10.48047/jocaaa.2025.34.12.55

ChatGPT as a Cognitive Partner: Fostering Critical Thinking in Mathematical Word Problem Solving

Wafiq Hibi

https://orcid.org/0000-0003-2897-0175
Sakhnin Academic College for Teacher Education
Sakhnin College, 805, Sakhnin, 3081000 Israel
Contact e-mail: wafiq.hibi@gmail.com
Wafiq.hibi@sakhnin.ac.il

Abstract

The current study explored the effect of integrating ChatGPT into mathematics instruction on the development of critical thinking skills among sixth-grade students in the Arab society in Israel. As artificial intelligence tools increasingly influence education, understanding their pedagogical value for young learners has become essential. The research employed a controlled experimental design using a pre-test/post-test format with two groups: an experimental group (28 students) who engaged in guided learning activities incorporating ChatGPT while solving mathematical word problems, and a control group (31 students) who studied the same content using traditional instructional methods. Both groups included male and female students with similar socioeconomic and academic backgrounds.

Critical thinking skills were assessed through two complementary instruments: (1) a culturally adapted Likert-scale questionnaire based on the framework developed by Rodríguez Rojas et al. (2024), and (2) eight open-ended questions designed according to Facione's (2015) model of critical thinking, focusing on analysis, inference, interpretation, and explanation. Quantitative data were analyzed using paired and independent t-tests to determine both within-group and between-group differences. Qualitative data were analyzed through content analysis to identify patterns of reasoning and argumentation in students' written responses.

The quantitative findings demonstrated a statistically significant improvement in the experimental group's mean critical-thinking scores between the pre- and post-tests ($p = 0.012$), whereas no significant change was found in the control group ($p = 0.432$). The comparison of change between groups indicated a moderate advantage for the experimental group, approaching statistical significance ($p \approx 0.07$). Qualitative analyses supported these results: students who learned with ChatGPT produced longer, more coherent, and evidence-based explanations, used precise mathematical and logical terminology, compared alternative solution paths, and reflected on the validity of information provided by the AI. In contrast, the control group tended to provide shorter and more descriptive answers, showing limited depth of reasoning.

10.48047/jocaaa.2025.34.12.55

The study concludes that when used with careful pedagogical guidance, ChatGPT can serve as a powerful cognitive partner that enhances students' engagement and promotes higher-order thinking, even at the elementary school level. However, the results also highlight the need for ongoing teacher mediation to prevent over-reliance on AI and to ensure that learners maintain independent judgment and evaluative skills. Further research with broader samples and across different subjects is recommended to examine the long-term effects and cross-cultural applicability of AI-supported learning environments.

Keywords

ChatGPT; critical thinking; mathematics education; word problem solving; Arab society in Israel

Introduction

In recent years, critical thinking has become one of the most essential competencies for meaningful learning and effective participation in a rapidly changing world. Critical thinking involves the ability to analyze information, evaluate arguments, draw evidence-based conclusions, and consider alternatives in a systematic and reflective manner (Hibi, 2021b; Hibi, 2021c; Hibi, 2022a; Hibi, 2022d; Facione, 2015; Larson et al., 2024). As the availability of information expands exponentially and students encounter multiple sometimes conflicting sources, teachers face the challenge of equipping their learners with the skills required to critically evaluate and responsibly engage with knowledge (Hibi, 2017; Hibi, 2018; Hibi, 2022b; Hibi, 2022c; Rodríguez-Rojas et al., 2024).

The emergence of artificial intelligence (AI), and particularly of ChatGPT as a generative language model, has introduced new opportunities for teaching and learning. AI-based tools allow for interactive dialogue, hypothesis generation, exposure to multiple perspectives, and the exploration of alternative reasoning paths all of which can support the development of critical thinking skills (Guo & Lee, 2023). However, researchers have also warned against overreliance on such tools, which may weaken learners' independent reasoning and critical judgment if not properly guided (Hibi, 2020; Hibi, 2021a; Hibi, 2022e; Hibi, 2022f; Larson et al., 2024; Park, Kim, & Lee, 2021).

