Journal of Mathematics and Science Teacher
2025, 5(3), em084
e-ISSN: 2752-6054
https://www.mathsciteacher.com/

MODESTUM

Research Article **OPEN ACCESS**

Strands and cognitive demand levels: Examining university entrance exam questions across three countries

Seyedehkhadijeh Azimi Asmaroud [1*], Yasemin Gunpinar [2], Sebnem Atabas [3], Maryam Zolfaghari [4]

[1] Virginia State University, Petersburg, VA, USA
[2] The University of Texas at Tyler, Tyler, TX, USA
[3] University of Saint Joseph, West Hartford, CT, USA
[4] Kent State University, Kent, OH, USA
***Corresponding Author:** Sazimi@vsu.edu

Citation: Azimi Asmaroud, S., Gunpinar, Y., Atabas, S., & Zolfaghari, M. (2025). Strands and cognitive demand levels: Examining university entrance exam questions across three countries. *Journal of Mathematics and Science Teacher, 5*(3), em084. https://doi.org/10.29333/mathsciteacher/16847

ARTICLE INFO

Received: 09 Apr. 2025
Accepted: 12 Aug. 2025

ABSTRACT

University entrance exams (UEEs) play a crucial role in higher education admissions worldwide, influencing curriculum design and teaching practices. This study examines the mathematics questions in UEEs in Iran, Turkey, and the USA from 2021 to 2024, comparing their content based on National Council of Teachers of Mathematics strands and Cognitive Demand Level framework. The results of the study showed that algebra comprised a higher percentage of questions compared to other content areas across all the countries. Additionally, the UEEs questions in all three countries predominantly consisted of procedural-level questions (either procedures without connections or procedures with connections). Based on these findings, suggestions and recommendations were provided to enhance the balance of question types and promote deeper conceptual understanding in assessments.

Keywords: textbook analysis, cognitive demand level, university entrance exams

INTRODUCTION

One significant type of assessment across many countries is university entrance exams (UEEs). These assessments are administered nationwide in many countries and affect the school education system (Borji & Sánchez, 2019; Otaç, 2019; Zhang, 2016). After high school, students in many countries must take UEEs to qualify for admission into higher education (Coles et al., 2023; Gölogluu Demir & Kaplan Keles, 2021). In some places, like China (Jiang et al., 2019; Zhang, 2016), Iran (Borji & Sánchez, 2019), Turkey (Otaç, 2019), and Japan (Watanabe, 2015), these exams are the primary or sole criterion for university acceptance. While the USA does not have a standardized national UEE comparable to those countries discussed, many postsecondary institutions consider students' performance on high-stakes tests (e.g., scholastic aptitude test [SAT]) in their admission process (Moore et al., 2019). With the high importance of these exams on students' admission to universities, sample questions from previous years' UEEs are often used as key resources for students to practice problems and prepare effectively for the exam (Ontong & Bruwer, 2020).

Institutions aim to admit students with the highest potential to meet the demands of their programs, and admitted students are expected to achieve strong outcomes upon program completion. From an institutional standpoint reliable admission tools are essential for making valid decisions about student selection (Akış, 2020). However, the overemphasis on standardized exams brings the risk of narrowing the curriculum and putting teachers under pressure to focus on preparing students for the exam rather than focusing on critical thinking and conceptual understanding of the mathematical ideas (Coles et al., 2023; Gölogluu Demir & Kaplan Keles, 2021). Furthermore, numerous scholars have emphasized the importance of students' conceptual understanding of mathematics, leading to various efforts aimed at improving their mathematical thinking skills (e.g., Henningsen & Stein, 1997; Hurrell, 2021).

Given that assessment tasks influence teachers' instructional practices (Coles et al., 2023; Gölogluu Demir & Kaplan Keles, 2021) and that the selection of assessment tools plays a crucial role in shaping students' learning, this study examines how UEE questions in Turkey, Iran, and the USA require students to engage in mathematical thinking and problem-solving. It also explores the similarities and differences in the types of questions used in these exams, highlighting variations in cognitive demand across the three countries.

Research Questions

This study explores the types of questions and the cognitive demand level (CDL) of the questions used in these countries. Below are the research questions we aim to address in this study:

1. Which mathematics topics were included in the UEEs in Iran, Turkey, and the USA?
2. What are the CDLs of questions in UEEs in Iran, Turkey, and the USA from 2021 to 2024?
3. What, if any, are differences in the cognitive demand of UEE questions across the three countries and shifts from 2021 to 2024?

LITERATURE REVIEW

Global Perspective on Curriculum Reform in Mathematics Education

Globalization has become increasingly prominent in mathematics education, shaping educational discourse and curriculum changes. According to the Peterson Institute for International Economics (2024), globalization refers to the increasing interconnectedness of economies, cultures, and populations through trade, technology, investment, migration, and information exchange. Adjustments made in one country or region are often mirrored in other countries within a few years (Atweh & Clarkson, 2002; Hossain, 2022).

The growing interconnectedness driven by globalization, often fueled by technological advancements, has also intensified the focus on educational practices and outcomes across nations. International comparative studies like TIMSS and PISA have raised awareness among countries about each other's mathematics curricula as well as highlighting the global strengths and weaknesses in student performance (Shimizu & Vithal, 2023). These results significantly influenced curriculum reform in various ways across different countries. These school mathematics reforms are often conducted with changes in all different aspects of the curriculum such as mathematics content, pedagogical approaches, and assessment and examinations including but not limiting to UEEs (Shimizu & Vithal, 2023). In Iran and Turkey, UEEs have been influenced by international assessments and curriculum changes, while the TIMSS performance is seen as one of the driving forces in curriculum changes (Bartolini Bussi et al., 2023; Eren, 2024). In Turkey, similar to Iran, 2015 PISA results showed that these two countries might have low-performing schools (Atac, 2019), while The TIMSS 2023 results show a decline in USA math scores and improvements in Turkey. The result of this assessment affected the education system in many countries. For example, in Turkey, the content of the college examination system has altered over the recent years (Atac, 2019) and it has included skill-based questions in UEEs to increase success in international exams (Eren, 2024). While TIMSS and PISA provide important international benchmarks, UEEs remain the most decisive factor in university admissions, making them the primary focus of student preparation. Given their significance, understanding how these exams align with broader curriculum reforms is essential to evaluating their role in shaping students' mathematical competencies.

