JOURNAL OF CULTURAL ANALYSIS AND SOCIAL CHANGE

Journal of Cultural Analysis and Social Change,
2025, 10(2), 3361-3376
ISSN: 2589-1316

Evaluation of the Performance of Saudi Students in the International Mathematical Olympiad (IMO) Between 2019–2023: A Benchmarking Comparison with their Peers in the G20 Countries

Ahmed Mohammed Nasser Alsager[1*]

[1] *Associate Professor – Deanship of the Common First Year – King Saud University Kingdom of Saudi Arabia. Address: Al-Malqa District, Riyadh, Kingdom of Saudi Arabia. ORCID: https://orcid.org/0009-0005-5221-5032*

*Corresponding Author: aalsaqer1@ksu.edu.sa

Citation: Alsager, A. M. N. (2025). Evaluation of the Performance of Saudi Students in the International Mathematical Olympiad (IMO) Between 2019–2023: A Benchmarking Comparison with their Peers in the G20 Countries, *Journal of Cultural Analysis and Social Change*, 10(2), 3361-3376. https://doi.org/10.64753/jcasc.v10i2.2105

Published: November 19, 2025

ABSTRACT

This study aimed to evaluate the performance of Saudi students in the International Mathematical Olympiad (IMO) from 2019 to 2023, through benchmarking their results against peers from G20 countries. The objective was to identify the extent of competitiveness of Saudi participants at the international level and to highlight performance gaps that require intervention. The descriptive-analytical method was employed as the research approach, supported by statistical techniques including arithmetic mean, standard deviation, Z-score, and T-score. The study sample consisted of 30 Saudi students (28 males and 2 females) who officially participated in the IMO under the supervision of the King Abdulaziz and His Companions Foundation for Giftedness and Creativity (Mawhiba). Six students participated annually, carefully selected after several qualifying stages and examinations. Raw scores were standardized to enable valid comparisons and classification of performance into high, medium, or low levels. The findings indicated that Saudi students generally performed at an "average" level, with the exception of 2020, which recorded a "low" level. Although some improvement appeared in 2022, negative Z-scores confirmed that overall performance remained below the global average, reflecting a persistent gap with G20 countries. Performance varied significantly by problem type: students achieved stronger results on Problem 1 (P1), while struggling with more complex problems such as P3 and P4. Differences also emerged across mathematical domains, with "average" outcomes in number theory, polygons, probability, and circles, but "low" outcomes in sequences, truth tables, and quadrilaterals. In light of these results, the study recommends more specialized training programs, instructional strategies aligned with IMO complexity, and enriched curricula with challenging problems across diverse mathematical domains to bridge performance gaps and strengthen competitiveness.

Keywords: International Mathematical Olympiad (IMO), G20 countries, Benchmarking Comparison.

INTRODUCTION

Mathematics is considered a fundamental language of science, characterized by a logical structure and specific rules that help improve designs and minimize errors. It also contributes to conducting theoretical calculations with precision, which saves time and resources and aids in problem analysis and forecasting future behavior. Moreover, mathematics plays a crucial role in individual development, as it fosters logical and scientific thinking that can be applied in daily and professional life. By enhancing critical and creative thinking, and developing both inductive and deductive reasoning skills, mathematics equips individuals with the ability to approach problems logically and analyze matters deeply, understanding the causes behind various phenomena, confirming that nothing happens without a reason (Camarena Gallardo, 2014).

When discussing mathematics education in schools, we typically refer to regular math classes and the standard curriculum. However, mathematics competitions and Olympiads are also considered an essential component of general education. For gifted students, they offer a means to challenge themselves, express their abilities, and compare their performance with peers (Avotina & Šuste, 2016). Campbell and Walberg argue that competitions operate based on a set of assumptions that underpin a logic for talent discovery. These assumptions may be summarized as follows: competitions are necessary because many schools lack the curricula and resources needed to challenge exceptional students; competitions attract participants with outstanding abilities; competitions stimulate early talent development; and once these talents are cultivated, they are expected to contribute to society (Campbell & Walberg, 2010).

As research on science competitions has shown, participation in such contests has a positive impact on students' academic and professional trajectories. Participants in these competitions are more likely to pursue careers in science, and successful competitors, especially those who perform well at the national level, often achieve outstanding results in their studies and professional lives. Furthermore, these students frequently exhibit strong motivation and self-discipline and often come from supportive family environments that nurture their interest in science. Many successful participants have also shown that regular classroom instruction was unchallenging for them, highlighting their need for more rigorous academic challenges (Tschisgale et al., 2024).

Mathematics competitions are among the most important tools for identifying and developing exceptional mathematical talent. Participation in such contests helps students develop critical and creative thinking skills, as they are challenged to solve complex problems that require high levels of innovation and intelligence. Additionally, these competitions help foster interest in science and technology, ultimately contributing to the development of a strong foundation of future scientists and researchers (Navarro Cendejas, 2017).

International benchmarking comparisons, such as those provided by TIMSS and PISA studies, are considered essential tools for assessing students' academic performance on a global scale. Through these comparisons, it becomes possible to analyze the gaps in educational attainment across countries and to identify strengths and weaknesses within educational systems (Shimizu & Vithal, 2023).

Based on this analytical framework, the same methodology can be applied to evaluate the performance of Saudi students in the International Mathematical Olympiad (IMO) between 2019 and 2023, and to compare their performance with that of their peers in the G20 countries.

This approach will contribute to understanding the differences in performance levels and to enhancing knowledge of the effectiveness of the educational strategies adopted in Saudi Arabia in comparison to international performance standards. It will also help identify the areas that require development and improvement in order to achieve better results in global competitions.

The International Mathematical Olympiad (IMO) is one of the most prestigious global mathematics competitions. It was established in 1959 and brings together students from numerous countries around the world. The Olympiad is held at an international level, with students selected to participate following intensive training provided by national teams in each participating country. The primary goal of the IMO is to inspire students to explore mathematics and challenge their understanding within a global framework. It also aims to foster friendships through the interaction of students from various nations, thereby promoting international cooperation and creating opportunities for the exchange of ideas and the global dissemination of mathematical knowledge (Santiago & Alves, 2022).

Each year, participating countries send six of their top high school mathematics students to compete in the IMO. The selection process varies by country but typically involves multiple rounds of increasingly difficult competitions. Many countries also organize intensive training camps for students who achieve the highest scores, in preparation for the international event. Currently, more than 100 countries participate in the IMO, and over the past sixty years, more than 18,000 contestants have taken part in this competition (Berg, 2021).

Problem Statement:

A study conducted by Canellidis & Sofianopoulou (Canellidis & Sofianopoulou, 2022), noted that Saudi Arabia demonstrated a below-average efficiency level when analyzing the performance of countries participating in the International Mathematical Olympiad (IMO). The efficiency of each country was evaluated based on outputs such as the number of medals earned, while taking into account input resources like population size and GDP per capita.

