Knowledge E
enriching | engaging | empowering

Check for updates

Research Article

Evidence and Promises of AI Predictions to Understand Student Approaches to Math Learning in Abu Dhabi K12 Public Schools

الأدلة والوعود الخاصة بتنبؤات الذكاء الاصطناعي لفهم مقاربات الطلبة لتعلم الرياضيات في المدارس الحكومية في أبوظبي

Xin Miao, Pawan Kumar Mishra, and Ali Nadaf

Alef Education, Abu Dhabi, UAE

Corresponding Author:
Xin Miao
xin@alefeducation.com

Abstract

Transforming the education system and building highly skilled human capital for a sustainable and competitive knowledge economy have been on the UAE's top policy agendas for the last decade. However, in the UAE, students' math performance on the Program for International Student Assessment (PISA) has not been promising. To improve the quality of schooling, a series of malleable predictive factors including the contributions of self-system, metacognitive skills, and instructional language skills are selected and categorized under student approaches to math learning. These factors are hypothesized as both predictors and outcomes of K12 schooling. Through the analysis using machine learning technique, XGBoost, a latent relationship between student approaches to math learning and math diagnostic test performance is uncovered and discussed for students from Grade 5 to Grade 9 in Abu Dhabi public schools. This article details how the analysis results are applied for student behavior and performance prediction, precise diagnosis, and targeted intervention design possibilities. The main purpose of this study is to diagnose challenges that hinder student math learning in Abu Dhabi public schools, uncover R&D initiatives in AI-driven prediction and EdTech interventions to bridge learning gaps, and to counsel on national education policy refinement.

الملخص

كان تحويل نظام التعليم وبناء رأس المال البشري عالي المهارة من أجل اقتصاد معرفي مستدام وتنافسي أحد أهم جداول أعمال السياسات في الإمارات العربية المتحدة على مدى العقد الماضي. ومع ذلك، لم يكن أداء الرياضيات لطلبة الإمارات العربية المتحدة في برنامج التقييم الدولي للطلبة (بيسا) واعدًا. لتحسين جودة التعليم، يتم تحديد وتصنيف سلسلة من العوامل التنبؤية المرنة بما في ذلك مساهمات النظام الذاتي والمهارات ما وراء المعرفية والمهارات اللغوية التعليمية في إطار مناهج الطلبة لتعلم الرياضيات. يتم افتراض هذه العوامل على أنها تنبؤات ونتائج التعليم المدرسي في مرحلة الصف ١٢. من خلال التحليل باستخدام تقنية التعلم الآلي، XGBoost، تم الكشف عن العلاقة الكامنة بين مناهج الطلبة لتعلم الرياضيات وأداء الاختبار التشخيصي للرياضيات للطلبة من مرحلة الصف الخامس إلى الصف التاسع في المدارس الحكومية في إمارة أبوظبي. توضح هذه الورقة كيفية تطبيق نتائج التحليل على سلوك الطالب وتوقع الأداء والتشخيص الدقيق وإمكانيات تصميم التدخل المستهدف. إن الغرض الرئيسي من هذه الدراسة هو تشخيص التحديات التي تعيق تعلم الطلبة لمهارات مادة الرياضيات في المدارس الحكومية في أبوظبي، والكشف عن مبادرات البحث والتطوير في التنبؤ المدفوع بالذكاء الاصطناعي وتدخلات تكنولوجيا التعليم لسد فجوات التعلم، وتقديم المشورة بشأن تنقيح سياسة التعليم الوطنية.

Keywords: *Human capital, Predictive learning factors, AI prediction, The instructional core, Self-system, Metacognition, Instructional language*

الكلمات المفتاحية: رأس المال البشري، عوامل التعلم التنبؤية، التنبؤ بالذكاء الاصطناعي، النواة التعليمية، النظام الذاتي، ما وراء المعرفة، اللغة التعليمية

🔓 **OPEN ACCESS**

1. Introduction

Building domestic human capital for a competitive knowledge economy has been one of the top policy agendas in the United Arab Emirates (UAE) for the past 10 years. In 2010, two milestone UAE federal policies on human capital development were initiated: first, the UAE Vision 2021 aims to build a first-rate education system with two key components: (a) the ranking of Emirati students as among the best in the world and (b) digitizing the education system by the use of smart devices and systems (UAE Vision 2021, 2009); second, the Ministry of Education 2010–2020 Strategy aims to accomplish a score of 10 out of 10 in strategic initiatives that encompass student outcomes, student school life, student equality, student citizens and administrative effectiveness to bring significant improvement in the education system (UAE Ministry of Education, 2010). In 2008, the Abu Dhabi Emirate outlined the importance of developing a highly skilled, highly productive workforce through education, training, and skill development in Abu Dhabi Economic Vision 2030 (Government of Abu Dhabi, 2008). Aligned with these policy goals, UAE has set a considerable part of its budget to develop the education sector. In 2010, the education sector received approximately 2.7 billion USD federal funds in the country (Farah, 2012); in 2018, the education sector received AED 10.4 billion from the federal budget (UAE Ministry of Finance, 2018); and in 2019, the Ministry of Education received AED 4 billion to transform the Vision 2021 into reality (UAE Ministry of Finance, 2019).

With investment and policy support, UAE has made key achievements. According to UNESCO public data on the UAE, the net enrollment rate for primary education improved from 86.53% in 2012 to 95.03% in 2017, while the net enrollment rate in secondary education reached 92.8% in 2017. Since 1975, the literacy rate among the population aged 15 years and older in the UAE has increased considerably, reaching 93.23% in 2015. Regarding the progress and completion in education, school life expectancy ISCED (the International Standard Classification of Education) 1–8 (years) for primary to tertiary education was 14.34 in 2017, the percentage of repeaters in primary education was 17% in 2015, and primary to secondary transition rate was 99.93% in 2013 (UNESCO, n.d.).