Learning in Mathematics

As curriculum and assessment reforms continue to shape educational systems, increasing attention has been given to students' conceptual understanding of mathematics and their ability to think, reason, and solve problems effectively (Eren, 2024; Henningsen & Stein, 1997). These changes reflect a broader shift in mathematics education, emphasizing the development of deeper cognitive skills rather than the memorization of procedures). The education vision document, which emphasizes 21st century skills, highlights the need for students to develop problem-solving and critical-thinking skills (Eren, 2024, p. 47). Similarly, educational priorities in many countries have shifted toward active and creative learning processes, instead of viewing mathematics as a static collection of facts and procedures (Joklitschke et al., 2022). There has also been a greater emphasis on conceptual knowledge, as seen in various reform efforts worldwide (Crooks & Alibali, 2014; Hussein & Csikos, 2023). Recent reform efforts in the USA, such as the standards from the National Council of Teachers of Mathematics (NCTM) and the common core state standards (CCSS), have emphasized the importance of students developing both conceptual understanding and procedural skills (Crooks & Alibali, 2014). This perspective shapes not only what students should learn but also the activities that both students and teachers should engage in. It also determines what students are required to be assessed on and how they are prepared to enter higher-level education in colleges.

Importance of Mathematical Tasks

The importance of implementing high-quality mathematical tasks has never been more evident, as they not only shape students' learning opportunities (Feldman et al., 2016; Hsu & Yao 2023; Smith & Stein, 1998) but also serve as key tools for classroom instruction to enhance students' learning (Shimizu et al., 2010). Tasks that involve deep mathematical thinking, including tasks that require analysis and synthesis, give students chances to develop mathematical skills such as problem-solving, critical thinking, and creativity (Paredes et al., 2020; Yeo, 2007). Research also shows that people who understand the concepts behind a procedure are more likely to successfully generalize it and apply it to new and unfamiliar problems (Crooks & Alibali, 2014).

Building on Doyle's (1988) work, Stein and Smith (1998) proposed the task analysis guide comprised of four categories of tasks that create different learning opportunities for students; thus tasks that are at different levels of cognitive demand:

(1) memorization,

(2) procedures without connections (PwoC),

(3) procedures with connections (PwC), and

(4) doing mathematics.

According to this framework, memorization tasks involve producing previously learned facts, rules, formulae, or definitions. There is no need to use even a procedure to solve memorization tasks. It is highly dependent on memorizing the previously learned information. PwoC tasks require using procedures that are specifically called for or that are highly obvious based on previous instructions. These tasks do not require students to understand how the procedure works or to explain their thinking but focus on producing correct answers by using the procedure being called. These two levels of tasks are named low-level tasks. The next two levels of cognitive demand comprise high-level tasks. In the PwC tasks, although there may be some suggested pathways, there is not an explicit procedure to follow. These tasks require a deeper understanding of mathematical concepts and a higher level of cognitive effort to provide an answer. Doing mathematics tasks require complex and non-algorithmic thinking. Students explore and understand the nature of mathematical concepts, processes, or relationships.

Therefore, studies have shown that higher-level cognitive demand tasks are associated with deeper learning (Gilbert, 2016; Paredes et al., 2020) and the ability to transfer knowledge to new situations (Henningsen & Stein, 1997). Students who are exposed to tasks with high CDL demonstrate greater mathematical proficiency and problem-solving skills (Paredes et al., 2020; Silver et al., 1995). Conversely, low-level tasks often lead to a limited understanding of mathematical concepts (Boaler, 2002). Yeo (2007) stated that procedural tasks and word problems in which students practice what they have been taught earlier are not mathematically rich tasks. Considering the importance of implementing high-quality mathematical tasks, analyzing CDLs of the tasks lies in their impact on students' mathematical learning and understanding. One aim of the current study is to assess the level of cognitive demand of the tasks that students are assessed on in the UEE, which plays a significant role in shaping students' preparation for higher education. By examining these tasks, we aim to better understand how they contribute to students' mathematical development and whether they align with the standards of deep learning and problem-solving emphasized in contemporary educational reforms. For this study, we focused on three countries: Iran, Turkey, and the USA Additionally, we discuss university entrance assessments in other countries to provide a broader international perspective. The following section provides more details about the types of assessments in these countries.

University Entrance Exams in the Iran, USA, and Turkey

Every year, millions of high school students in different countries take the national UEEs, and their results are the main factor determining entry into universities (Borji & Sánchez 2019; Hong & Choi, 2011; Kusayanagi, 2013).

Iran's UEEs are conducted twice a year by the National Organization of Educational Testing and include questions on topics covered in high school textbooks, such as literature, science subjects (e.g., physics, chemistry, and biology for science tracks), religious studies, language, and mathematics specialized subjects are also covered depending on the chosen track (Parviz, 2023). The mathematics portion consists of 40 multiple-choice questions out of approximately 200 total questions (Borji & Sanchez, 2019). Due to the limited number of spots in top universities, these exams are extremely competitive, with students often preparing for years.