Within the Arab society in Israel, the integration of AI into education presents both promise and challenge. Schools in this context operate within unique linguistic, cultural, and social frameworks, where students often navigate between Arabic, Hebrew, and English. These multilingual demands can add cognitive load to learning processes, particularly in subjects such as mathematics that require precise logical reasoning. At the same time, limited access to technological resources in some Arab communities (Agbaria, 2018; Hibi, 2021d; Hibi, 2021e; Hibi, 2021f; Hibi, 2022g; Israel

10.48047/jocaaa.2025.34.12.55

State Comptroller, 2019) underscores the importance of carefully designed digital interventions that promote equity and empowerment. Research indicates that structured use of open-ended questions based on Facione's framework can effectively measure and foster critical thinking, especially in elementary schools (Sarwanto et al., 2021). Incorporating ChatGPT into this process may offer a means to bridge learning gaps and enhance reflective reasoning, but empirical evidence is still scarce particularly among younger learners.

The present study therefore aims to examine the influence of integrating ChatGPT into the learning of mathematical word problem solving on the development of critical thinking skills among sixth-grade students (both boys and girls) in the Arab society in Israel. By comparing an experimental group using ChatGPT-supported instruction with a control group engaged in traditional learning, the study seeks to determine whether AI-assisted learning environments can foster measurable improvements in students' reasoning and evaluative thinking. The findings are expected to contribute both theoretically and practically to the growing body of research on AI in education, with implications for culturally responsive pedagogy and equitable access to innovative learning technologies.

1. Theoretical Background

1. Critical Thinking: Definitions and Conceptual Models

Critical thinking is widely recognized as one of the foundational skills of 21st-century education. While its philosophical roots trace back to classical logic and reasoning, its contemporary educational interpretation emphasizes analytical and reflective processes that enable learners to make rational, evidence-based decisions (Hibi, 2024a; Hibi, 2024b; Hibi, 2024c; Hibi, 2025d; Larson et al., 2024).

Facione (2015) defines critical thinking as a purposeful, self-regulatory process involving interpretation, analysis, evaluation, inference, and explanation accompanied by intellectual virtues such as open-mindedness and curiosity. Similarly, Ennis (1985) views it as "reasonable, reflective thinking focused on deciding what to believe or do," and Halpern (1998) conceptualizes it as a combination of cognitive skills and dispositional traits.

Paul and Elder (2014) highlight the metacognitive dimension, arguing that critical thinkers consciously apply intellectual standards such as clarity, accuracy, relevance, and logical consistency to their reasoning.

Recent research has extended these definitions to the digital age, emphasizing skills related to evaluating online sources, identifying bias, and interacting responsibly with artificial intelligence (Larson et al., 2024).

Rodríguez-Rojas et al. (2024) developed an empirical scale (CTES) measuring both cognitive abilities (e.g., analysis and inference) and thinking dispositions (e.g., intellectual perseverance and flexibility), offering a comprehensive framework for assessing critical thinking development.

10.48047/jocaaa.2025.34.12.55

2. Critical Thinking in Elementary Mathematics Education

In elementary education, critical thinking provides a foundation for reasoning, problem-solving, and independent learning. Research shows that nurturing critical thinking at an early age enhances students' ability to handle complex problems, encourages creativity, and promotes active engagement (Sarwanto et al., 2021). In mathematics specifically, solving word problems requires students not only to apply computational skills but also to interpret situations, identify relevant data, evaluate alternative strategies, and justify their solutions all of which are integral to critical thinking (Hibi, 2024d; Hibi, 2024e; Hibi, 2025c; Hibi, 2025e; Norton & Zhang, 2023).
Integrating reflective dialogue, reasoning tasks, and metacognitive questioning into mathematics lessons helps children develop habits of mind such as curiosity, precision, and persistence. These habits are particularly valuable in multicultural learning environments, where diverse linguistic and cultural backgrounds influence the ways students construct meaning and reason mathematically (Ayalon & Even, 2022; Hibi, 2025a; Hibi, 2025b; Hibi, 2025f; Hibi, 2025g).

In the Arab society in Israel, teaching mathematics involves navigating multiple languages and educational expectations. While students learn in Arabic, they must also understand mathematical terminology in Hebrew and English, which can complicate reasoning and problem interpretation (Hibi, 2025h; Spolsky & Shohamy, 1999). Encouraging critical thinking through mathematical problem solving not only improves academic performance but also empowers students by strengthening their sense of agency, confidence, and participation in broader Israeli society.

3. The Role of Technology and Artificial Intelligence in Promoting Critical Thinking

Digital technologies have transformed learning environments, offering interactive and individualized experiences. When used thoughtfully, such tools can stimulate inquiry, comparison of ideas, and reflection core elements of critical thinking (Zawacki-Richter et al., 2019).
Artificial intelligence (AI) tools, such as ChatGPT, provide immediate feedback, alternative explanations, and multiple perspectives that encourage learners to engage in deeper reasoning. According to Hattie (2009), feedback is among the most influential factors in learning outcomes, and AI-based interaction enables continuous, personalized feedback loops.
However, studies also caution that excessive reliance on AI may reduce cognitive engagement and hinder independent reasoning (Kosmyna et al., 2025). Therefore, teacher mediation is crucial: educators must guide students to verify the accuracy of information, question the validity of responses, and maintain critical awareness while using AI-generated content (Park, Kim, & Lee, 2021).

