully and effectively support the needs of gifted students.

An effective solution to support the optimal development of gifted students, particularly those gifted in mathematics, is to provide challenging mathematical tasks. Such tasks are instrumental in enhancing the abilities of gifted students. Singer et al. (2016) describes the Sheffield model as incorporating heuristic steps that foster open-ended problem-solving and problem-posing as challenging tasks. These steps include: (a) connecting the task to students' prior knowledge, (b) investigating the problem, (c) evaluating findings, (d) communicating results, and (e) generating new questions for further exploration [11]. The steps above are non-sequential, allowing students to approach the problem non-linearly and explore solutions creatively. Figure 1 provides an illustration of these five steps.

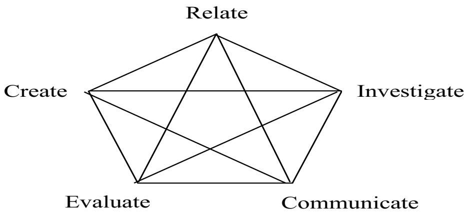


Fig. 1. Steps of challenging mathematical task

The potential of gifted students can be cultivated by engaging them in challenging mathematical tasks that require creative, critical, and reflective thinking to derive solutions. Therefore, this study aims to explore the development of Higher-Order Thinking Skills (HOTS) in mathematics learning using challenging mathematical tasks. Using a single-subject approach with gifted students, the research seeks to understand how specially designed mathematical tasks can enhance advanced skills in analysis, evaluation, and problem-solving. The focus on gifted students offers insights into teaching strategies and approaches that can effectively support the achievement of HOTS.

2 Methods

This study employs a single-subject research method based on behavior modification theory, measuring variables in the same subject across different conditions. The conditions in this study include baseline (A) and intervention (B) conditions. Baseline conditions involve measuring the target behavior in a natural setting before any intervention, while intervention conditions measure the target behavior during the application of the intervention [12].

This study employs an A-B-A reversal design. Initially, the target behavior is continuously measured under the first baseline condition (A1) for a set period, followed by measurement during the intervention condition (B). A second baseline condition (A2) is then applied to control for the intervention, allowing for conclusions about a functional relationship between the independent and dependent variables [12].

In the baseline condition, gifted students were administered a higher-order thinking skills (HOTS) instrument 3-5 times until data stabilized. During the intervention condition, students first completed a challenging mathematical task (CMT) followed by the HOTS instrument. Once HOTS data stabilized in the intervention phase, CMT administration ceased, and after a period, the HOTS instrument was re-administered 3-5 times. This approach allows conclusions on the effectiveness of CMT in enhancing HOTS in gifted students.

The study participants were three students identified as gifted based on Gagne's IQ criteria, which defines giftedness within an IQ range of 112 to 160[13]. The research subjects' identities are as follows: (1) G1, female, 13 years old, IQ 128, a national semi-finalist in the math and science Olympiad (2014) and first place in the Qur'an Tahfizh competition at the Regency level (2014); (2) G2, male, 14 years old, IQ 114, first place in the National Madrasah Science Competition (2018) and City Madrasah Science Competition (2014, 2017); (3) G3, female, 14 years old, IQ 116, first place in the LKBBT District Level (2017).

The research instrument used was a test designed to measure higher-order mathematical thinking skills across the dimensions of analysis, evaluation, and creation. Student responses were assessed using a rubric adapted from Bosch [14]. The HOTS test was validated for content and construct through expert judgment. This test, developed as a 33-item essay format, measures indicators across the dimensions of analysis, evaluation, and creation. Ensuring research instruments' quality requires both validity and reliability. The content validity of the HOTS test was determined using Lawshe's Content Validity Ratio (CVR) method [15]. Content validity was assessed by a panel of 8 mathematics experts, comprising 4 mathematics lecturers and 4 mathematics teachers. CVR-Lawshe formula.

$$CVR = \frac{\left(N_e - \frac{N}{2}\right)}{\frac{N}{2}} \tag{1}$$

Here, CVR (Content Validity Ratio) represents content validity, where N is the total number of panelists and Ne is the number of panelists indicating the item as essential. Table 1 presents the results of the content and empirical validity of the HOTS test items.