In the USA, institutions use different tests depending on the state where they are located (Koretz et al., 2016). SAT is one of these tests that has long been a key element of college admissions, serving as a standardized assessment of academic readiness (Krishnaveti & Rawat, 2024). The SAT exams in the USA were recently revised to focus more on what students learn during high school (Koretz et al., 2016). A study by Krishnaveti and Rawat (2024) showed a "significant decrease in the difficulty of the SAT math section over time, alongside a decline in students' math performance" (p. 1) from 2018 to 2023. Furthermore, this study revealed that before 2016, the key focus areas of SAT were "arithmetic, numbers and operations, algebra, functions, geometry and data analysis" (p. 6) and later it included questions related to trigonometry and complex numbers. After the 2016 versions, the penalty for incorrect answers was eliminated and there was a reduction in the number of math sections from three to two while the number of questions increased from 54 to 58 (Krishnaveti & Rawat, 2024). The College Board (2015) explained,

> The redesigned SAT will be better and more clearly aligned to best practices in classroom instruction so that the most effective preparation for the SAT is the development of the ELA/literacy and math skills taught in great courses across the disciplines … No longer will the SAT stand apart from the work of teachers in their classrooms (p. 13).

In contrast to countries such as Iran and Turkey, in the USA, the significance of SAT exams is relatively low since students' scores are considered alongside "admission essays, high school grades, and extracurricular activities" (Watanabe, 2015, p. 84).

The UEE system in Turkey, known as the Higher Education Institutions Exam (YKS), is a standardized test administered by the Measuring, Selection and Placement Center (ÖSYM). It serves as the sole pathway for students to gain admission to universities within the Turkish education system. Candidates' scores from the exam are combined with their high school GPA to determine university placements. Participation in the exam has steadily increased over recent years, with 2,607,715 high school graduates taking it in 2021, 3,243,334 in 2022, and 3,527,443 in 2023 (Dogru Tercihler, 2024). The exam content is based on the curriculum taught to students throughout their basic and secondary education in Turkey (Otaç, 2019, p. 14). The YKS consists of three main sections: The Basic Proficiency Test (TYT), the Field Proficiency Test (AYT), and the Foreign Language Test (YDT). The first session, TYT, comprises 125 questions, including 40 math questions. The AYT features 166 questions, including 40 math questions. Both sections are multiple-choice exams, with each question offering five answer options. The initial phase of the exam, the Basic Proficiency Test (TYT), is common to all students, regardless of their academic focus, and includes questions on Turkish, social sciences, basic mathematics, and science.

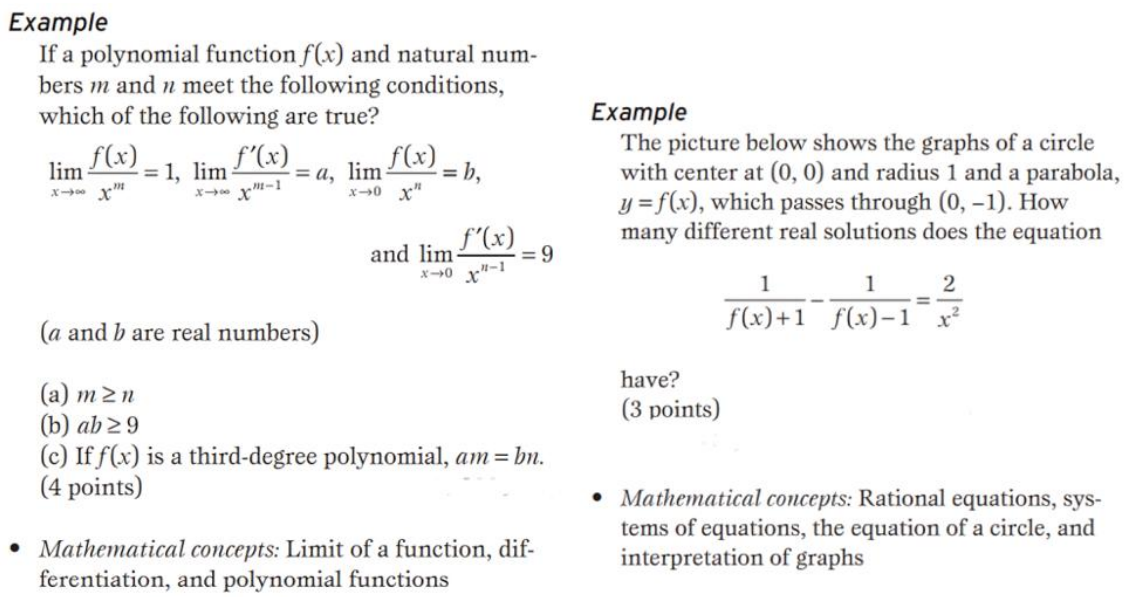


Figure 1. PwC questions in Korean UEEs (Hong & Choi, 2011)

Types of Assessments in Other Countries

UEEs play a major role in shaping mathematics education in many countries (Hong & Choi, 2011), yet there has been a limited number of studies examining these types of examinations (Borji & Sánchez, 2019; Watanabe, 2015).

In their study of the Korean college entrance examination, Hong and Choi (2011) describe the characteristics of the mathematics questions. This exam, which includes 30 questions–21 multiple-choice and 9 constructed-response questions–covers topics in high school–level mathematics "such as precalculus (e.g., trigonometry, functions, equations, and inequalities), differential and integral calculus, discrete mathematics, and probability and statistics" (p. 209). Additionally, questions that assess reasoning and problem-solving are rewarded with 3 or 4 points, compared to 2 points for questions that assess computation and understanding. This study also showed that the type of problem-solving and reasoning questions in these exams requires students to make connections between different mathematics topics, as categorized in PwC level tasks in Smith and Stein (1998) framework. **Figure 1** shows two examples of the types of questions in the Korean college entrance examination mentioned in Hong and Choi (2011).