4. Integrating ChatGPT in Mathematics Learning: Potential and Challenges

10.48047/jocaaa.2025.34.12.55

ChatGPT, as a generative language model, has the potential to act as a "thinking partner" for students. In the context of mathematics education, it can assist learners in rephrasing problems, generating alternative solution paths, and reflecting on logical reasoning. Guo and Lee (2023) demonstrated that structured integration of ChatGPT enhanced students' questioning ability and evidence-based reasoning. Similarly, Hikmah and Walida (2024) found that AI-assisted instruction promoted organization and clarity in academic writing, though deeper analysis still required human guidance. In mathematics classrooms, ChatGPT can be used to encourage Socratic dialogue, help students verify their reasoning, and provide diverse representations of a problem. Yet, challenges remain: AI may produce incorrect or biased information ("hallucinations"), and students might overtrust its answers. Therefore, effective use of ChatGPT demands guided pedagogy teachers must encourage learners to critically evaluate the system's responses, justify their solutions, and compare multiple reasoning strategies. When implemented under these conditions, ChatGPT can strengthen both cognitive and metacognitive aspects of critical thinking, making it a meaningful tool in mathematics education.

3. Methodology

Research Design

The study adopted a quasi-experimental design using a pre-test/post-test format with an experimental and a control group. This design was chosen to examine, in a controlled manner, the causal effect of integrating ChatGPT into mathematics instruction specifically, in the teaching of word problem solving on the development of students' critical thinking skills.
According to Creswell and Creswell (2018), such a design enables both within-group (pre- vs. post-test) and between-group comparisons, allowing researchers to infer the effect of the intervention while controlling for external factors. The study also applied a mixed-methods approach, combining quantitative and qualitative analyses to gain a more comprehensive understanding of the phenomenon. Quantitative data provided measurable indicators of change, while qualitative data offered rich insights into students' reasoning and reflective processes during learning.
This methodological pluralism aligns with recommendations by Johnson and Onwuegbuzie (2004), who emphasize that mixed methods provide a deeper and more nuanced perspective on complex educational phenomena than single-method approaches.

Participants

The study was conducted in a public elementary school serving the Arab society in Israel. Participants were 59 sixth-grade students (ages 11–12), including both boys and girls. The sample was divided into two intact classes:

10.48047/jocaaa.2025.34.12.55

- Experimental group: 28 students who participated in ChatGPT-supported mathematics lessons.
- Control group: 31 students who learned the same mathematical content through traditional instruction.

All participants came from similar socioeconomic backgrounds and had comparable prior academic achievement levels. The inclusion of both genders aimed to reflect a realistic learning environment and to enable the exploration of potential gender-neutral effects of the intervention. Participation was voluntary, and parental consent was obtained from all students prior to data collection.

Research Instruments

Two complementary tools were employed to assess students' critical thinking skills before and after the intervention:

1. Critical Thinking Questionnaire (CTQ):
   A 20-item Likert-scale questionnaire adapted culturally and linguistically to Arabic, based on the Critical Thinking Evaluation Scale (CTES) developed by Rodríguez-Rojas et al. (2024).
   The CTQ assessed two main dimensions:
   - Cognitive skills: analysis, evaluation, and inference.
   - Dispositional traits: intellectual curiosity, open-mindedness, and persistence.
     Items were rated on a four-point scale (1 = strongly disagree to 4 = strongly agree).
     Internal reliability was confirmed using Cronbach's alpha for both pre-test and post-test data.
2. Open-Ended Problem-Solving Tasks:
   A set of eight open-ended mathematics problems was developed to evaluate manifestations of critical thinking during word problem solving. The questions were aligned with Facione's (2015) model, focusing on four cognitive processes analysis, inference, interpretation, and explanation.
   Example items included tasks requiring students to justify their reasoning, compare alternative solutions, and explain mathematical relationships.
   Students' written responses were later analyzed qualitatively to identify indicators of higher-order reasoning and reflective thinking.