Based on the results of the Data Envelopment Analysis (DEA), the researcher extracted data for the G20 countries as presented in the previous study and conducted a comparative analysis with Saudi Arabia. The findings revealed significant variation in efficiency levels: China, India, South Korea, and Russia achieved an optimal efficiency score of 1.0, indicating full and effective utilization of available resources. The United States recorded a very high efficiency score of 0.88, followed by Italy with a relatively high score of 0.80. Indonesia achieved an efficiency score of 0.72, while the United Kingdom scored 0.70. Germany registered a moderate efficiency level at 0.68, whereas France showed a relatively lower efficiency of 0.58. Brazil recorded a score of 0.53, followed by

Saudi Arabia with a relatively low efficiency score of 0.44. This indicates that Saudi Arabia has not yet fully utilized its available resources, presenting substantial opportunities for improving its performance in international competitions.

In light of the above, Saudi Arabia has faced a challenge in achieving outstanding performance in the IMO during the period 2019–2023, particularly when compared to the G20 countries, as the results reflect noticeable differences in performance levels. This research aims to objectively evaluate the performance of Saudi students through a benchmarking comparison with their peers in the G20 countries, in order to determine their competitive position and efficiency level in this international competition.

The problem of the study can thus be summarized by the following main research question:

How can the performance of Saudi students in IMO be evaluated based on a benchmarking comparison with their peers in the G20 countries during the period 2019–2023?

Sub-Questions

What is the performance level of Saudi students in solving IMO problems based on a benchmarking comparison with the performance of students from the G20 countries between 2019 and 2023?
What is the performance level of Saudi students on each individual IMO problem, according to a benchmarking comparison with students from the G20 countries between 2019 and 2023?
What is the performance level of Saudi students in IMO exams, based on a benchmarking comparison categorized by problem topics between 2019 and 2023?

Study Objectives

The current study aims to:

1. Evaluate the performance level of Saudi students in solving IMO problems between 2019 and 2023, through a benchmarking comparison with students from G20 countries.
2. Analyze the performance of Saudi students on each IMO problem between 2019 and 2023, in comparison with the performance of G20 students.
3. Determine the performance level of Saudi students in IMO tests, categorized by mathematical topics, based on benchmarking data from 2019 to 2023.

Significance of the Study

1. To provide a scientific benchmarking diagnosis that reveals the standing of Saudi students in the IMO compared to their peers in the G20 countries, information that can be utilized by policymakers and educational supervisors for developing educational strategies.
2. To identify strengths and weaknesses in the performance of Saudi students, thereby helping direct efforts to improve their mathematical skills.
3. To contribute valuable input for designing specialized training programs aimed at enhancing student performance in topics that require development.
4. To support Saudi Arabia's national vision by raising the competitive efficiency of its students in international scientific competitions.

Study Delimitations

Topical Delimitation:

This study focuses specifically on the topics of the IMO test questions from the years 2019 to 2023, which is the period during which this study was conducted.

Temporal Delimitation:

The study is confined to the timeframe from 2019 to 2023, during which the performance of Saudi students in the IMO is analyzed and compared with the performance of students from the G20 countries within the defined period.

Human Delimitation:

The study is limited to Saudi students who officially participated in the IMO between 2019 and 2023, under the supervision of the King Abdulaziz and His Companions Foundation for Giftedness and Creativity (Mawhiba). A total of 30 students participated, including 28 males and 2 females, with an average of six students representing the Kingdom each year. These students were meticulously selected after passing through multiple qualifying stages and tests.

Terminology

***International Mathematical Olympiad (IMO)*:**

The IMO is an annual international mathematics competition held in a different country each year. Each participating country sends a team of six pre-university students, accompanied by a team leader, deputy leader, and observers. The competition spans two days, during which students are required to solve three problems per day over a period of 4.5 hours. The Olympiad exam consists of six questions, each worth seven points, for a maximum total score of 42 points. The problems typically cover topics in algebra, geometry, number theory, and combinatorics (Djukić et al., 2006).

Operational Definitions

The IMO is defined operationally in this study as an international competition held between 2019 and 2023, in which 112 countries participated, each represented by a delegation of six students, including the G20 countries.

Group of Twenty (G20)

The G20 is an international economic forum composed of 19 countries plus the European Union. The Member States represent the world's major economies and include: Argentina, Australia, Brazil, Canada, China, France, Germany, India, Indonesia, Italy, Japan, Mexico, Russia, Saudi Arabia, South Africa, South Korea, Turkey, the United Kingdom, the United States, and the European Union ("Energy poverty affects women the most: At a G20 event, Brasil launches clean cooking initiatives," 2024).

***Benchmarking*:**

The Cambridge Dictionary defines benchmarking as: "the process of measuring the quality of something by comparing it with something else that is accepted as a standard" ("Benchmarking,").
Benchmarking is a method of performance comparison that often aims to identify good—or more ambitiously, best—practice methods. As such, it is also used to diagnose performance issues and identify areas of strength (*Benchmarking in higher education*, 1998).

***Operational Definition of Benchmarking*:**

Benchmarking is operationally defined as the process of analyzing and comparing the performance of Saudi students in the IMO between 2019 and 2023 with that of their G20 counterparts. This is conducted by calculating z-scores and T-scores for each group, and comparing means and standard deviations of performance, in order to assess the extent to which Saudi students meet global performance standards, and to identify performance gaps and potential areas for improvement.

Theoretical Framework

In its early days, the IMO was a much smaller competition than it is today. In 1959, seven countries gathered to compete in the first IMO: Bulgaria, Czechoslovakia, the German Democratic Republic, Hungary, Poland, Romania, and the Soviet Union. Since then, the competition has been held annually. Gradually, additional countries from the Eastern Bloc joined, followed by Western European nations, and eventually the Olympiad expanded to include many countries from all over the world and from every continent.
Each participating country sends a team of six contestants, and each contestant competes individually (without any help or collaboration). Countries also send a team leader, who participates in the selection of problems and remains isolated from the team until the end of the contest, as well as a deputy leader, who accompanies and supervises the contestants.

The IMO takes place over two days, with students given 4.5 hours each day to solve three problems, making a total of six problems. The problems cover areas such as algebra, geometry, number theory, and combinatorics, among others. Typically, the first problem of the day is the easiest, while the last is the most difficult—though notable exceptions exist. Each problem is worth 7 points, for a maximum total score of 42.
The scoring is determined by a jury, based on negotiations between problem coordinators (appointed by the host country) and the team leader and deputy, who advocate for their contestants. This system ensures a relatively objective assessment, with score differences rarely exceeding two or three points (Djukić et al., 2006; Jung & Lee, 2021).