In their analysis of the UEEs in Iran and Spain Borji and Sánchez (2019) found that most UEE questions in both countries focused on algebraic representation and a very small percentage focused other representations and their relationship.

THEORETICAL PERSPECTIVE

This study utilizes the framework developed by Smith and Stein (1998) to analyze the CDLs of UEE questions. Building on Doyle's (1988) work, Stein and Smith (1998) proposed the task analysis guide comprised of four categories of tasks at different levels of cognitive demand:

(1) memorization,

(2) PwoC,

(3) PwC, and

(4) doing mathematics.

These levels are used to describe the complexity and cognitive effort required to complete mathematical tasks.

Memorization tasks are not ambiguous, require minimal cognitive effort, no deep understanding of mathematical concepts, and no connection to mathematical ideas, and involve recalling formulas, facts, rules, or definitions. There is no need to use even a procedure to solve memorization tasks. It is highly dependent on memorizing the previously learned information. PwoC tasks involve performing routine procedures, do not require understanding the underlying concepts, and require limited cognitive demand. Tasks at this level require using procedures that are specifically called for or that are highly obvious based on the previous instruction. These tasks do not require students to understand how the procedure works or to explain their thinking but focus on producing correct answers by using the procedure being called. Students do not make connections to mathematical concepts at this level. PwC tasks involve procedures that are linked to conceptual understanding or the use of the procedure to develop a deeper level of understanding. Although there may be some suggested pathways to solve a problem, there is not an explicit procedure to be followed. These tasks require some reasoning and understanding of the relationships between mathematical ideas and require some degree of cognitive effort. Doing mathematics requires non-algorithmic thinking and problem-solving. These tasks involve exploring, conjecturing, self-monitoring, and solving non-routine problems. They often require considerable cognitive effort, and students need to analyze the tasks.

Table 1. Comparison of the number of questions in each NCTM content strand

Strands from NCTM		Iran					Turkey					SAT USA				
		2021	2022	2023	Total	%	2021	2022	2023	Total	%	Test 1	Test 2	Test 3	Total	%
Numbers & operations		12	7	3	22	15.20	4	5	9	18	15.00	5	2	2	9	5.56
Algebra		23	21	16	60	**41.40**	20	20	15	55	**45.80**	39	38	37	114	**70.37**
Geometry		13	8	8	29	20.00	7	5	6	18	15.00	4	6	4	14	8.64
Measurement		0	3	3	6	4.10	6	7	7	20	16.70	3	3	6	12	7.41
Data analysis & probability		4	5	4	13	9.00	1	1	1	3	2.50	3	5	5	13	8.02
Other topics	Set theory	1	3	3	7	4.80	1	1	1	3	2.50	0	0	0	0	0.00
	Graph theory	2	2	2	6	4.10	0	0	0	0	0.00	0	0	0	0	0.00
	Logic	0	1	1	2	1.40	1	1	1	3	2.50	0	0	0	0	0.00
Total		55	50	40	145	100	40	40	40	120	100	54	54	54	162	100

Smith and Stein (1998) categorized memorization and PwoC tasks as lower-level cognitive demands, while PwC and doing mathematics tasks were classified as higher-level cognitive demands.

METHODS

This study is part of a broader research initiative aimed at examining UEEs and the associated mathematics curricula across multiple countries. The present research focuses on the UEEs of three countries: Iran, Turkey, and the USA. These countries were chosen due to the authors' expertise in their educational systems and linguistic familiarity. The analysis spans three consecutive years (2021, 2022, and 2023) for the Iranian and Turkish UEEs, while for the USA, the latest publicly released SAT practice test (test 1, test 2, and test 3) by the College Board (2015) was employed, given the restricted access to actual SAT exam content.

The total number of questions analyzed from the UEEs was 145 for Iran, 120 for Turkey, and 162 for the SAT. Iranian UEEs showed a decreasing trend in the number of questions each year, starting with 55 in 2021, reducing to 50 in 2022, and 40 in 2023. Conversely, Turkish UEEs consistently included 40 multiple-choice questions per year, while the SAT practice test contained 54 items comprising both multiple-choice and short-answer questions. The coding of Iranian UEEs questions was conducted by the first and third authors, while the Turkish exam questions were coded by the next two authors due to their proficiency in the language. SAT questions were reviewed collaboratively by all authors to ensure a balanced and comprehensive analysis.

To categorize the questions, the NCTM strands were utilized, organizing them into numbers and operations, algebra, data analysis and probability, geometry, and measurement. The analysis framework was designed to capture variations in the representation of these strands across the UEEs of the three countries. While the majority of the questions aligned with the established NCTM categories, certain advanced topics that extended beyond the traditional strands were grouped under "other topics." These included set theory, graph theory, and logic.

The study then applied the task analysis guide by Smith and Stein's (1998) to analyze the complexity of the questions across all strands, and questions were categorized into the four levels mentioned by them: Memorization, PwoC, PwC, and doing mathematics. To maintain coding reliability, the authors jointly analyzed a 10% random sample of the questions, achieving high inter-rater reliability (87.5% for Iran, 86.67% for Turkey, and 84.61% for SAT). Discrepancies were resolved through collaborative discussion.

After coding the data, descriptive statistics were applied to summarize the distribution of topic coverage across the UEEs of the three countries and to categorize questions according to the task analysis guide. This quantitative analysis was complemented by a qualitative exploration of the questions, focusing on their CDLs to identify patterns, differences, and similarities across the countries.

RESULTS

In this section, we present the results of the data analysis for the UEE questions from Iran, Turkey, and the USA in 2021-2023. The findings are structured according to the research questions. First, we provide an analysis of the content covered in the UEE questions based on the NCTM strands. Following this, we compare the CDL required by the UEE questions in the three countries. We then analyze the types of connections present in the PwC-level questions, and we offer examples to illustrate the differences in question types at the PwC and doing mathematics levels in two tables.