Procedure

The intervention lasted **six consecutive weeks** and followed three main stages:

1. Week 1: Pre-Test:
   Both groups completed the Critical Thinking Questionnaire (CTQ) to establish

10.48047/jocaaa.2025.34.12.55

baseline levels. Students were informed that the test was not graded and served solely for research purposes.

2. Weeks 2–5: Instructional Intervention:
   - The experimental group engaged in guided problem-solving sessions using ChatGPT as a learning assistant. Students worked in small groups, posing questions to ChatGPT such as "How can I approach this problem in another way?" or "Can you explain why this method works?"
     After receiving responses, students discussed the accuracy and logic of ChatGPT's explanations, compared them with textbook solutions, and evaluated which reasoning strategy was more valid. The teacher acted as a pedagogical mediator, guiding discussions and encouraging critical verification rather than passive acceptance.
   - The control group studied the same mathematical problems using textbooks and teacher-led explanations, without any use of AI tools. Class discussions focused on standard mathematical reasoning and peer collaboration, maintaining a traditional instructional structure.
3. Week 6: Post-Test:
   Both groups completed the CTQ again under identical conditions. In addition, students' written responses to the open-ended tasks from all sessions were collected for qualitative analysis.

This structure ensured that both groups were exposed to identical content, duration, and teacher involvement, differing only in the integration of ChatGPT as a cognitive tool.

Data Analysis

1. Quantitative Analysis:
   Data were analyzed using SPSS software.
   - Descriptive statistics (means, standard deviations) were computed for pre- and post-test scores.
   - Paired-sample t-tests examined within-group improvements.
   - Independent-sample t-tests compared change scores between groups.
   - Effect sizes (Cohen's d) were calculated to estimate the practical significance of observed differences.
2. Qualitative Analysis:
   Students' written explanations in the open-ended tasks were analyzed using content analysis based on Facione's (2015) categories: analysis, inference, interpretation, and explanation.
   Two independent raters coded all responses, achieving satisfactory inter-rater reliability after discussion and consensus. Representative excerpts illustrating different levels of critical thinking were identified.
   Triangulation between quantitative and qualitative data strengthened the

10.48047/jocaaa.2025.34.12.55

validity of interpretations, ensuring that statistical improvements corresponded with observable evidence of deeper reasoning.

Ethical Considerations

The research complied with institutional ethical standards for educational studies. Approval was obtained from the school administration and local education authorities. Parents received detailed information about the study's goals, procedures, and confidentiality measures. All data were anonymized; no individual student was identifiable in reports.
The intervention was designed to integrate naturally into the regular mathematics curriculum without imposing extra workload. Students were also introduced to the concept of AI literacy, emphasizing that ChatGPT is a supportive tool rather than an unquestionable source of truth encouraging responsible and critical use of technology.

4. Results

This section presents the quantitative and qualitative findings of the study, which examined the effect of integrating ChatGPT into mathematics instruction—specifically in solving word problems on the development of critical thinking skills among sixth-grade students in the Arab society in Israel.

Quantitative Findings

Descriptive Statistics

Table 1 displays the mean scores and standard deviations of students' critical thinking levels before and after the six-week intervention.
The experimental group, which learned with ChatGPT, showed an increase in mean critical thinking scores from 2.91 (SD = 0.38) in the pre-test to 3.10 (SD = 0.36) in the post-test an average gain of 0.19 points.
In contrast, the control group's mean scores rose only slightly, from 2.94 (SD = 0.40) to 2.98 (SD = 0.37) a marginal change of 0.04 points.

These results suggest a clear upward trend in the experimental group's performance compared with the nearly stable scores in the control group. Although the numerical difference appears modest, such changes are consistent with the magnitude of effects typically observed in short-term classroom interventions aimed at developing higher-order thinking skills.

Table 1. Mean and Standard Deviation of Critical Thinking Scores by Group and Test Time

10.48047/jocaaa.2025.34.12.55

Group	Time of Measurement	Mean	SD
Experimental	Pre-test	2.91	0.38
Experimental	Post-test	3.10	0.36
Control	Pre-test	2.94	0.40
Control	Post-test	2.98	0.37

Statistical Significance Tests

Paired-sample *t*-tests were performed to determine whether the observed changes were statistically significant within each group.

- For the experimental group, the improvement was statistically significant ($t(27) = -2.69$, $p = 0.012$).
- For the control group, the change was not significant ($t(30) = -0.80$, $p = 0.432$).

An independent-sample *t*-test comparing the gain scores between the two groups revealed a marginally significant advantage for the experimental group ($t(57) = 1.85$, $p \approx 0.07$).