Table 1. The results of the content validity and the item validity of the CTS test

Dimension	Indicators	Item	CVR	Min	Conclusion	Decision
Analyze	Identify the elements contained in a relationship	1	0.71	0.99	Revised	Used
		2	1.00	0.99	Valid	Used
		3	1.00	0.99	Valid	Used
		4	1.00	0.99	Valid	Used
		5	1.00	0.99	Valid	Used
	Verifying the accuracy of element relationships and interactions within the problem and making decisions as a solution.	6	1.00	0.99	Valid	Used
		7	1.00	0.99	Valid	Used
		8	1.00	0.99	Valid	Used
	Reorganize certain rules related to how to solve the problem	9	0.71	0.99	Revised	Used
		10	1.00	0.99	Valid	Used
		11	1.00	0.99	Valid	Used
Evaluate	Make judgments or considerations based on criteria and standards	12	0.71	0.99	Revised	Used
		13	1.00	0.99	Valid	Used
		14	1.00	0.99	Valid	Used
		15	1.00	0.99	Valid	Used
		16	1.00	0.99	Valid	Used
		17	1.00	0.99	Valid	Used
		18	1.00	0.99	Valid	Used

	Accepting or rejecting a statement based on pre-determined criteria	19	1.00	0.99	Valid	Used
	Detect the procedural appropriateness of a problem	20	1.00	0.99	Valid	Used
		21	1.00	0.99	Valid	Used
Create	Assemble elements or parts into a new structure	22	1.00	0.99	Valid	Used
		23	0.71	0.99	Revised	Used
		24	1.00	0.99	Valid	Used
		25	1.00	0.99	Valid	Used
	Devising a way to solve the problem	26	1.00	0.99	Valid	Used
		27	1.00	0.99	Valid	Used
		28	1.00	0.99	Valid	Used
		29	1.00	0.99	Valid	Used
	Generalize an idea	30	1.00	0.99	Valid	Used
		31	1.00	0.99	Valid	Used
		32	1.00	0.99	Valid	Used
		33	0.71	0.99	Revised	Used

The analysis results in Table 1 indicate that all 33 items are content valid, confirming the instrument's suitability for collecting HOTS data.

The analysis technique for single-subject experimental research involves simple descriptive statistical analysis and visual analysis. Descriptive analysis includes trend measures such as mean and median. Visual condition analysis comprises six components: condition length, trend estimation, trend stability, data trace, stability level and range, and level of change. Visual analysis between conditions includes five components: the number of variables changed, trend shifts and effects, stability changes, level changes, and data overlap [12].

3 Results and Discussion

The percentage of HOMT scores for subjects G1, G2, and G3 across baseline 1 (A1), intervention (B), and baseline 2 (A2) conditions is shown in Figure 2. As illustrated, G1, G2, and G3 achieved the highest average HOTS scores during the intervention phase. The average HOTS scores for G1 and G3 in baseline 2 were higher than in baseline 1, while G2's average HOTS in baseline 2 was lower than in baseline 1.

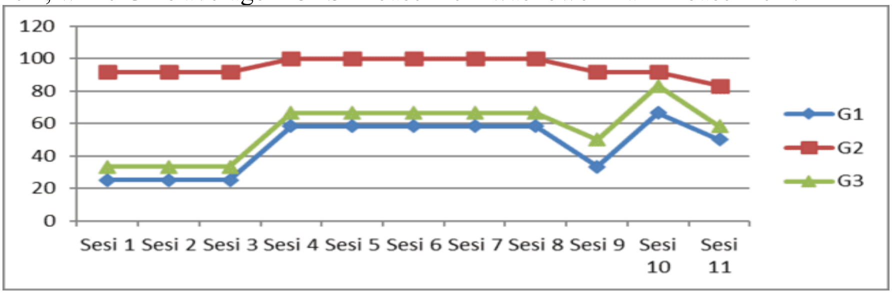


Fig. 2. Summary of HOTS percentage data across all conditions (A-B-A) for subjects G1, G2, and G3

3.1 Higher Order Thinking Skills in baseline condition 1 (A1)

Baseline 1 (A1) represents the initial data collection stage, consisting of 3 sessions, each with a maximum duration of 90 minutes. In each session, students were given 3 HOTS questions covering indicators of analyzing, evaluating, and creating. The HOTS topics were geometry in session 1, statistics and algebra in session 2, and numbers in session 3. Examples of questions and responses for the analyzing indicator by gifted students G1, G2, and G3 in the baseline 1 (A1) condition are shown in Figure 3.

Question/Problem for analyzing indicators: "In a village, there is a mother who has 9 children with always the same age difference of 15 months. If the age of the first child is now 6 times the age of the last child, then:

1) List the elements needed to determine the age of the fifth child!
2) Suppose x is the age of the last child, make a math sentence based on the problem above!
3) Find the age of the fifth child!"