Mathematics Question Content

Table 1 shows the content covered in the UEE questions based on the NCTM strands. The total number of questions in Iran UEE decreases by five each year, whereas in Turkey and the SAT exams, the number of questions remains constant.

Regarding the total number of questions in each category through all three years, it was shown that algebra comprised a higher percentage of questions compared to other content areas across all the countries, with 41.1% in Iran, 45.8% in Turkey, and 70% in the SAT exams. The SAT exams, in particular, included more algebra content questions than the exams from the other two countries. Geometry was the second most dominant category in Iran and the USA, accounting for 20% and 8.64% of the questions, respectively. In the Turkish and USA UEEs, categories other than algebra had nearly the same percentage of questions. In Turkey,

Table 2. Number of the questions in each CDL in three countries

CDL of all questions		SAT USA					Iran					Turkey				
		Test 1	Test 2	Test 3	Total	%	2023	2022	2021	Total	%	2023	2022	2021	Total	%
		n (%)					n (%)					n (%)				
Low level	Memorization	3 (5)	2 (4)	1 (2)	6	3.7	1 (3)	0 (0)	0 (0)	1	1.0	0 (0)	0 (0)	0 (0)	0	0.0
	PwoC	32 (60)	35 (65)	34 (63)	101	62.3	8 (20)	5 (10)	1 (2)	14	10.0	9 (22)	4 (10)	5 (12)	18	15.0
High level	PwC	19 (35)	17 (31)	18 (33)	54	33.3	23 (57)	25 (50)	33 (60)	81	56.0	29 (73)	23 (58)	25 (63)	77	64.0
	Doing mathematics	0 (0)	0 (0)	1 (2)	1	0.6	8 (20)	20 (40)	21 (38)	49	34.0	2 (5)	13 (32)	10 (25)	25	21.0
Total		54	54	54	162	100	40	50	55	145	100	40	40	40	120	100

Table 3. Example of questions with the PwC CDL and the same NCTM strand with different connection types

Country	Example of linear equations	Type of the connection
Iran (2021, Q54)	If the two lines x + y = 1 and x - y = 3 are the diagonals of a circle and the line 4x + 3y + 5 = 0 is tangent to it, what is the closest distance of the point (4, -2) from the circle?	Connection of system of linear equations to geometry-coordinate plane and measurement
Turkey (2023, Q38)	On the perpendicular coordinate plane, the line 2x + y = 12 and a line d intersect at point A (4, 4). These two lines divide each circle, whose center is point A (4,4), into four regions of equal area. Accordingly, which of the following is the equation of line d? (A) -2x + y = -4 (B) x - 3y = -8 (C) 3x + y = 16 (D) x + 2y = 12 (E) x - 2y = -4	Connection of system of linear equations to geometry-coordinate plane and measurement
USA (practice test 1, Q8)	A teacher is creating an assignment worth 70 points. The assignment will consist of questions worth 1 point and questions worth 3 points. Which equation represents this situation, where x represents the number of 1-point questions and y represents the number of 3-point questions? (A) 4xy = 70 (B) 4(x + y) = 70 (C) 3x + y = 70 (D) x+ 3y = 70	Connection to the real-world situation

measurement, geometry, and numbers and operations each represented about 15%-16%, while in the USA, geometry, measurement, and data analysis and probability each accounted for approximately 7%-8% of the questions. The distribution of questions in the Iran UEEs varied drastically across different categories.

Furthermore, data analysis revealed that 10.34% of the questions in Iran's UEEs and 5% of the questions in Turkey's UEEs fall into content areas not covered by the NCTM strands. Specifically, Turkey's UEEs included questions on set theory and logic, and Iran's UEEs included questions on set theory, graph theory, and logic.

Cognitive Demand Levels in UEEs

Table 2 shows the comparison of the CDL required by the UEE questions in the three countries. As **Table 2** shows, in the USA UEEs, the majority of questions were categorized in lower levels; comprising over 65% of the total number of questions in memorization or lower-level procedural questions (62.3 % PwoC and 3.7% Memorization). Additionally, the UEEs questions in the USA predominantly were procedural-level questions (either PwoC or PwC), 95.6% of the total with questions involving PwoC being a substantial portion (62.3% of all questions). In Iran and Turkey, there was a noticeable emphasis on higher-level cognitive demands, with approximately 90% and 85% of questions, respectively, falling into PwC or doing mathematics level. In both countries, a high percentage of the higher CDLs were concentrated in the PwC category. **Table 2** does not reflect students' performance levels; they only indicate the level of cognitive demand required to solve the questions. One might assume that Iranian and Turkish students would typically perform better in mathematics, yet data from various global tests suggest otherwise, revealing a different trend.

To understand the underlying causes of this disparity, we delved into the nature and construction of the test questions in **Table 3** and **Table 4**. Our focus on **Table 3** was primarily on PwC questions for two reasons:

(a) they constitute the majority of questions in both Iran and Turkey, accounting for over 50% of the total and

(b) as defined by Smith and Stein (1998) they are designed to connect with multiple contents and representations, among other elements.