Saudi Arabia's Participation in the IMO

Saudi Arabia first joined the IMO as a scientific observer in Japan 2003, without any student participation, in accordance with IMO regulations. The official student participation began in subsequent years:

2004 – Greece, 2005 – Mexico, 2006 – Slovenia, 2007 – Vietnam, 2008 – Madrid, Spain

In 2009, Saudi Arabia participated in Germany as an observer only, without student contestants.("Olympiad in mathematics (IMO).", 2024)

Factors Contributing to the Difficulty of Mathematical Problems

Several factors contribute to increasing the difficulty of mathematical problems. Based on a review of previous studies, the following factors can be identified:

Task Complexity:

The complexity of mathematical tasks is one of the most important factors influencing problem difficulty. Problems that require multiple steps to solve or involve multiple mathematical elements increase the cognitive load on the student. Lannie and Martens (Lannie & Martens, 2004) noted that "as task difficulty increases, students tend to prefer time-based reinforcement because they feel challenged to achieve the required accuracy in more complex tasks" (Lannie & Martens, 2004).

Interaction Between Difficulty and Motivation:

Motivation plays a key role in students' willingness to solve math problems. According to Lynch and Patten (Lynch et al., 2013), "students assigned easier tasks show significant improvement in performance due to increased confidence and motivation, which enhances their engagement" (Lynch et al., 2013). Conversely, students facing more difficult tasks may experience decreased motivation, negatively affecting their performance.

Math Anxiety:

Anxiety toward mathematics can increase the perceived difficulty of problems. "Students with high math anxiety perceive math tasks as more difficult compared to those with low anxiety" (Lynch et al., 2013), which in turn affects their problem-solving strategies.

Working Memory:

Working memory is crucial for solving complex mathematical problems, as students need to hold and process information simultaneously. Research indicates that "high intrinsic cognitive load directly impacts students' working memory, thereby increasing the difficulty of solving problems" (Lannie & Martens, 2004).

Immediate Feedback:

Receiving immediate feedback helps students quickly adjust their problem-solving strategies, enhancing performance on more complex tasks. Lynch and Patten (Lynch et al., 2013) stated that "immediate feedback enables students to identify errors more rapidly, reducing frustration and increasing motivation to persist" [p. 12].

Individual Differences:

Students' abilities to solve mathematical problems vary based on their academic backgrounds and mathematical intelligence. High-ability students can handle complex problems more effectively, while others may struggle even with simpler tasks. Lannie and Martens (Lannie & Martens, 2004) note that "task difficulty may vary among students depending on their level of mathematical skill and ability to manage problem complexity" [p. 55].

Problem-Solving Strategies:

The strategies students use to solve problems differ between high performers and those who face difficulties. High-performing students tend to employ diverse and innovative strategies, whereas lower-performing students often rely on simpler, more traditional strategies that may not be effective for complex problems. According to Lynch and Patten (Lynch et al., 2013), "students using complex and innovative strategies are better equipped to handle difficult problems compared to those who depend on basic, traditional methods" [p. 9].

Cognitive Load Theory

One of the most prominent theories explaining the difficulty of mathematical problems is the Cognitive Load Theory which posits that working memory is affected by three types of cognitive load: intrinsic load, extraneous load, and germane load. According to the theory, the difficulty of problems depends on the complexity of information that must be processed in working memory. "Intrinsic cognitive load relates to the complexity inherent in the task itself, while extraneous load concerns how the information is presented and solved" (Lannie & Martens, 2004).

Cognitive Load Theory also highlights the burden that mathematical problems place on working memory, especially in complex tasks. Sweller explains that traditional teaching methods often increase cognitive load, hindering the transfer of new information into long-term memory. The theory suggests linking prior knowledge with new concepts and reducing unnecessary cognitive demands to improve learning efficiency and problem-solving performance (Moussa, 2021). In contexts such as the IMO, solving problems with high cognitive load requires substantial capacity to manage and analyze information, thereby increasing the difficulty and making the problems a significant challenge for students.

Summary

A combination of interrelated factors plays a major role in determining the difficulty level of mathematical problems faced by students. From task complexity to motivation, math anxiety, individual differences, and working memory, each factor significantly affects students' ability to successfully solve mathematical problems.

Cognitive Load Theory provides a useful framework for understanding how these factors influence problem difficulty, especially in environments demanding high levels of analytical thinking like the IMO. Based on these insights, curricula and teaching strategies can be improved to provide better support to students and enhance their performance in solving complex mathematical problems.

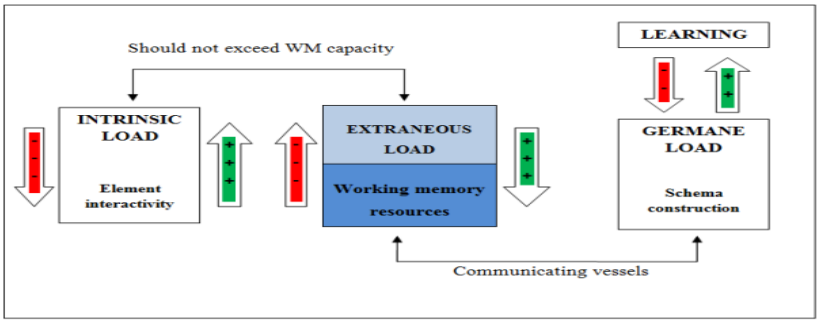


Figure 1: The Relationship Between Cognitive Load and Learning

The green arrows (with positive signs) represent cognitive processes that support learning and the development of problem-solving skills, while the red arrows (with negative signs) represent cognitive processes that hinder learning and the development of problem-solving skills. (Dhlamini & Mogari, 2011)

STUDY METHODOLOGY AND PROCEDURES

Methodology

This study adopts a comparative benchmarking approach, along with the descriptive analytical method. It analyzes the quantitative performance of Saudi students in the IMO from 2019 to 2023 and compares it to the performance of their counterparts in the G20 countries. This design relies on the analysis of pre-existing data from the IMO website, with the aim of measuring gaps and determining the level of efficiency and effectiveness in comparison with other countries.

Study Procedures

1. Collecting raw data for each year from the official IMO website for the period 2019–2023. This includes performance data for each country, such as student rankings, total scores, and detailed results for each test problem.
2. Filtering the results to extract the performance data of the G20 countries and determining each country's global ranking.
3. Organizing the data to show the rankings of students from each country per year, while taking into account the overall performance ranking of each country.
4. Calculating the average performance using the Z-score for each country during the specified period to measure each country's relative performance compared to the global average.
5. Presenting IMO test problems to three secondary school mathematics teaching experts and asking them to classify the questions according to the subject areas covered in the secondary school mathematics curriculum.

6. Comparing the results between Saudi Arabia and the G20 countries to answer the study questions and determine each country's relative annual performance.