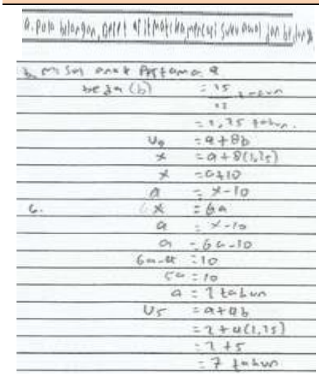
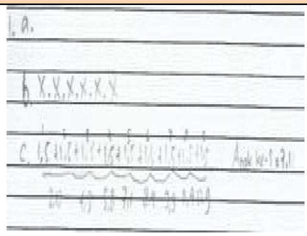
Answer G1:	Answer G2:	Answer G3:
		

Fig. 3. Example answers G1, G2 and G3 baseline 1 (A1) for analyzing indicators

As shown in Figure 3, subject G1 was unable to correctly identify the problem-solving elements, establish relationships, or provide a solution. G1 only formulated a simple mathematical sentence, stating, "the age of the last child is six times the age of the first child," without incorporating additional information, such as the age differences among the 9 children. Subject G2 successfully identified the necessary elements, determined relationships, developed a solution, and applied the concept of an arithmetic sequence to find the age of the fifth child. Subject G3, however, did not attempt the problem.

Figure 4 provides examples of questions and responses to HOTS questions from gifted students G1, G2, and G3 in the baseline 1 (A1) condition for the evaluation indicator.

Question/Problem for evaluating: "A block-shaped water tank that is 1/3 filled has sides whose areas are 20 dm^2, 48 dm^2, and 60 dm^2. If 200 liters of water is put into the basin, will the water in the basin spill out? Prove it!"

Answer G1:	Answer G2:	Answer G3:

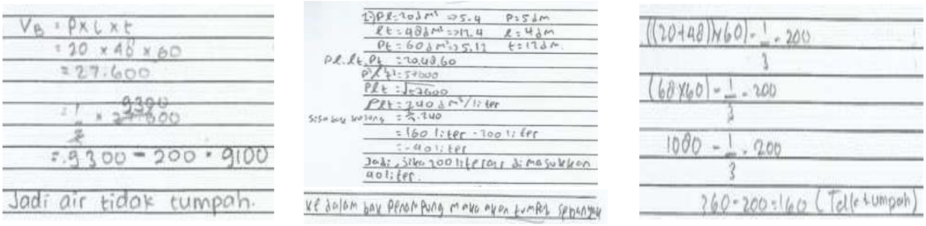


Fig. 4. Example answers G1, G2 and G3 baseline 1 (A1) for evaluating indicators

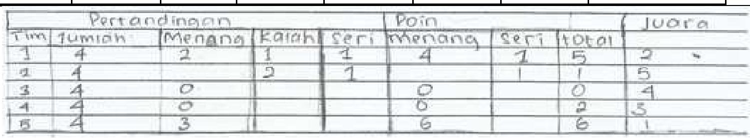
Based on Figure 4, subject G1 misinterpreted the problem, mistaking the side area of the prism for its volume and incorrectly calculating whether water would spill from the reservoir by immediately subtracting 200 liters from one-third of the prism's volume. Subject G2, however, provided the correct answer with a sound rationale, interpreting "side area of the prism" accurately. G2 determined the prism's dimensions by identifying numbers that satisfy the multiplication requirements for length, width, and height. Rather than directly multiplying these dimensions to find the volume, G2 calculated the product of the side areas and applied the square root to derive the volume formula. G2 then correctly identified the remaining unfilled volume as two-thirds of the prism's volume, determining whether water would spill by subtracting 200 liters from this unfilled volume. Meanwhile, subject G3 made calculation errors by adding two side areas, multiplying by the side area, then by one-third, and finally subtracting 200, resulting in an incorrect answer.

Figure 5 provides examples of questions and responses to HOTS questions from gifted students G1, G2, and G3 in the baseline 1 (A1) condition for the creation indicator.

Question/Problem for creating indicators: "In a Hockey match, a team gets 2 points if they win, 1 point if they draw, and 0 points if they lose. There are 5 teams that play each other once each. A computer failed to print some parts of the following match results. Complete the table with the correct numbers to get the 1st to 5th place! (If you get teams with the same total points, the team with the most wins is considered superior)!