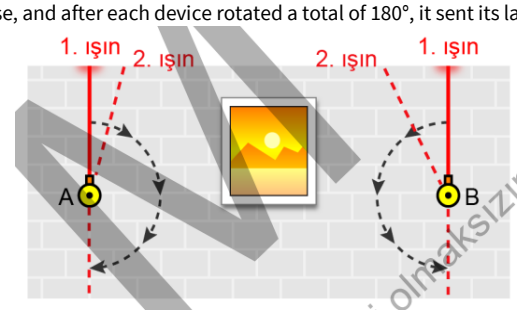
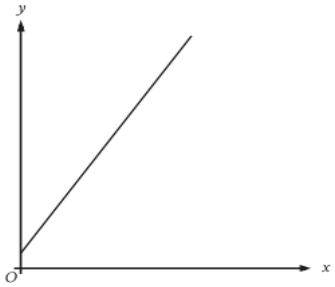
In **Table 3**, we compare questions with the same CDL, PwC, and the same content within the NCTM strand. Algebra was selected for this analysis due to its predominant representation across all three assessments, making it a key area for comparison. Linear Equations which were categorized in the Algebra strand were chosen to show that questions in the USA were more in real-world scenarios but did not interconnect multiple contents, whereas Turkey and Iran questions connected multiple concepts. In **Table 4**, we provide examples of doing mathematics-level questions in three countries to illustrate the variations in these assessment approaches and highlight the differences in question styles and formats.

As **Table 4** shows, the questions in Iran and Turkey at doing mathematics-level are non-routine problems that require students to apply their knowledge in multiple concepts creatively and persistently to arrive at a solution. Doing mathematics level questions from the USA requires students to apply their understanding to select the correct contextual interpretation of the slope.

CONCLUSION AND DISCUSSION

The results of this study revealed that algebra accounted for a larger proportion of questions compared to other content areas. Researchers highlight that readiness for college-level math and technical jobs depends on students acquiring a fundamental understanding of algebra by the end of high school (Loveless, 2013). A study by Otten et al. (2024) showed that 91% of the tasks

Table 4. Comparing doing mathematics level questions in UEE in three countries

	Example of the question in doing mathematics Level
Iran (2022, Q24)	The intersection of the asymptotes of the homographic function $y = \frac{ax+3}{(a+1)x+(a-1)}$ is the minimum point of the function $y = \frac{3}{2}x^2 + x + \frac{5}{6}$. At what point does the graph of this homographic function cross the x-axis? (A) 3 (B) -3 (C) $\frac{3}{2}$ (D) $-\frac{3}{2}$
Turkey (2022, Q32)	Laser devices A and B are placed on a wall where a painting exhibited in a museum hangs. Each of these devices sends its first beam along the wall surface, perpendicular and upward from the ground, and each time it turns to a fixed angle specific to that device, it sends its next beam along the wall surface. As shown in the figure below, device A moved clockwise, device B moved counterclockwise, and after each device rotated a total of 180°, it sent its last ray and stopped.  *Note*. "ışın" means "ray" in English. The 2nd ray coming out of device A and the 3rd ray coming out of device B are perpendicular to each other; The 8th ray coming out of device A and the 5th ray coming out of device B are perpendicular to each other. Accordingly, what is the total number of rays coming out of device A?
SAT (practice test 3, Q9)	 The graph represents the total charge, in dollars, for an electrician for x hours of work. The electrician charges a onetime fee plus an hourly rate. What is the best interpretation of the slope of the graph? (A) The electrician's hourly rate (B) The electrician's onetime fee (C) The maximum amount that the electrician charges (D) The total amount that the electrician charges

implemented in algebra class lesson plans in the USA, across a wide variety of school districts and a diverse group of teachers, fall in the category of PwoC and there was a grand total of 0.2 % in doing mathematics tasks. Thus, we can conclude that all three UEEs can serve as a good benchmark for the knowledge required from students for college-level mathematics.

Considering the emphasis of the standards and research on the importance of developing both conceptual understanding and procedural skills in students (Crooks & Alibali, 2014), the findings of this study suggest including more PwC questions or doing mathematics in the UEEs. Especially since, as mentioned in the results for the SAT exams, most of the questions were categorized at the PwoC-level. This stands in contrast to recent reform efforts in the USA, such as the standards put forth by the NCTM and the CCSS which emphasize the importance of fostering students' conceptual understanding alongside their procedural skills (Crooks & Alibali, 2014).

One suggestion regarding the level of the questions is that questions about reasoning and problem-solving could have more weight than those focused on computation and memorization, similar to Korea's UEE (Hong & Choi, 2011). This approach would result in exam evaluations that better represent students' skills in conceptual understanding and reasoning. Future studies could also further explore the types of connections emphasized in the questions.

The results of our analysis suggest that Iran and Turkey had a higher proportion of high-level CDL questions in their UEEs. However, despite the presence of these higher-level questions, student performance in these countries remains relatively low. This discrepancy indicates that the complexity of test questions alone does not necessarily correspond to higher student achievement. One possible explanation is that, in the USA, the level of questions in textbooks is more closely aligned with those in the UEEs, potentially providing a more effective assessment tool. This hypothesis prompted further comparative analysis of the alignment between textbook questions and UEE questions across different countries, aiming to explore its potential impact on student performance.

Further studies need to be conducted to explore how the mathematics questions in the UEEs align with the goals of the designers and planners of textbooks and curriculum, as well as how effectively these exams prepare students for the college-level curriculum by examining students' performance on UEEs. Moreover, as studies such as Henningsen and Stein (1997) show, classroom-based factors can affect maintaining the level of the task, and students' engagement with mathematical tasks that

were set up to encourage high-level mathematical thinking can decline during their classroom experiences (e.g., Agterberg et al., 2022).

Furthermore, our analysis revealed that some of the questions in Iran and Turkey's UEEs fall into content areas that were not covered by the NCTM strands. This discrepancy highlights a potential misalignment between the international national curricula and the content assessed in these exams. These findings are aligned with the findings of other research such as Borji and Sánchez (2019) that showed that UEEs in Spain cover topics such as "vectors, matrixes, systems of equations, equations of lines and planes, dot and cross products" (p. 7) which is not part of the content covered by the NCTM strands. Also, a study by Hong and Choi (2011) showed Korean UEEs cover topics in trigonometry, "differential and integral calculus, discrete mathematics, and probability and statistics" (p. 209) which is not the expectations from SAT. Moreover, our results are aligned with the results of other studies which show that the SAT puts an increase "focus on data analysis, graphs, and word problems and puts a smaller focus on geometry-related questions" (Krishnaveti & Rawat; 2024, p. 6).