Study Sample

The study sample consists of data related to the performance of students from the G20 countries in the IMO during the period from 2019 to 2023. This includes country rankings, total points scored, and the overall performance level of each country. The purpose of analyzing this data is to compare the quantitative performance of Saudi students with that of their peers in the G20 countries based on the available results for the specified years.

Statistical Measures Used

Using the Statistical Package for the Social Sciences (SPSS-V.28), the following statistical measures were employed:

- ***Frequencies and Percentages:*** Used to describe the data and determine rankings.
- ***Arithmetic Mean***: The sum of values divided by the number of values, used to confirm the ranking of values.
- ***Ranks:*** Utilized to classify values.
- ***Standard Deviation:*** Calculated to determine the degree of deviation, variation, or dispersion of scores from the general mean.
- ***Z-Score:*** Used to convert raw scores into standardized scores based on the mean and standard deviation of the distribution, allowing the value of each score to be better interpreted.
- ***T-Score***: Converts the Z-score (which may include decimals) into standardized scores with a fixed standard deviation and mean, offering more precise and adjusted standardized values that account for averages and value limits.
- ***Minimum Performance Threshold Formula Applied to T-Scores:*** Using the formula (Mean ± Standard Deviation), where the mean is 50 and the standard deviation is 10. This formula resulted in the following classification:

Low	Medium	High
Below 40	40–60	Above 60

The use of both Z-score and the adjusted T-score in this study aligns with the benchmarking methodology, as each metric offers distinct insights for performance evaluation. The Z-score is used to determine the position of individual values (e.g., student scores) within a given distribution based on the mean and standard deviation. In contrast, the T-score allows for precise comparisons between group performance means while adjusting for potential differences in variance or sample sizes.

The integration of both metrics enhances the accuracy of benchmarking in this research. As such, the study provides a standardized evaluation of Saudi students' performance in the IMO compared to their peers in the G20 countries, contributing to more accurate and objective results.

STUDY FINDINGS

The findings of the study are presented through answers to its sub-questions as follows:

First Question:

What is the performance level of Saudi students in solving problems from the IMO based on a benchmarking comparison with students from the G20 countries between 2019 and 2023?

To answer this question, the researcher extracted the performance data of G20 countries from the official source and conducted an analytical comparison with the performance of students from Saudi Arabia.

Table (1) presents the relevant results and indicators, categorized by year, based on the total scores achieved by the Saudi team (consisting of 6 students) out of the total possible score of 252 points.

Benchmarking analysis was applied using Z-scores and T-scores to estimate Saudi Arabia's relative performance within the context of G20 countries. This approach provides an objective foundation for assessing performance differences and identifying trends over time.

Table 1. Annual Performance of Saudi Students in the IMO according to Benchmarking Indicators (Z-Score and Adjusted T-Score), 2019–2023

Performance Level	Adjusted T-Score	Z score	Total Score (Out of 252)	Year
medium	42.6	-0.7	124	2019
low	32.0	-1.8	82	2020
medium	42.0	-0.8	90	2021
medium	48.1	-0.2	168	2022
medium	41.0	-0.9	130	2023

Table Interpretation

The table illustrates the indicators and results of Saudi students' performance in the IMO from 2019 to 2023. It includes the total score, Z-score, adjusted T-score, and the performance evaluation level. Based on a benchmarking analysis, the following key points can be concluded:

- ***Relative Improvement over the Years:*** Despite fluctuations in total scores across the years, the overall performance level of Saudi students remained within the "medium" range for most years, except in 2020, when the performance level dropped to "low".
- 2020 ***Marked the Lowest Performance Level:*** In 2020, Saudi students achieved a total score of 82, the lowest among the five years. The Z-score was -1.8, indicating that the performance was significantly below the average of other participating countries. The adjusted T-score was 32.0, the lowest among the five years, reflecting a notable decline in performance.
- ***Noticeable Improvement in 2022:*** In 2022, the students recorded the highest total score (168 points), and the Z-score reached -0.2, indicating a performance closer to the overall average compared to previous years. The adjusted T-score was 48.1—the highest score across the five-year span—signifying a clear improvement.
- ***Consistency at a Medium Level in Most Years*****:** Except for 2020, students consistently achieved a "medium" performance level from 2019 to 2023, suggesting a relatively stable performance with room for further improvement toward higher achievement levels.
- ***Interpretation of the Z-score:*** The negative Z-scores in all years indicate that performance remained below the global average of participating countries. However, a smaller negative value (e.g., -0.2 in 2022) reflects closer alignment with the overall average performance.

These findings reveal the importance of developing preparation programs both qualitatively and quantitatively. Remaining at an "average" level without progressing to "above average" or "high" indicates a structural gap related to the quality of training and the methods used for selecting and qualifying students, as well as a need to broaden the participation base in order to increase opportunities for identifying outstanding mathematical talent. Accordingly, these findings do not merely reflect the current state of performance but also open the way for formulating more effective educational and training policies that could enhance international competitiveness and elevate the Kingdom's standing in global scientific forums.

Research Question Two

What is the performance level of Saudi students in solving each question of the IMO from 2019 to 2023, based on a benchmarking comparison with students from the G20 countries?

This section addresses the second research question, which aims to analyze the performance level of Saudi students on each individual question of the IMO during the period from 2019 to 2023. The analysis is conducted through a benchmarking comparison with the performance of students from the G20 countries.

In this analysis, the statistical indicators Z-Score and T-Score were used to estimate the relative position of Saudi students' performance within the normal distribution of their G20 peers' performance on each question. These indicators reflect how close or distant the Saudi students' scores are from the G20 average, and the degree of deviation from that average.

This analysis serves as a foundation for the next section (Research Question Three), which focuses on classifying the IMO questions according to the mathematical topics covered in the Saudi national mathematics curriculum. This classification will help establish a connection between the nature of the local academic content and the domains of international assessment.

Table (2): Detailed Performance of Saudi Students on IMO Questions for the Year 2019 according to Benchmarking Indicators (Z-Score and Adjusted T-Score) Compared to G20 Students

Performance Level	Rank	Adjusted T-Score	Z score	Raw Score (Saudi Students)	Question
medium	1	53.00	0.30	41	P1
medium	2	49.90	-0.01	35	P5
medium	3	43.80	-0.62	0	P6
medium	4	43.60	-0.64	19	P2
medium	5	41.70	-0.83	0	P3
medium	6	40.90	-0.91	29	P4

The performance of Saudi students in the 2019 IMO showed variation across the six questions. They achieved the highest performance on question P1, ranking first, followed by question P5 (Probability), which ranked second. This reflects a performance level comparable to that of students from G20 countries. In contrast, their performance was weaker on questions P6, P2, P3, and P4, indicating that their scores on these questions were below the average of G20 students — particularly on P3 and P4, which points to challenges associated with problems of a structural nature or those requiring advanced skills in abstract thinking and in-depth analysis. Thus, it may be said that the 2019 findings reveal strengths concentrated in direct and basic questions, alongside a need for further development in areas related to complex problems that demand higher levels of mathematical reasoning.