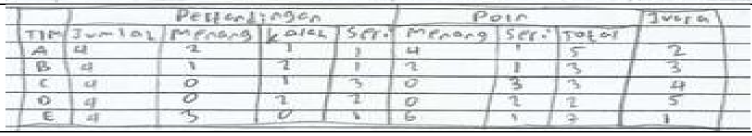
Team	Match				Points			Champion
	Total	Win	Lose	Series	Win	Series	Total	
A	4	2	1	1	4	1	5	
B	4		2	1				
C	4	0			0			4
D	4	0					2	
E	4	3						

Answer G1:



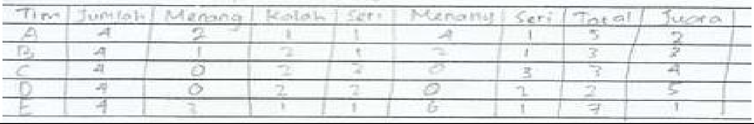
Tim	Pertandingan: Jumlah	Menang	Kalah	Seri	Poin: Menang	Seri	Total	Juara
1	4	2	1	1	4	1	5	2
2	4		2	1		1	1	5
3	4	0			0		0	4
4	4	0			0		2	3
5	4	3			6		6	1

Answer G2:



Tim	Pertandingan: Jumlah	Menang	Kalah	Seri	Poin: Menang	Seri	Total	Juara
A	4	2	1	1	4	1	5	2
B	4	1	2	1	2	1	3	3
C	4	0	1	3	0	3	3	4
D	4	0	2	2	0	2	2	5
E	4	3	0	1	6	1	7	1

Answer G3:



Tim	Jumlah	Menang	Kalah	Seri	Menang	Seri	Total	Juara
A	4	2	1	1	4	1	5	2
B	4	1	2	1	2	1	3	3
C	4	0	2	2	0	3	2	4
D	4	0	2	2	0	2	2	5
E	4	3	1	1	6	1	7	1

Fig. 5. Example answers G1, G2 and G3 baseline 1 (A1) for creating indicators

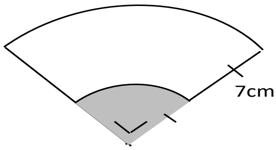
The results in Figure 5 show that Subject G1 seems to create a new structure inaccurately because there are still some empty columns. Subject G2 can create a new structure that is precise and complete based on the elements available. Furthermore, subject G3 was able to complete all empty columns, but there was an error in determining the number of defeats of team C. Subject G3 also did not make the top of the table in the problem, so it was difficult to distinguish between the number of wins, losses, and draws with the points earned.

3.2 Higher Order Thinking Skills in intervention condition (B)

The intervention in this study consisted of five sessions of challenging mathematical tasks (CMT). The CMT topics and HOTS instruments covered geometry in sessions 4 and 5, statistics in session 6, algebra in session 7, and numbers in session 8. HOTS ability scores were collected immediately following each CMT intervention.

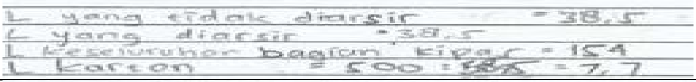
Figure 6 provides examples of questions and responses to HOTS questions from gifted students G1, G2, and G3 in the intervention (B) condition for the analyzing indicator.

Question/Problem for analyzing indicators:
Ahmad has a rectangular piece of cardboard with a size of 20x25 cm, which will be made into a fan shape (the part that is not shaded), as shown below!

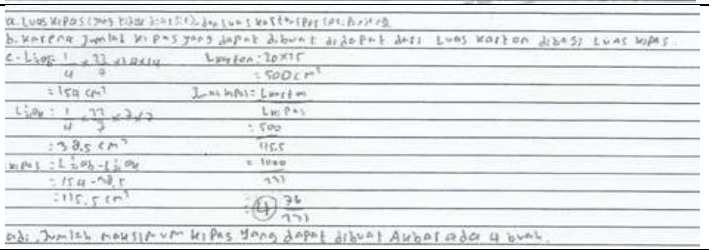


a. List the elements needed to determine the maximum number of fans that Ahmad can make!
b. Explain the relationship between the elements in (a) and the maximum number of fans that Ahmad can make!
c. Based on the elements mentioned in (a) and (b), determine the maximum number of fans that Ahmad can make!