These findings indicate that the integration of ChatGPT produced a measurable and positive effect on students' critical thinking, though the difference between groups reached only borderline significance. Given the short duration of the intervention, this improvement can be interpreted as a meaningful pedagogical outcome.

Table 2. Statistical Significance of Critical Thinking Score Differences

Group	Mean (Pre)	Mean (Post)	Δ Change	*t*(df)	*p*
Experimental	2.91	3.10	+0.19	-2.69 (27)	0.012*
Control	2.94	2.98	+0.04	-0.80 (30)	0.432
Between-group difference	–	–	–	1.85 (57)	0.07

*p < 0.05

Effect size calculations (Cohen's *d*) suggested a moderate effect for the experimental group ($d \approx 0.45$), consistent with medium-strength educational interventions.

Qualitative Findings

To complement the quantitative results, students' written responses to the eight open-ended word problems were analyzed for evidence of critical thinking processes specifically analysis, inference, interpretation, and explanation (Facione, 2015).

10.48047/jocaaa.2025.34.12.55

Emergent Patterns

Several distinct patterns emerged from the content analysis:

1. Analytical reasoning:
   Students in the experimental group demonstrated greater ability to break down complex word problems into smaller logical steps. Many identified hidden assumptions, clarified relationships between quantities, and selected relevant data before solving the problem.
   Example (translated from Arabic):
   "If the train's speed changes, the time must also change. We need to find the proportional relationship before calculating."
2. Inference and justification:
   Experimental group students were more likely to provide reasons for their chosen strategies, often referencing or questioning ChatGPT's suggestions. They compared alternative methods (e.g., algebraic vs. arithmetic reasoning) and justified their choices.
   In contrast, control group answers were typically procedural, describing solution steps without explanation.
3. Interpretation and reflection:
   ChatGPT-supported learners tended to express metacognitive awareness. Some explicitly mentioned checking whether ChatGPT's responses were logical or accurate, demonstrating the emergence of reflective skepticism an essential aspect of critical thinking.
4. Explanation and communication:
   Students in the experimental group wrote longer, more structured, and linguistically coherent explanations. They used mathematical terminology more precisely and articulated conclusions clearly, while control group responses were shorter and more descriptive.

Summary of Findings

The convergence of quantitative and qualitative data strengthens the study's validity. Both strands indicate that integrating ChatGPT into the learning of mathematical word problems when guided by the teacher enhanced students' analytical reasoning, depth of explanation, and critical evaluation of information.
Although the statistical improvement was moderate, the qualitative evidence reveals richer reasoning patterns and greater engagement among students who used ChatGPT compared with those who did not.

5. Discussion and Conclusions

5.1 Discussion

10.48047/jocaaa.2025.34.12.55

Discussion and Conclusions

The purpose of this study was to investigate whether the integration of ChatGPT into mathematics instruction specifically in solving word problems could foster the development of critical thinking skills among sixth-grade students in the Arab society in Israel. The findings revealed that the experimental group, which engaged in guided interaction with ChatGPT, demonstrated significant improvement in critical thinking scores and exhibited richer, more analytical reasoning compared with the control group. Although the overall between-group difference was only marginally significant statistically ($p \approx 0.07$), the consistency between quantitative and qualitative evidence provides strong pedagogical and theoretical insights.

Interpretation of Findings

The results align with previous studies indicating that AI-supported learning environments can stimulate higher-order cognitive processes when implemented under structured pedagogical guidance (Guo & Lee, 2023; Zawacki-Richter et al., 2019). In this study, students who interacted with ChatGPT not only improved their ability to solve mathematical problems but also developed meta-cognitive skills such as questioning, verifying, and reasoning through alternative approaches. These outcomes suggest that ChatGPT, when used as a thinking partner, can promote analytical engagement and reflective dialogue two core components of critical thinking (Facione, 2015; Halpern, 1998).

The findings also demonstrate that the quality of learning outcomes depends heavily on teacher mediation. In the experimental group, students were explicitly guided to challenge ChatGPT's responses, assess their accuracy, and justify their reasoning. This structured questioning prevented passive consumption of AI-generated answers and instead cultivated skepticism, evaluation, and evidence-based thinking. These results echo Park, Kim, and Lee's (2021) argument that teachers play a vital role as cognitive coaches in AI-mediated learning.

The qualitative analysis further revealed that students developed a greater awareness of the reasoning process itself a hallmark of metacognitive growth. Many students explicitly reflected on *why* a particular strategy worked, rather than simply applying it. Such reflection represents an essential bridge between computational proficiency and genuine mathematical understanding.