This suggests that the question where Saudi students performed most similarly to their international peers was P1, while the weakest performance was observed in question P4.

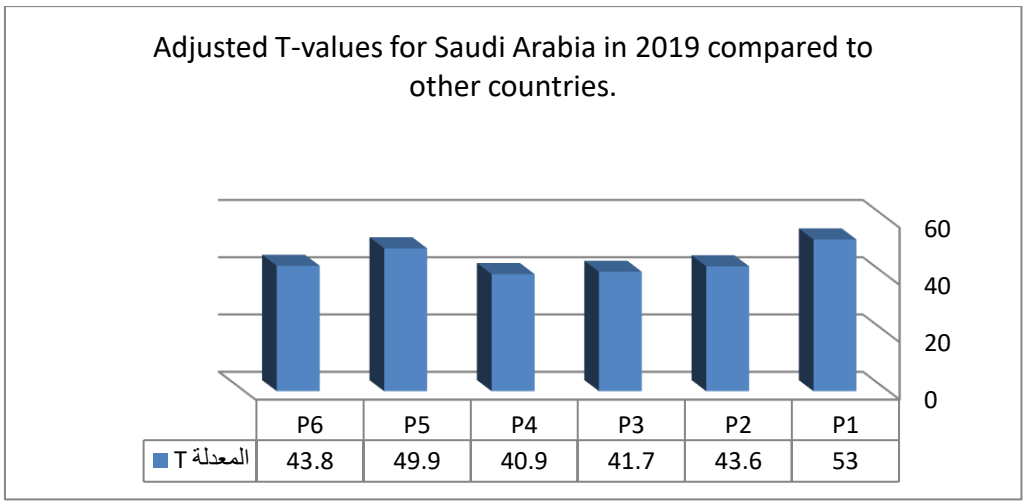


Table (3): Detailed Performance of Saudi Arabian Students on the IMO 2020 Problems according to Benchmarking Indicators (Z-Score and T-Score) Compared to G20 Countries' Students

Performance Level	Rank	Adjusted T-Score	Z score	Raw Score (Saudi Students)	Question
medium	1	45.40	-0.46	8	P3
medium	2	42.80	-0.72	1	P6
low	3	33.30	-1.67	6	P2
low	4	30.00	-2.00	24	P4
low	5	29.10	-2.09	8	P5
low	6	26.40	-2.36	35	P1

The results of Saudi students in the 2020 IMO showed a decline in performance across all six questions compared to the G20 countries. Their best performance was on question P3 (T = 45.40), ranking first, followed by question P6 (T = 42.80) in second place. The remaining questions saw weaker performance, with question P2 (T = 33.30) ranking third, question P4 (T = 30.00) fourth, question P5 (T = 29.10) fifth, and the poorest performance on question P1 (T = 26.40) in sixth place. This indicates a clear gap between the performance of Saudi students and their G20 peers on these questions.

This means that the question on which Saudi students performed closest to other countries was question P3, while the question with the lowest performance was question P1.

The findings also revealed a gap between Saudi students and their peers from other G20 countries, reflecting weaker performance on basic and introductory problems compared with relative strength in some of the more complex ones. This points to an imbalance in the acquisition of skills, as students display readiness in certain areas of advanced mathematical thinking while showing weakness in the fundamentals, which constitute a key entry point for achieving integrated performance in international competitions.

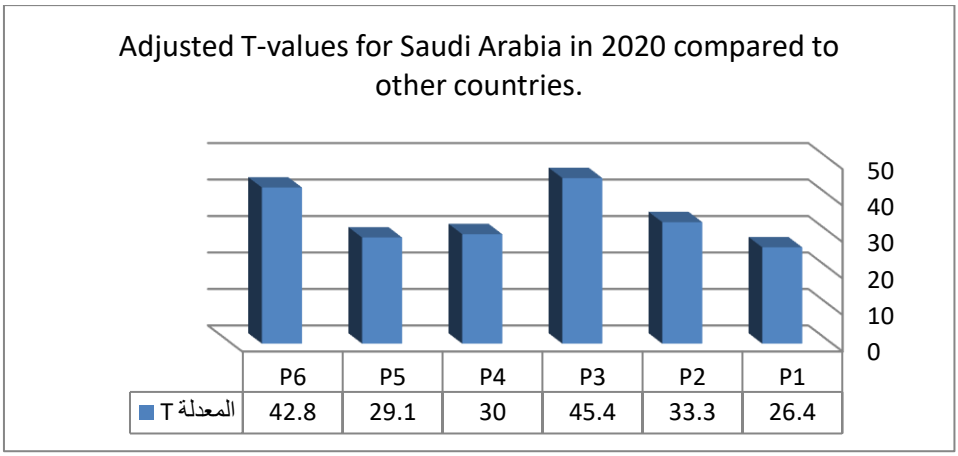


Table (4): Detailed Performance of Saudi Students in the IMO 2021 Questions according to Benchmarking Indicators (Z-Score) and Adjusted Score (T-SCORE) Compared to Students from the G20 Countries

Performance Level	Rank	Adjusted T-Score	Z score	Raw Score (Saudi Students)	Question
medium	1	57.40	0.74	42	P4
medium	2	45.00	-0.50	2	P3
medium	3	42.00	-0.80	0	P2
medium	4	41.90	-0.81	17	P5
medium	5	41.30	-0.87	0	P6
low	6	33.60	-1.64	29	P1

The results of Saudi students in the 2021 IMO showed a notable variation in performance across the six questions compared to G20 countries. They achieved their best and relatively comparable performance with students from G20 countries in question P4 (T = 57.40), ranking first, followed by question P3 (T = 45.00) in second place. Question P2 (T = 42.00) ranked third, followed by question P5 (T = 41.90) in fourth place, and question P6 (T = 41.30) in fifth place. The weakest performance was in question P1 (T = 33.60), ranking sixth, reflecting a disparity in performance levels across different questions.

This means that the question where the students performed most suitably and closest to other countries was question P4, while the lowest performance was on question P1.

This indicates the persistence of challenges in dealing with introductory or basic problems. Such a disparity suggests that performance in 2021 was marked by clear strengths in certain areas, alongside a need to strengthen competencies in basic questions so as to achieve a greater balance in overall performance.