Answer G1:



Answer G2:



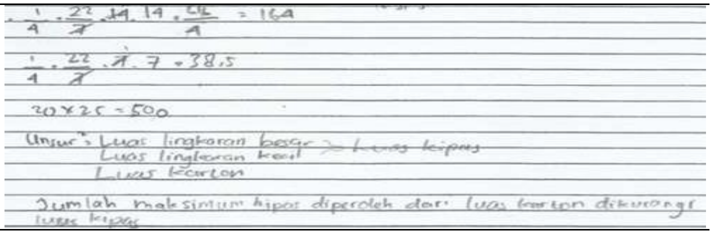
Answer G3:	

Fig. 6. Example answers G1, G2 and G3 in intervention (B) for analyzing indicators

The results in Figure 6 indicate that subject G1 listed the areas of the shaded part, unshaded part, fan, and carton without explaining the underlying concepts, leading to an incorrect answer. Subject G2 systematically and accurately determined the maximum number of fans by calculating the area of one-fourth of the large circle, 1.4 small circles, the fan, and the carton, then dividing the carton's area by the fan's area. The answer was clear, complete, and well-organized. Subject G3 documented calculations and identified necessary elements and their relationships, though the response was still incomplete.

Figure 7 provides examples of questions and responses to HOTS questions from gifted students G1, G2, and G3 in the intervention (B) condition for the evaluation indicator.

Question/Problem for evaluating indicators: "Cahya arranged five equal-sized squares into a flat shape with an area of 405 cm2 and a minimum perimeter. It is 1 m long rope that is enough to go around all sides of the flat! Prove it!"

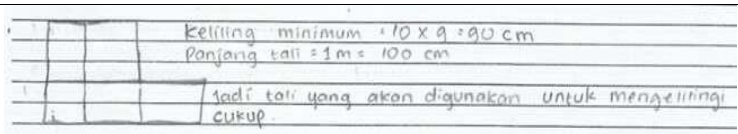
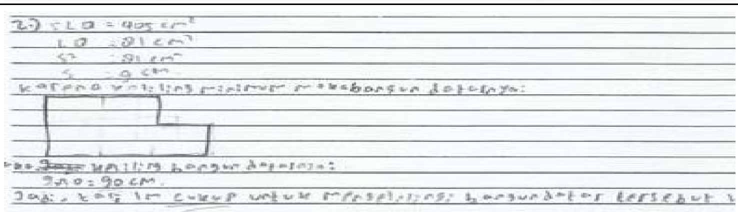
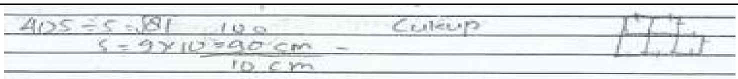
Answer G1:	
Answer G2:	
Answer G3:	

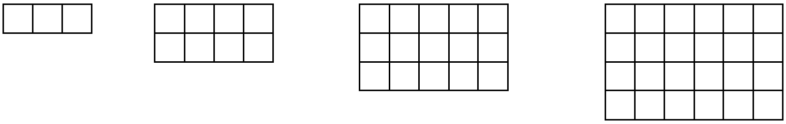
Fig. 7. Example answers G1, G2, and G3 intervention (B) for evaluating indicators

The results in Figure 7 indicate that subject G1 provided a correct answer but failed to explain how the side length of the square was determined or the meaning of the shape drawn. Subject G2 offered a correct and well-reasoned answer, calculating the square's area from the problem data, determining the side length, illustrating the shape, and calculating the minimum perimeter. In contrast, subject G3 only documented the calculation without explaining what was being calculated or the significance of the shape drawn.

Figure 8 provides examples of questions and responses to HOTS questions from gifted students G1, G2, and G3 in the intervention (B) condition for the creation indicator.

Question/Problem for creating indicators: Find the number of squares in the nth structure!

1st structure 2nd Structure 3rd Structure 4th Structure



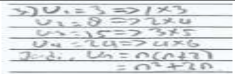
Answer G1:	Answer G2:	Answer G3:
		

Fig. 8. Example answers G1, G2, and G3 intervention (B) for creating indicators

The findings in Figure 8 reveal that subject G1 only identified the nth term without explaining the reasoning, obscuring the generalization process. Subject G2 provided a complete and correct generalization, detailing the number of squares in each structure, identifying the number pattern, and determining the nth arrangement. Subject G3 offered ideas and a generalization, but the solution steps were less precise, with the nth term presented without explanation, thus lacking a clear generalization process.