In the USA, the SAT plays a less significant role in college admissions (Watanabe, 2015) compared to countries like Japan, Iran, and Turkey. Thus, further research is needed on other exams, such as the common admission test, as well as additional admission factors in the USA, such as GPA, extracurricular activities, and personal statements, to make comparisons with other countries more comprehensive and meaningful.

Preparation for UEEs in Iran and Turkey is highly stressful, as these exams are the primary factor in university admissions, unlike the USA, where holistic evaluations reduce the weight of standardized tests like the SAT. This difference raises equity concerns and highlights the impact of high stakes testing on students' well-being. Future research could examine factors such as socio-economic background, access to resources, and preparation strategies, as well as how these influence student performance and quality of life during the critical years of preparing for the tests.

Author contributions: SAA: conceptualization, supervision, methodology, project administration, resources, validation, writing – original draft; **YG:** conceptualization, methodology, project administration, resources, validation, writing – review & editing; **SA:** methodology, resources, validation, writing – review & editing; **MZ:** methodology, resources, validation, writing – review & editing. All authors have agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Ethical statement: The authors stated that the study relied exclusively on publicly available resources and did not involve human participants, patient data, or any sensitive/confidential information, ethical approval was not applicable.

AI statement: The authors stated that no AI technologies were used in any part of this study.

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Agterberg, D. A., Oostdam, R. J., & Janssen, F. J. J. M. (2022). From speck to story: Relating history of mathematics to the cognitive demand of tasks. *Educational Studies in Mathematics, 110*, 49-64. https://doi.org/10.1007/s10649-021-10093-6

Akış, D. (2020). *Predicting high school students' academic achievement in international and national programs based on admission examination: A case study from Turkey* [Master's thesis, Bilkent University]. Bilkent University Institutional Repository. http://hdl.handle.net/11693/53711

Atac, C. (2019). *The challenge of pre-examination system in Turkey for teacher candidates and potential solutions to the challenge* [Master's thesis, University of Jyväskylä].

Atweh, B., & Clarkson, P. (2002). Globalized curriculum or global approach to curriculum reform in mathematics education. *Asia Pacific Education Review, 3*, 160-167. https://doi.org/10.1007/BF03024909

Bartolini Bussi, M. G., Milinkovic, J., & Gooya, Z. (2023). The role of values and culture in past mathematics curriculum reforms. In Y. Shiizu, & R. Vithal (Eds.), *Mathematics curriculum reforms around the world. New ICMI study series* (pp. 87-99). Springer. https://doi.org/10.1007/978-3-031-13548-4_6

Boaler, J. (2002). *Experiencing school mathematics: Traditional and reform approaches to teaching and their impact on student learning*. Routledge. https://doi.org/10.4324/9781410606365

Borji, V., & Sánchez, A. (2019). An exploratory analysis of the representations of functions in the university entrance exam in Spain and Iran. *Eurasia Journal of Mathematics, Science and Technology Education, 15*(8), Article em1727. https://doi.org/10.29333/ejmste/106258

Coles, A., Rodríguez-Muñiz, L. J., Mok, I. A. C., Ruiz, Á., Karsenty, R., Martignone, F., Osta, I., Ferretti, F., & Nguyen, T. T. A. (2023). Teachers, resources, assessment practices: Role and impact on the curricular implementation process. In Y. Shiizu, & R. Vithal (Eds.), *Mathematics curriculum reforms around the world. New ICMI study series* (pp. 291-321). Springer. https://doi.org/10.1007/978-3-031-13548-4_18

College Board. (2015). *SAT practice tests*. SAT Suite of Assessments. https://satsuite.collegeboard.org/practice/practice-tests/paper

Crooks, N. M., & Alibali, M. W. (2014). Defining and measuring conceptual knowledge in mathematics. *Developmental Review, 34*(4), 344-377. https://doi.org/10.1016/j.dr.2014.10.001

Dogru Tercihler. (2025). Üniversite sınavına başvuran ve yerleşen aday sayıları [Number of candidates who applied for and were placed in the university entrance exam]. *Dogru Tercihler*. https://dogrutercihler.com/universite-sinavina-basvuran-ve-yerlesen-aday-sayilari/

Doyle, W. (1988). Work in mathematics classes: The context of students' thinking during instruction. *Educational Psychologist*, *23* 2), 167-180. https://doi.org/10.1207/s15326985ep2302_6

Eren, A. Y. (2024). New generation questions in mathematics education. *Current and Advanced Research in Science and Math Education I, 43*.

Feldman, Z., Thanheiser, E., Welder, R. M., Tobias, J. M., Hillen, A. F., & Olanoff, D. (2016). When is a mathematical task a good task? In L. Hart, S. Oesterle, S. S. Auslander, & A. Kajander (Eds.), *The mathematics education of elementary teachers: Issues and strategies for content courses* (pp. 9-24). Information Age Publishing.