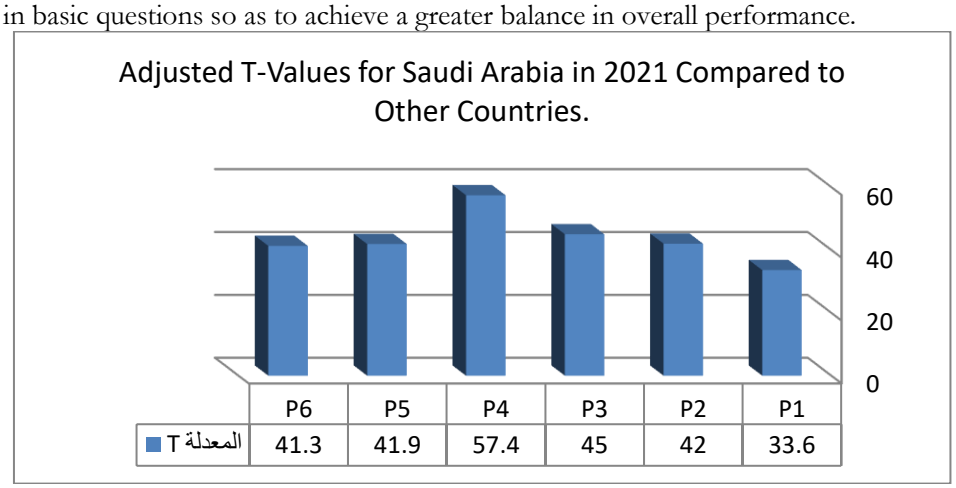


Table (5): Detailed Performance of Saudi Students in the IMO 2022 Questions According to Benchmarking Indicators (Z-Score) and Adjusted T-Score Compared to G20 Countries

Performance Level	Rank	Adjusted T-Score	Z score	Raw Score (Saudi Students)	Question
medium	1	57.90	0.79	42	P1
medium	2	54.90	0.49	42	P4
medium	3	53.10	0.31	40	P2
medium	4	51.70	0.17	36	P5
medium	5	44.40	-0.56	6	P3
medium	6	42.50	-0.75	2	P6

The results of Saudi students in the IMO 2022 showed good performance in some questions and slight variation when compared to students from G20 countries. The students achieved their best performance in question P1 (T = 57.90), which ranked first, followed by question P4 (T = 54.90) in second place, and question P2 (T = 53.10) in third place. Question P5 (T = 51.70) came in fourth, while question P3 (T = 44.40) ranked fifth. The lowest performance was recorded in question P6 (T = 42.50), which ranked sixth.

This indicates that the question where Saudi students performed most comparably to their peers from other countries was question P1, whereas the weakest performance was in question P6, which reflects the persistence of challenges in some areas of more complex mathematical thinking. Overall, however, the 2022 findings indicate a qualitative shift in performance, with a clear ability to achieve tangible convergence with the international average, demonstrating the evident impact of ongoing training and preparation efforts.

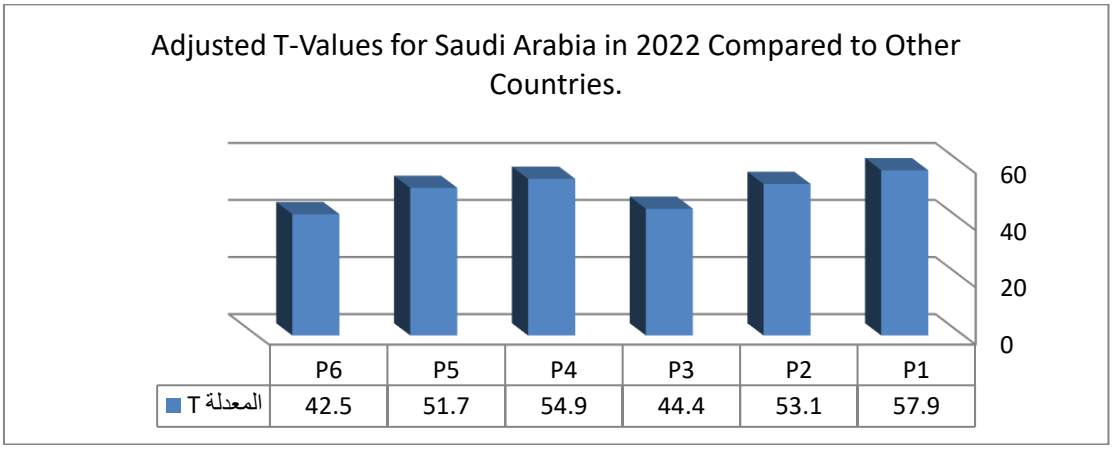
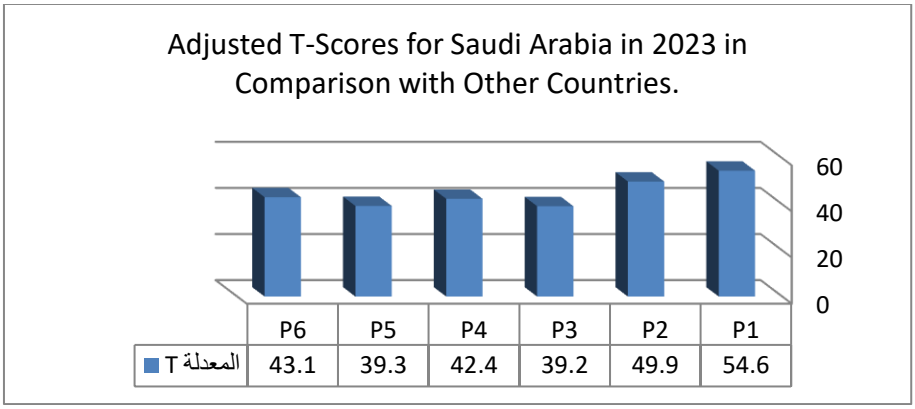


Table (6): Detailed Performance of Saudi Arabian Students in the IMO Questions for the Year 2023 according to Benchmarking Indicators (Z-Score) and Adjusted T-Score Compared to G20 Countries Students

Performance Level	Rank	Adjusted T-Score	Z score	Raw Score (Saudi Students)	Question
medium	1	54.60	0.46	42	P1
medium	2	49.90	-0.01	33	P2
medium	3	43.10	-0.69	0	P6
medium	4	42.40	-0.76	36	P4
low	5	39.30	-1.07	16	P5
low	6	39.20	-1.08	3	P3

The results of Saudi students in the 2023 IMO showed a variation in performance across the six questions compared to students from G20 countries. The best performance was in question P1 (T=54.60), which ranked first, followed by question P2 (T=49.90) in second place, and question P6 (T=43.10) in third place. Question P4 (T=42.40) came in fourth place, reflecting a relatively comparable performance to students from the G20 countries. However, weaker performance was observed in question P5 (T=39.30), which ranked fifth, and the lowest performance was in question P3 (T=39.20), which ranked sixth. This indicates that the performance in these questions was below that of G20 students, especially in questions P5 and P3.