3.3 Higher Order Thinking Skills in baseline condition 2 (A2)

The baseline 2 (A2) condition represents the final phase in the A-B-A single-subject research design, conducted after completing the CMT intervention. Baseline 2 includes three sessions following the same procedure as baseline 1, where students answer three HOTS questions per session within a 90-minute limit. The HOTS topics covered were geometry in session 9, statistics and algebra in session 10, and numbers in session 11.

Figure 9 provides examples of questions and responses to HOTS questions from gifted students G1, G2, and G3 in baseline 2 (A2) for the analyzing indicator.

Question/Problem for analyzing indicators: " Farhan used 1/3 of his money and lost 2/3 of the rest. The remaining money is now only IDR 12,000.

1) If x is Farhan's first money, make a math sentence from the above information!
2) Determine the amount of Farhan's money by using the math sentence in point (a)!

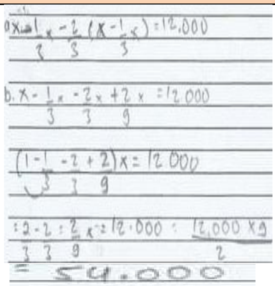
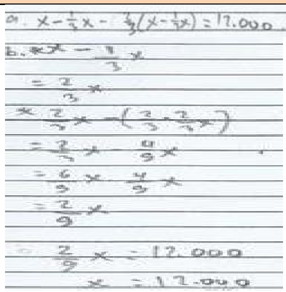
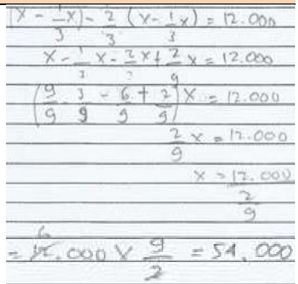
Answer G1:	Answer G2:	Answer G3:
		

Fig. 9. Example answers G1, G2 and G3 baseline 2 (A2) for analyzing indicators

The findings in Figure 9 indicate that subject G1 accurately identified specific rules, formulated mathematical sentences, and solved the resulting equations correctly. Subject G2 also identified the rules accurately and completely but made an error in the final calculation. Subject G3 successfully identified the rules, formulated mathematical sentences, and accurately solved the equations formed.

Figure 10 provides examples of questions and responses to HOTS questions from gifted students G1, G2, and G3 in the baseline 2 (A2) condition for the evaluation indicator.

Question/Problem for evaluating: "The numbers 2, 3, 4, 7, and 8 will form an odd number consisting of five numbers with no repeating numbers. Is the statement "the difference between the largest and smallest number is divisible by 6" acceptable? Prove it!"

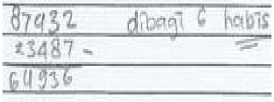
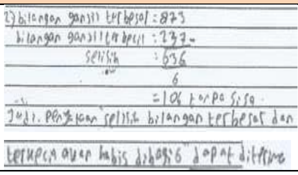
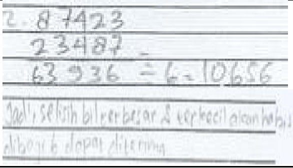
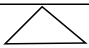
Answer G1:	Answer G2:	Answer G3:
		

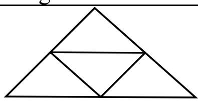
Fig. 10. Example answers G1, G2 and G3 baseline 2 (A2) for evaluating indicators

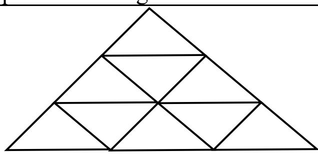
The findings in Figure 10 indicate that subject G1 provided a correct answer but lacked clear and complete reasoning. Subject G2's response was imprecise, with insufficient justification for accepting or rejecting the statement. Subject G3 answered correctly and indicated whether the statement in the problem was accepted but only documented the calculation without clarifying what was being calculated.

Figure 11 presents examples of questions and responses to HOTS questions from gifted students G1, G2, and G3 in the baseline 2 (A2) condition for the creation indicator.

Question/Problem for creating indicators: "Look at the equilateral triangle below!


1st pattern


2nd pattern


3rd pattern

Given that the length of one side of the triangle in the 2nd pattern is 16 cm, what is the perimeter of the triangle in the nth pattern?"