Gilbert, M. C. (2016). Relating aspects of motivation to facets of mathematical competence varying in cognitive demand. *The Journal of Educational Research, 109*(6), 647-657. https://doi.org/10.1080/00220671.2015.1020912

Göloglu Demir, C., & Kaplan Keles, Ö. (2021). The impact of high-stakes testing on the teaching and learning processes of mathematics. *Journal of Pedagogical Research, 5*(2), 119-137. https://doi.org/10.33902/JPR.2021269677

Henningsen, M., & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education, 28*(5), 524-549. https://doi.org/10.5951/jresematheduc.28.5.0524

Hong, D. S., & Choi, K. M. (2011). Korean college entrance exams: An inside look. *The Mathematics Teacher, 105*(3), 208-213. https://doi.org/10.5951/mathteacher.105.3.0208

Hossain, M. (2022). Diffusing "destandardization" reforms across educational systems in low- and middle-income countries: The case of the world bank, 1965 to 2020. *Sociology of Education, 95*(4), 320-339. https://doi.org/10.1177/00380407221109209

Hsu, H.-Y., & Yao, C.-Y. (2023). A review of the mathematical tasks framework and levels of cognitive demand. In J. Cai, G. J. Stylianides, & P. A. Kenney (Eds.), *Research studies on learning and teaching of mathematics. Research in mathematics education* (pp. 219-252). Springer. https://doi.org/10.1007/978-3-031-35459-5_10

Hurrell, D. (2021). Conceptual knowledge or procedural knowledge or conceptual knowledge and procedural knowledge: Why the conjunction is important to teachers. *Australian Journal of Teacher Education, 46*(2), 57-71. https://doi.org/10.14221/ajte.2021v46n2.4

Hussein, Y. F., & Csíkos, C. (2023). The effect of teaching conceptual knowledge on students' achievement, anxiety about, and attitude toward mathematics. *Eurasia Journal of Mathematics, Science and Technology Education, 19*(2), Article 12938. https://doi.org/10.29333/ejmste/12938

Jiang, C., Kim, D. H., Wang, C., & Wang, J. (2019). Premises and challenges of high-stakes examinations: National higher education entrance examination mathematics test scores in China. *Journal of Applied Educational and Policy Research, 4*(1).

Joklitschke, J., Rott, B., & Schindler, M. (2022). Notions of creativity in mathematics education research: A systematic literature review. *International Journal of Science and Mathematics Education, 20*, 1161-1181. https://doi.org/10.1007/s10763-021-10192-z

Koretz, D., Yu, C., Mbekeani, P. P., Langi, M., Dhaliwal, T., & Braslow, D. (2016). Predicting freshman grade point average from college admissions test scores and state high school test scores. *AERA Open, 2*(4), Article 2332858416670601. https://doi.org/10.1177/2332858416670601

Krishnaveti, V., & Rawat, S. (2024). GPT takes the SAT: Tracing changes in test difficulty and math performance of students. *arXiv*. https://doi.org/10.48550/arXiv.2409.10750

Kusayanagi, C. (2013). Constructing and understanding an incident as a social problem: A case study of university entrance exam cheating in Japan. *Human Studies, 36*(1), 133-148. https://doi.org/10.1007/s10746-013-9256-2

Loveless, T. (2013). *The algebra imperative: Assessing algebra in a national and international context*. Brookings.

Moore, R., Sanchez, E., & San Pedro, S. (2019). *College entrance exams: How does test preparation affect retest scores?* ACT, Inc.

Ontong, J. M., & Bruwer, A. (2020). The use of past assessments as a deductive learning tool? Perceptions of students at a South African university. *South African Journal of Higher Education, 34*(2), 177-190. https://doi.org/10.20853/34-2-3834

Otaç, B. C. (2019). *Understanding the notion of choice in the process of university entrance exam in Turkey: A case study at METU* [Master's thesis, Middle East Technical University].

Otten, S., Araujo, Z., & Baah, F. (2024). Mathematical tasks from 141 secondary algebra lessons: A preponderance of procedures without connections. In *Proceedings of the 46th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Kent State University.

Paredes, S., Caceres, M. J., Diego-Mantecon, J., Blanco., T. F., & Chamoso, J. M. (2020). Creating realistic mathematics tasks involving authenticity, cognitive domains, and openness characteristics: A study with pre-service teachers. *Sustainability, 12*(22), Article 9656. https://doi.org/10.3390/su12229656

Parviz, M. (2023). Reflecting on the consequences of the Iranian university entrance examination: A systematic-narrative hybrid literature review. *Discover Education, 2*(22), 1-11. https://doi.org/10.1007/s44217-023-00046-x

Peterson Institute for International Economics. (2024). What is globalization? *PIIE*. https://www.piie.com/microsites/globalization/what-is-globalization

Shimizu, Y., & Vithal, R. (2023). *Mathematics curriculum reforms around the world: The 24th ICMI study*. Springer. https://doi.org/10.1007/978-3-031-13548-4

Shimizu, Y., Kaur, B., Huang, R., & Clarke, D. (2010). The role of mathematical tasks in different cultures. In Y. Shimizu, B. Kaur, R. Huang, & D. Clarke (Eds.), *Mathematical tasks in classrooms around the world* (pp. 1-14). Brill. https://doi.org/10.1163/9789460911507_002

Silver, E. A., Smith, M. S., & Nelson, B. S. (1995). The QUASAR project: Equity concerns meet mathematics education reform in the middle school. In W. G. Secada, E. Fennema, & L. B. Adajian (Eds.), *New directions for equity in mathematics education* (pp. 9-56). Cambridge University Press.

Smith, M. S., & Stein, M. K. (1998). Reflections on practice: Selecting and creating mathematical tasks: From research to practice. *Mathematics Teaching in the Middle School, 3*(5), 344–350. https://doi.org/10.5951/MTMS.3.5.0344

Watanabe, M. E. (2015). Typology of abilities tested in university entrance examinations: Comparisons of the United States, Japan, Iran and France. *Comparative Sociology, 14*(1), 79-101. https://doi.org/10.1163/15691330-12341339

Yeo, J. B. W. (2007). *Mathematical tasks: Clarification, classification and choice of suitable tasks for different types of learning and assessment*. National Institute of Education, Nanyang Technological University, Singapore.

Zhang, Y. (2016). *National college entrance exam in China: Perspectives on education quality and equity*. Springer. https://doi.org/10.1007/978-981-10-0510-7