This means that the question in which students performed most comparably to other countries was question P1, while the weakest performance was in question P3, which indicates the continued existence of a gap in advanced mathematical thinking skills and in the ability to tackle more complex problems. Thus, it may be said that the 2023 findings reflect a combination of limited strengths in certain basic areas, alongside a persistent need to strengthen capacities in more complex problems in order to achieve greater balance in overall performance.



Summary of the Results for Question Two

Based on the performance of Saudi students in the IMO from 2019 to 2023, there is a clear variation in performance across the different questions. Students achieved relatively good results in some questions, while their performance in others was below average when compared to students from the G20 countries.

In most cases, question P1 was where Saudi students performed best, often showing results close to those of other countries. However, their performance declined in other questions—particularly in questions P3 and P4—which consistently showed weaker outcomes across multiple years.

The performance gap was most evident in the years 2020 and 2023, where the adjusted T-SCORE for some questions revealed a significant disparity between Saudi students and their G20 peers.

Overall, the results highlight the need to enhance students' skills in tackling more complex questions and challenges that require advanced mathematical competencies, in order to ensure balanced performance across all IMO questions.

In answering the third question, we will explore the topics of the problems and their difficulty classifications.

Question Three:

What is the performance level of Saudi students in the IMO tests between 2019 and 2023, based on a benchmarked comparison classified by problem topics?

To answer this question, the researcher analyzed the detailed results extracted from the IMO databases and categorized them according to the mathematical topic to which each problem belongs.

This analysis relied on the statistical indicator of the adjusted score, the T-SCORE, to measure the average relative performance of students in each topic. This provides an accurate understanding of their strengths and weaknesses across different sub-topics.

Table (7) presents the average adjusted T-Scores of student performance by topic over the five-year period, along with the performance levels classified based on standardized benchmark values.

Table:(7) Average T-Scores of Saudi students' performance by topic in the IMO during the period 2019–2023, with performance levels based on benchmark comparisons.

No.	**Sub-topic**	**Questions**										
			2019/ P3	2019/ P5	2021/ P1	2021/ P5	2022/ P4	2022/ P5	2022/ P6	Average	Performance Level	
1	Probability	Question No. & Year	T	41.7	49.9	33.6	41.9	54.9	51.7	42.5	45.17	medium
2	Circle	Question No. & Year	T	43.6	43.8	57.4	57.9	49.9	-	-	50.52	medium
3	Sequences	Question	T	29.1	42.0	39.2	42.4	39.3	-	-	38.4	low

		No. & Year										
4	Number Theory	Question No. & Year	T	33.3	41.3	53.1	-	-	-	-	42.55	medium
5	Polygons	Question No. & Year	T	45.0	43.1	-	-	-	-	-	44.05	medium
6	Analytical Geometry	Question No. & Year	T	44.4	-	-	-	-	-	-	44.4	medium
7	Distance Formula	Question No. & Year	T	42.8	-	-	-	-	-	-	42.8	medium
8	Functions	Question No. & Year	T	53.0	-	-	-	-	-	-	53.0	medium
9	Analysis	Question No. & Year	T	54.6	-	-	-	-	-	-	54.6	medium
10	Factorial	Question No. & Year	T	40.9	-	-	-	-	-	-	40.9	medium
11	Permutation	Question No. & Year	T	45.4	-	-	-	-	-	-	45.4	medium
12	Truth Tables	Question No. & Year	T	30.0	-	-	-	-	-	-	30.0	low
13	Quadrilateral	Question No. & Year	T	26.4	-	-	-	-	-	-	26.4	low

This table shows the performance of Saudi students in the subtopics of IMO tests between 2019 and 2023 as follows:

Probability: Students' performance in this topic remained at a medium level throughout the years, reflecting relative stability and acceptable proficiency with an average T-Score of 45.17.

Circle Geometry: Students showed good performance in circle geometry, maintaining a medium level from 2019 to 2023, with an average T-Score of approximately 51.02, indicating a satisfactory understanding of geometric concepts and their ability to solve related problems.

Sequences: Students exhibited a clear weakness in this topic, with results classified mostly as low level, averaging a T-Score of 38.4. This indicates significant challenges in dealing with sequence problems and highlights the need to intensify efforts to improve future results.

Number Theory: The average performance in number theory was at a medium level, with an average T-Score of 42.55. Although this shows a tendency toward improvement, students need to strengthen their understanding to achieve better results.

Polygons: Students demonstrated a medium level of performance in polygons, with an average T-Score of 44.05, indicating a good grasp of the concepts in this topic.

Topics Appearing only Once between 2019 and 2023:

Analytic Geometry: Students showed a medium level of performance with an average T-Score of 44.4 in 2023.
Distance Formula: Students performed at a medium level with an average T-Score of 42.8 in 2020.
Functions: Students showed a medium level with an average T-Score of 53 in 2019.
Calculus: Students demonstrated a medium level of performance with an average T-Score of 54.6 in 2023.
Product: Students had a medium performance with an average T-Score of 40.9 in 2019.
Permutation: Students showed a medium performance with an average T-Score of 45.4 in 2020.
Truth Tables: Students showed a relatively low performance, with an average T-Score of 30 in 2020.
Quadrilateral: Students showed a relatively low performance, with an average T-Score of 26.4 in 2020.

Summary of Findings for Question Three:

Overall, students' performance in the subtopics from 2019 to 2023 showed significant variation. Some topics demonstrated relative stability and acceptable performance, such as Number Theory, Polygons, Probability, and Circle Geometry, where students achieved "medium" levels with average T-Scores of 42.55, 44.05, 45.17, and 51.02, respectively. This reflects a good grasp of these topics. However, there is still considerable room for improvement, especially in more complex problems.

On the other hand, students showed clear weakness in some topics like Sequences, where performance was at a "low" level in most years with an average T-Score of 38.4. This indicates evident challenges in dealing with these topics and highlights the need to intensify educational efforts and provide additional support to improve performance.

For topics that appeared only once during the study period, such as Analytic Geometry, Distance Formula, Functions, Calculus, Product, and Permutation, students demonstrated average performance, indicating an acceptable understanding of these concepts.

Conversely, students performed poorly in topics such as Truth Tables with an average T-Score of 30, and Quadrilaterals with an average T-Score of 26.4, showing notable difficulty in handling these areas.