Implications for Education in the Arab Society in Israel

This study contributes original empirical evidence from the Arab society in Israel, a context that has been underrepresented in AI and mathematics education research. In many Arab schools, limited access to technological resources and linguistic challenges in multilingual environments can hinder the development of advanced

10.48047/jocaaa.2025.34.12.55

cognitive skills (Agbaria, 2018; Spolsky & Shohamy, 1999). The present findings suggest that, when accessible and properly scaffolded, AI tools such as ChatGPT may help bridge these educational gaps. By enabling students to explore problems interactively and reason in their native language, the tool supports both cognitive and linguistic development.

Furthermore, the inclusion of both boys and girls in the study underscores that the positive effects of AI-assisted learning are not gender-specific. This supports a broader educational goal of promoting equal opportunities in digital and cognitive skill development within the Arab community and across the Israeli education system.

Limitations

Despite its contributions, the study has several limitations that should be acknowledged.

First, the sample size was relatively small and drawn from a single school, limiting the generalizability of the findings. Future research should involve multiple schools and larger samples across different geographic regions and socioeconomic backgrounds.

Second, the duration of the intervention six weeks was relatively short. Longitudinal studies are needed to assess whether the observed improvements in critical thinking persist over time or are influenced by novelty effects.

Third, while the study focused on mathematical word problems, the results may vary across other subjects that demand different types of reasoning, such as science or social studies. Future research should therefore explore cross-disciplinary applications of ChatGPT in elementary education.

Finally, although the study employed a culturally adapted Arabic version of ChatGPT, variations in the model's linguistic accuracy may have influenced the depth of student interaction. Continued advancements in multilingual AI models will likely enhance the quality of such educational applications.

5.2 Conclusions

The findings of this study indicate that the guided use of ChatGPT can serve as a meaningful pedagogical tool for promoting critical thinking in mathematics learning among elementary students in the Arab society in Israel. The combination of real-time feedback, multiple perspectives, and dialogic engagement enables learners to move beyond procedural problem solving toward deeper analytical reasoning.

However, the effectiveness of ChatGPT depends critically on the teacher's role as a facilitator of inquiry and reflection. Without structured guidance, there is a risk that students may rely uncritically on AI generated responses, undermining independent thought.

10.48047/jocaaa.2025.34.12.55

Overall, this research contributes to the growing body of evidence supporting the pedagogical potential of generative AI when integrated ethically and thoughtfully into classroom practice. It highlights the importance of pedagogical design, cultural adaptation, and teacher mediation in transforming AI from a source of information into a catalyst for intellectual growth.

Further large-scale studies across diverse educational settings are recommended to validate and extend these findings, paving the way for equitable and critical use of AI in primary education.

References

1. Abu-Asbah, K. (2005). *Arab education in Israel: Dilemmas of identity and cultural belonging.* Tel Aviv: Ramot Publishing.
2. Agbaria, A. K. (2018). Arab education in Israel: Between equality and politics. *British Journal of Educational Studies, 66*(4), 447–464. https://doi.org/10.1080/00071005.2018.1424315
3. Ayalon, M., & Even, R. (2022). Cultural and linguistic influences on mathematical reasoning in multicultural classrooms. *Journal of Mathematics Education Research, 34*(2), 155–171.
4. Bekerman, Z., & Zembylas, M. (2017). *Teaching coexistence: Education and reconciliation in Israel and beyond.* Cambridge University Press.
5. Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review, 31*(1), 21–32.
6. Bruner, J. S. (1966). *Toward a theory of instruction.* Harvard University Press.
7. Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.
8. Ennis, R. H. (1985). A logical basis for measuring critical thinking skills. *Educational Leadership, 43*(2), 44–48.
9. Facione, P. A. (2015). *Critical thinking: What it is and why it counts.* Insight Assessment.
10. Guo, Z., & Lee, J. (2023). Enhancing student reasoning through generative AI: Structured integration of ChatGPT in higher education. *Computers & Education, 200*, 104787. https://doi.org/10.1016/j.compedu.2023.104787
11. Halpern, D. F. (1998). Teaching critical thinking for transfer across domains. *American Psychologist, 53*(4), 449–455.
12. Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* Routledge.
13. Hibi, W. (2017). Beliefs and attitudes about mathematics among students and their relationship to their achievement in mathematics. Journal of Harbin Institute of Technology, 49(11), 47-57. http://hebgydxxb.periodicales.com/index.php/JHIT/article/view/879
14. Hibi, W. (2018). The effect of origami's educational and geometrical interaction on sixth graders. Journal of Harbin Institute of Technology, 50(11), 33-40. http://hebgydxxb.periodicales.com/index.php/JHIT/article/view/735
15. Hibi, W. (2020). Polynomial coloring applications. *Yantu Gongcheng Xuebao/Chinese Journal of Geotechnical Engineering, 42*(7), 1-4.