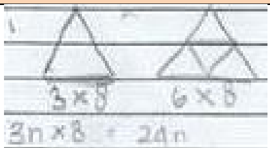
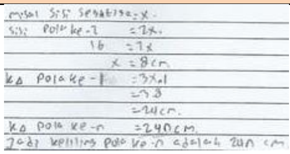
Answer G1:	Answer G2:	Answer G3:
		

Fig. 11. Example answers G1, G2 and G3 baseline 2 (A2) for creating indicators

The findings in Figure 11 indicate that subject G1 provided a generalization by determining the perimeter of the 1st and 2nd pattern triangles but did not explain the process for finding the side length. Subject G2 identified the side length and perimeter for the 1st pattern and then directly calculated the perimeter for the nth pattern, making the generalization process less apparent. Subject G3 demonstrated the ability to generalize an idea but showed imprecision in the solution steps, only recording the perimeter of the nth pattern triangle without clarifying the reasoning, which obscured the generalization process.

Table 2. Summary of visual analysis results under conditions

	Subject G1			Subject G2			Subject G3		
Condition	A_1	B	A_2	A_1	B	A_2	A_1	B	A_2
Length	3	5	3	3	5	3	3	5	3
Direction	Flat	Flat	Increase	Flat	Flat	Decrease	Flat	Flat	Increase
Stability (%)	Stable 100	Stable 100	Variable 33.33	Stable 100	Stable 100	Variable 66.67	Stable 100	Stable 100	Variable 66.67
Data Trace	Flat	Flat	Increase	Flat	Flat	Decrease	Flat	Flat	Increase
Stability	Stable	Stable	Variable	Stable	Stable	Stable	Stable	Stable	Variable
Level and Range (%)	25-25	58.3-58.33	33.3-66.7	91.7-91.7	100-100	91.67-83.3	33.3-33.3	66.7-66.7	50-83.33
Level Changes (%)	25-25 (0)	58.3-58.33 (0)	50-33.33 (+16.67)	91.7-91.7 (0)	100-100 (0)	91.67-83.3 (-17.67)	33.3-33.33 (0)	66.6-66.67 (0)	58.33-50 (+8.33)

Further visual analysis between conditions is presented in Table 3.

Table 3. Summary of visual analysis results between conditions

	Subject G1		Subject G2		Subject G3	
Comparison of Conditions	B/A_1 2 : 1	A_2/B 3 : 2	B/A_1 2 : 1	A_2/B 3 : 2	B/A_1 2 : 1	A_2/B 3 : 2
Number of Variables	1	1	1	1	1	1
Changes in Directional Tendencies & The Effects	Flat to Flat	Flat to Increase	Flat to Flat	Flat to Decrease	Flat to Flat	Flat to Increase
Changes in Stability Trends	Stable to Stable	Stable to Variable	Stable to Stable	Stable to Variable	Stable to Stable	Stable to Variable
Changes in Data Level (%)	58.33-25 = 33.33 (+)	58.33-33.33 = 25 (-)	100-91.67 = 8.33 (+)	100-91.67 = 8.33 (-)	66.67-33.33 = 33.34 (+)	66.67-50 = 16.67 (+)
Data *overlap* (%)	0	0	0	0	0	0

Overall, the study findings showed that subjects G1, G2, and G3 achieved their highest HOTS scores during the intervention phase. For subjects G1 and G3, HOTS scores in baseline 2 (A2) exceeded those in baseline 1 (A1), while G2's scores were lower in baseline 2 than in baseline 1. This indicates that the CMT intervention improved HOTS in G1 and G3, with both achieving over 58% accuracy, while G2 reached 100%

accuracy. The HOTS levels achieved by subjects include analysis, evaluation, and creation, as outlined in Bloom's taxonomy. These skills were bolstered by CMT activities, which followed the stages of relating, investigating, evaluating, communicating, and creating. This finding aligns with research by Bakri and Bakar, indicating that mathematical ability correlates with students' HOTS. High-ability students could interpret meaning, form opinions, and draw conclusions; medium-ability students could interpret meaning and form opinions but struggled with conclusions; while low-ability students were unable to interpret meaning or draw conclusions [16].

The findings of this study align with Singer et al. (2016), who state that providing CMT enhances gifted students' skills and deepens their mathematical understanding [17]. Similarly, the results support Hendriana et al. (2014), who argue that selecting appropriate tasks fosters mathematical understanding, stimulates relationship building, encourages problem formulation and solving, enhances mathematical reasoning, and advances mathematical communication [18].