STUDY CONCLUSIONS:

1. Saudi students showed notable variability in their performance over the years, achieving improvement and good results in years such as 2019 and 2022. Conversely, the year 2020 witnessed a significant decline in results, coinciding with the COVID-19 pandemic and its associated impacts on education and learning methods. This highlights the major challenges in maintaining consistent progress during that exceptional period.
2. Overall, Saudi students performed best on the first question (P1) among the six questions in most years, with performance close to that of G20 countries. This question is generally easier, which typically leads to higher scores for students worldwide. However, they showed weaker performance on other questions, particularly the third (P3) and fourth (P4) questions in many years.
3. Saudi students demonstrated average performance in several subtopics such as Probability, Polygons, and Number Theory. Despite this moderate performance, there remains room for improvement to achieve higher mastery levels in these areas.
4. Students faced significant challenges in topics like Sequences, Truth Tables, and Quadrilaterals, where their results were below average (low). This weakness indicates the need to focus more intensively on these subjects in future to improve performance.
5. Saudi students need to enhance their skills in handling more complex questions and advanced mathematical challenges to ensure balanced and competitive performance with students from other countries across various questions and topics.

RECOMMENDATIONS

In light of the study's findings, the researcher recommends the following:

1. Enhancing qualitative training to improve performance in the upcoming years, while continuing to support and strengthen the factors that contributed to the notable improvement in 2022, aiming for greater stability and closer alignment with the average performance of participating countries.

2. Introducing more in-depth qualitative exercises and employing teaching strategies that simulate the level of complexity found in the IMO, to bridge the performance gap with the G20 countries and achieve balanced results across all questions.
3. Enriching mathematics curricula with problems that challenge students' abilities and stimulate their creativity and talents in various branches of mathematics, including number theory, geometry, algebra, statistics, probability, and combinatorics.
4. Including problem-solving skills and strategies as a core part of mathematics teacher preparation and qualification programs.
5. Considering the Mathematical Olympiad as an important tool to nurture gifted and talented students in mathematics, and expanding its implementation across all grade levels.
6. Strengthening training in topics where students showed weaknesses, such as sequences, truth tables, and quadrilaterals, through focused educational programs and additional support. At the same time, improving skills in topics with average performance, like circles and number theory, to achieve higher levels of mastery in complex problems.

Proposed Research Topics:

In light of the study's findings, the researcher suggests conducting the following studies:

1. Analysis of Saudi students' performance in the International Mathematical Olympiad: A normative comparative study across different mathematical topics.
2. Factors influencing the performance of Saudi students in specific topics during the International Mathematical Olympiad.
3. Mathematics teaching strategies and their impact on the performance of Saudi students in the International Mathematical Olympiad according to topics.
4. Global trends in Mathematical Olympiad performance: A topical comparison between Saudi students and students from other countries.
5. Evaluation of preparation strategies for the Mathematical Olympiad tests and their impact on the performance of Saudi students.
6. The role of extracurricular activities in improving Saudi students' performance in the International Mathematical Olympiad: A study of specific topics.
7. The impact of educational culture on Saudi students' performance in International Mathematical Olympiad topics: A comparative analysis.

Conflict-of-Interest Statement: All authors declare that there is no conflict of interests.
Funding: This research has not received grants from any public, commercial or not-for profit agency.

REFERENCES

Avotina, M., & Šuste, A. (2016). *What mathematics teachers know about problems of Mathematical Olympiads*. Proceedings of the 17th International Conference" Teaching Mathematics: Retrospective and Perspektives.

Benchmarking. (

Benchmarking in higher education. (1998). Commonwealth Higher Education Management Service.

Berg, A. (2021). Statistical Analysis of the International Mathematical Olympiad. *The Mathematical Intelligencer*, *43*(1), 105-112. http://doi.org/10.1007/s00283-020-10015-z.

Camarena Gallardo, P. (2014). La matemática social en el desarrollo integral del alumno. *Innovación educativa (México, DF)*, *14*(65), 143-149.

Campbell, J. R., & Walberg, H. J. (2010). Olympiad Studies: Competitions Provide Alternatives to Developing Talents That Serve National Interests. *Roeper Review*, *33*(1), 8-17. http://doi.org/10.1080/02783193.2011.530202.

Canellidis, V., & Sofianopoulou, S. (2022). Analytics Framework for Comparing National Performance Achievements in International Mathematical Olympiads. *Education Sciences, 12*(11). https://doi.org/10.3390/educsci12110838.

Dhlamini, J. J., & Mogari, D. (2011, 2011). *Using a cognitive load theory to account for the enhancement of high school learners' mathematical problem solving skills*. he ISTE 2011 International Conference: Towards Effective Teaching and Meaningful Learning in MST.

Djukić, D., Janković, V., Matić, I., & Petrović, N. (2006). *The IMO compendium: A collection of problems suggested for the international mathematical Olympiads: 1959–2004*. Springer.

Energy poverty affects women the most: At a G20 event, Brasil launches clean cooking initiatives. (2024). Available online: https://g20.org/en/?activeAccordion=c57ea798-8ff7-436f-9d0-7f8c03f97b03 (accessed on 5 October 2024).

Jung, J. Y., & Lee, J. (2021). After the International Mathematical Olympiad: The Educational/Career Decisions and the Development of Mathematical Talent of Former Australian Olympians. *Gifted Child Quarterly*, *65*(3), 235-261. http://doi.org/10.1177/0016986221991160.

Lannie, A. L., & Martens, B. K. (2004). Effects Of Task Difficulty And Type Of Contingency On Students' Allocation Of Responding To Math Worksheets. *Journal of Applied Behavior Analysis*, *37*(1), 53-65. http://doi.org/https://doi.org/10.1901/jaba.2004.37-53.

Lynch, R., Patten, J. V., & Hennessy, J. (2013). The impact of task difficulty and performance scores on student engagement and progression. *Educational Research*, *55*(3), 291-303. http://doi.org/10.1080/00131881.2013.825165.

Moussa, M. A. (2021). Predictive Validity of Cognitive Load Patterns in Mathematical Problem-Solving Stereotypical Thinking in the Inferential Statistics Course Among Psychology Department Students.

Navarro Cendejas, J. (2017). Talento matemático excepcional y destino profesional. Trayectorias de participantes mexicanos en olimpiadas internacionales de matemáticas. *Innovación educativa (México, DF)*, *17*(73), 49-77.

Olympiad in mathematics (IMO). (2024). Available online: https://www.mawhiba.org/Ar/programs/competitions/International-Competitions/International-Olympiad/Math/Pages/default.aspx (accessed on 11 September 2024).

Santiago, P., & Alves, F. R. V. (2022). Math for olympiad: a didactic proposal for high school from the perspective of the international mathematical olympiad. *Journal of Advanced Sciences and Mathematics Education*, *2*(2), 97-108.

Shimizu, Y., & Vithal, R. (2023). *Mathematics curriculum reforms around the world: The 24th ICMI study*. Springer Nature.

Tschisgale, P. L., Steegh, A., Petersen, S., Kubsch, M., Wulff, P., & Neumann, K. (2024). Are science competitions meeting their intentions? a case study on affective and cognitive predictors of success in the Physics Olympiad. *Disciplinary and Interdisciplinary Science Education Research*, *6*(1), 10. http://doi.org/10.1186/s43031-024-00102-y.