10.48047/jocaaa.2025.34.12.55

16. Hibi, W. (2021a). General uses in intermediate value theorems. Multicultural Education, 7(8), 436-439. http://ijdri.com/me/wp-content/uploads/2021/08/48.pdf
17. Hibi, W. (2021b). Isometric between spaces ℓ∞, ℓ1. Advances in Mechanic, 9(3), 432-439. https://advancesinmechanics.com/pdf/137-242.pdf
18. Hibi, W. (2021c). On regular graph, that has two types of faces that having two consecutive numbers of edges. *Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12*(14), 1566-1571. https://turcomat.org/index.php/turkbilmat/article/view/10479
19. Hibi, W. (2021d). The Beckman-Quarles theorem for rational spaces. *Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12*(7), 1907-1912. https://turcomat.org/index.php/turkbilmat/article/view/3120
20. Hibi, W. (2021e). The Beckman-Quarles theorem for rational spaces, Mappings of Q^5 to Q^5 that preserve distance. *Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12*(7), 1913-1917.
21. Hibi, W. (2021f). The Beckman-Quarles theorem for rational spaces: Mapping of Q^d to Q^d that Preserve distance 1. *Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12*(7), 1938-1949.
22. Hibi, W. (2022a). Assembling Planer Graphs to Service the Coloring Number. *Review of International Geographical Education Online, 12*(1), 28-31.
23. Hibi, W. (2022b). Attitudes and perceptions of lecturers at Sakhnin College towards the integration of online teaching and innovative teaching methods and innovative teaching technologies, *Journal of Positive Psychology and Wellbeing*, *6*(2), 2090-2105.
24. Hibi, W. (2022c). Definitions of ratio and proportionality terms and their relationship to cognitive and meta-cognitive orientation in the proportional thinking of preparatory level students. *Multicultural Education, Caddo Gap Press, 3145 Geary Boulevard, PMB 275, San Francisco, CA 94118 U.S.A*, *8*(2),12-34.
25. Hibi, W. (2022d). Excerpts on mathematics education in the Arab community in Israel. Caddo Gap Press, 3145 Geary Boulevard, PMB 275, San Francisco, CA 94118, U.S.A ISBN: 874-745-05-5478-3 (eBook). Library Number: 2458796325.
26. Hibi, W. (2022e). The educational and geometric psychological impact of origami on sixth graders. Journal of Positive Psychology and Wellbeing, 6(1), 2031-2039. https://journalppw.com/index.php/jppw/article/view/2893
27. Hibi, W. (2022f). The Positive Impact of Combining Daily Math with First-Class Equation Learning for Middle School Students, Yantu Gongcheng Xuebao Chinese Journal of Geotechnical Engineering. Vol. 44 No. 6 (2022), 31-44.
28. Hibi, W. (2022g). Relationships between faces in regular, connected, and planar graphs. *Yantu Gongcheng Xuebao/Chinese Journal of Geotechnical Engineering*, *44*(1), 1-5. https://ytgcxb.periodicales.com/index.php/CJGE/article/view/70