The research findings indicate that challenging mathematical tasks (CMT) offer an effective approach to enhancing higher-order mathematical thinking in gifted students. Despite efforts to ensure optimal conditions, certain factors proved challenging to control, such as students' physical and psychological states during testing, as well as individual backgrounds, which could impact results. Additionally, instrument consistency across conditions relied solely on validity, material, and indicator alignment. Future research should explore the impact of CMT on other subjects, its effects on different mathematical thinking skills, and its applicability across various educational levels.

4 Conclusions

This study concludes that implementing challenging mathematical tasks in stages—relate, investigate, communicate, evaluate, and create—effectively enhances higher-order thinking skills in gifted students, including their abilities to analyze, evaluate, and create. Initially, gifted students 1 and 3 showed very low proficiency in higher-order thinking, often struggling with complex question comprehension. Following the intervention, their skills improved to a sufficient level, demonstrating increased familiarity with complex tasks, despite minor errors in understanding and calculation. Gifted student 2, initially excelling, maintained very good performance throughout, though their post-intervention skill level slightly declined to a good category due to occasional mistakes. Further studies could explore the long-term impact of challenging tasks on higher-order thinking in diverse student populations and investigate optimal task sequencing and support to sustain skill development across varied proficiency levels.

Acknowledgments.

We would like to thank the support from the Department of Mathematics Education, Faculty of Education, UIN Syarif Hidayatullah Jakarta and the cooperation with gifted students in this single subject research.

Disclosure of Interests.

The authors declare that we have no competing commercial interests or personal relationships that could influence the performance of the research reported in this paper.

References

1. Zaiyar, M., & Rusmar, I.: Students' creative thinking skill in solving higher order thinking skills (HOTS) problems. Al-Jabar: Jurnal Pendidikan Matematika **11**(1), 111–120 (2020)
2. Brookhart, S.M.: How to assess higher-order thinking skills in your classroom. ASCD Alexandria, Virginia (2010)
3. Soemarmo, U.: Berpikir dan disposisi matematik serta pembelajarannya. Jurusan Pendidikan Matematika FPMIPA UPI, Bandung (2015)
4. Ramirez, R.P., Ganaden, M.S. (2008). Creative activities and students' Higher Order Thinking Skills. Journal Education Quarterly **66**(1), 22-23 (2008)
5. Davis, G. A.: Anak berbakat dan pendidikan keberbakatan. PT Indeks, Jakarta (2012).
6. Mhlolo, M. K.: Regular classroom teachers' recognition and support of the creative potential of mildly gifted mathematics learners. ZDM **49**(1), 81–94 (2017)
7. Purnomo, H., Sa'dijah, C., Cahyowati, E. T. D., Nurhakiki, R., Anwar, L., Hidayanto, E., Sisworo.: Gifted students in solving HOTS mathematical problems. AIP Conference Proceedings, **2330**(1), 040008 (2021)
8. Sheffield, L. J.: Dangerous myths about "gifted" mathematics students. ZDM **49** (1), 13-23 (2017)
9. Pikiran Rakyat. 1,3 Juta Anak Cerdas Istimewa, www.pikiran-rakyat.com last accessed 2017/08/1.
10. Setiawan, R., Septiarti, S. W.: A review on Indonesia policy in supporting gifted students' education. International Conference on Special and Inclusive Education (ICSIE 2018), 133-139 (2019)
11. Singer, F. M., Sheffield, L. J., Freiman, V., & Brandl, M.: Research on and activities for mathematically gifted students. Springer Nature (2016).
12. Sunanto, J., Takeuchi, K., & Nakata, H.: Pengantar Penelitian Dengan Subyek Tunggal. CRICED University of Tsukuba, Japan (2005)
13. Gagné, F.: A proposal for subcategories within gifted or talented populations. Gifted Child Quarterly **42**(2), 87-95 (1998)
14. Bosch, N.: Rubric for creative thinking skills evaluation. KCCL, Kansas (2008)
15. Lawshe, C.: A quantitative approach to content validity. Personnel Psychology **28**(1), 563-575 (1975)
16. Bakry., Bakar, M.N.: The Process of thinking among junior high school students in solving HOTS question. International Journal of Evaluation and Research in Education (IJERE) **4**(3), 138-145 (2015)
17. Singer, F. M., Sheffield, L. J., Freiman, V., & Brandl, M.: Research on and activities for mathematically gifted students. Springer Nature, Cham Switzerland (2016)
18. Hendriana, H., Soemarmo, U.: Penilaian Pembelajaran Matematika. PT Refika Aditama, Bandung (2014)