10.48047/jocaaa.2025.34.12.55

29. Hibi, W. (2024a). A lesson in philosophy of high school geometry: Strange circle and discrete space. The International Journal of Science, Mathematics and Technology Learning, 31(2), 45-69. https://doi.org/10.18848/2327-7971/CGP/v31i02/45-69
30. Hibi, W. (2024b). Mathematical distance learning in the Israeli Arab middle schools during COVID: Teachers' point of view. International Journal of Learner Diversity and Identities, 31(1). https://scholarsarchives.com/arcijld/wp-content/uploads/2024-31-1-7-1.pdf
31. Hibi, W. (2024c). Using smart applications to develop mathematical concepts among fourth-grade students with arithmetic learning difficulties. The International Journal of Science, Mathematics and Technology Learning, 32(1), 1. https://doi.org/10.18848/2327-7971/CGP/v32i01/1-28
32. Hibi, W. (2024d). Using the (4MAT) model in teaching the triangles similarity unit for the ninth grade. Journal of Positive Psychology and Wellbeing, 8(2), 108-125. https://www.journalppw.com/index.php/jppw/article/view/18237
33. Hibi, W. (2024e). Written Calculation Ability and Numerical Insight in Grades 6 and 7 among the Arab Sector in Israel. Mathematics Teaching-Research Journal. https://mtrj.commons.gc.cuny.edu/volume-16-n-5/
34. Hibi, W. (2025a). Bridging Home and School: Rethinking Everyday Mathematics Through the Lens of Cultural Commodity. International Journal of Multicultural Education, 27(1),33-52. https://doi.org/10.5281/zenodo.15728013
35. Hibi, W. (2025b). Developing Mathematics Education by Integrating Graph Theory into college Curricula. *Multicultural Education*, *11*(1),64-75. DOI: 10.5281/zenodo.14639723
36. Hibi, W. (2025c). How AI Tools Influence the Skills of Young Developers. *International Journal of Multicultural Education, 27*(1),33-52.
37. Hibi, W., & Khalifa, W. (2025d). Developing Numerical Understanding in Early Mathematics Subtraction. International *Journal of Multicultural Education*, *27*(1),14-32. https://doi.org/10.5281/zenodo.15728062
38. Hibi, W. (2025e). From Notes to Numbers: The Impact of Music Education on Mathematical Skills and Cognitive Development. *Acta Scientiae, revista Ensino de Ciênciase Matemática, 26*(1),*181-195.* https://www.periodicos.ulbra.org/index.php/acta/article/view/373
39. Hibi, W. (2025f). Institutional and Ideological Forces in Mathematics Discourse a Critical Perspective. *Acta Scientiae, revista Ensino de Ciências e Matemática, 26*(1),*164-180.* https://www.periodicos.ulbra.org/index.php/acta/article/view/373
40. Hibi, W. (2025g). Proof, Attitudes, and Gender: A Study in Early Math Education. Multicultural Education, 11(4),50-66. DOI: 10.5281/zenodo.15263408
41. Hibi, W. (2025h). Teaching Parallelograms Using Geometer's Sketchpad. International Journal of Stress Management, 32(1), 18-41. https://apa-ijsm.org/volume-32-issue-1/

10.48047/jocaaa.2025.34.12.55

42. Hikmah, N., & Walida, A. (2024). The role of ChatGPT in improving academic writing: Potentials and challenges. *International Journal of Education and Technology, 15*(2), 120–135.
43. Israel State Comptroller. (2019). *Report on digital infrastructure and access to technology in Arab schools.* Jerusalem: State Comptroller's Office.
44. Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14–26. https://doi.org/10.3102/0013189X033007014
45. Kosmyna, N., Singh, R., & Cummings, M. (2025). Cognitive engagement and overreliance in AI-assisted learning: Insights from MIT classroom studies. *AI in Education Journal, 45*(1), 22–39.
46. Larson, E., Peters, K., & Silva, R. (2024). Rethinking critical thinking in the age of AI: A conceptual framework. *Educational Review, 76*(3), 317–336. https://doi.org/10.1080/00131911.2023.2287415
47. Norton, S., & Zhang, Q. (2023). Word problems and reasoning: Linking problem structure with student cognition. *International Journal of Mathematical Education in Science and Technology, 54*(8), 1224–1241.
48. Park, J., Kim, S., & Lee, H. (2021). Teacher mediation in AI-based learning environments: Balancing guidance and autonomy. *Computers & Education, 174*, 104296. https://doi.org/10.1016/j.compedu.2021.104296
49. Paul, R., & Elder, L. (2014). *Critical thinking: Tools for taking charge of your professional and personal life* (2nd ed.). Pearson Education.
50. Piaget, J. (1972). *The psychology of the child.* Basic Books.
51. Rodríguez-Rojas, M. L., Hernández, C., & Vargas, D. (2024). Developing and validating a critical thinking evaluation scale (CTES) for multicultural contexts. *Journal of Educational Psychology, 116*(5), 890–907. https://doi.org/10.1037/edu0000865
52. Sarwanto, S., Mulyono, A., & Rahardjo, S. (2021). Using Facione's framework to assess critical thinking in elementary students. *International Journal of Instruction, 14*(4), 67–84.
53. Spolsky, B., & Shohamy, E. (1999). *The languages of Israel: Policy, ideology, and practice.* Multilingual Matters.
54. Suriano, D., Plebe, A., Acciai, G., & Fabio, R. A. (2024). Interactive AI and student curiosity: The role of feedback in developing analytical thinking. *Journal of Educational Technology Research, 72*(2), 98–112.
55. Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.
56. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education, 16*(1), 39. https://doi.org/10.1186/s41239-019-0171-0