Apart from accomplishments in providing access to education and eradicating illiteracy, the UAE is moving further toward improving the quality of schooling as the country aims to have local students rank among the best in the world. In addition, since 2009, the UAE has been actively aligning the quality of education with international standards by participating in international tests such as PISA (The Programme for International Student Assessment). PISA assesses whether 15-year-olds have acquired key knowledge and skills by the time they are nearing the end of the cycle of compulsory education before greater participation in modern societies. Economists Hanushek and Kimko (2000) argue that countries with higher math and science international test scores have higher rates of economic growth. One standard deviation in test performance results in a 1% difference in annual per capita GDP growth rate (Farah, 2012, p. 3).

However, according to the 2018 PISA report on UAE students' performance, students' mean performance in reading, math, and science has been well below the OECD

average since 2012. Student mean scores in reading and science have been dropping from 2012 to 2018, and math mean score remains mostly stable, with a small fluctuation in 2015, but only over a range of <15 score points. The gap between the highest- and lowest-achieving students widened in all three subjects (OECD, 2018).

Therefore, if the UAE federal government aims to build domestic human capital and improve student performance in PISA, then understanding how its young citizens approach academic learning and the contribution of their approaches to academic performance before the age of 15 is very important. PISA data on the UAE considers both public and private schools across seven Emirates. Among the seven emirates, the Emirate of Abu Dhabi has the highest number of Emirati students in the K12 public school system (Kippels & Ridge, 2019). This article focuses on understanding how Grade 5 to Grade 9 students approach math learning in Abu Dhabi public schools to find out potential reasons behind why 15-year-olds perform consistently low in the PISA math test.

Student approaches to learning are reflected in how the student approaches the instructional task. According to the "Instructional Core" proposed by City et al. (2009), the instructional task that the student is actually doing predicts performance, and "increases in student learning occurs only as a consequence of improvements in the level of content, teachers' knowledge and skill, and student engagement" (p. 23). The student approaches an instructional task by first directing the information processing inward to have domain-specific judgments about self and task (for instance, self-efficacy, affective reactions, motivation) and then outward at the required learning activities using metacognitive and cognitive strategies, according to the preliminary causal model of the role of the self-system in self-regulated learning (McCombs, 1986). Winne (1997) proposed the COPES model for self-regulated learning that incorporates conditions, operations, products, standards, and evaluation. Conditions consist of internal conditions including the following: knowledge about the topic, study tactics, motivational orientation, and external conditions such as difficulty of the level of learning content that the student perceives could influence internal conditions. Operations work on information following cognition and metacognition processes. Every operation generates products that are evaluated using standards. Basic cognitive operations consist of the following: (a) searching for information that meets standards; (b) monitoring by identifying whether information corresponds to standards; (c) assembling by joining previously separate information to identify a relationship; and (d) rehearsing by reinstating information in working memory and transforming the representation of given information (Winne, 1985; Winne, 2010).

Metacognition is a quality of thoughts and thinking, using the same fundamental cognitive operation processes. Throughout the self-regulated learning process, as Efklides et al. (2018) pointed out, the student's motivation and emotions play a key role in the following three aspects: (1) motivation and emotions are internal conditions that the student monitors in the initial phase of self-regulated work; (2) the standard that the student sets in metacognitive monitoring corresponds to the presence or level of motivation and emotions; (3) the student regulates cognition the same way as the student sets goals to regulate motivation and emotions.

Assuming that the student is the active self-regulatory learning agent while the role of the teacher is not dominant, student approaches to math learning within the Instructional Core essentially follow Winne's COPES (1997) processes including conditions, operations, products, standards, and evaluation as well as McComb (1986)'s model, both of which encompass internal conditions that cover self-system factors such as affective emotions, motivation, and metacognitive and cognitive operations. In addition, math instructional language, English, is included as an external condition because Emirati students learn math in their second language. Also, grit, coined by Duckworth et al. (2007), is added as a self-system factor/an internal condition because it reflects a key quality if an individual can persevere and have passion for long-term goals to become a high-skilled citizen that contributes to key human capital for a country like the UAE.

Although the data in this paper were collected from students, they are reflective of the interplay of the Instructional Core, as explained above. Due to the complexity of the contributions of learning factors in the education domain, instead of using traditional social science statistical analysis, this study uses machine learning techniques because the prediction model takes into account the contribution of each feature/learning factor in a complex way. This fits the nature and complexities of education, plus using machine learning and deep learning techniques to uncover the contribution of learning factors to student academic performance has been practiced by researchers for the past 10 years. Pardos et al. (2014) used Machine Learning algorithms (e.g., step regression, Naive Bayes, etc.) to study students' behavior and affective states using log data from a web-based learning platform and their effects on the end-of-year math performance. They concluded that confusion and boredom are negatively correlated with the learning outcomes, while engaged concentration and frustration have a positive association with the learning outcomes. Kumari et al. (2018) applied bagging, boosting, and voting ensemble methods on Decision Tree (ID3), Naive Bayes, K-Nearest Neighbor, Support Vector Machines to predict the student academic performance using e-learning system data. They showed that behavioral features are influential factors affecting the performance of students. Similarly, Kostyuk et al. (2018) used multiple machine learning algorithms like linear models, random forest, and gradient boosting to investigate several affective states and engagement-related behaviors in a mathematics-blended learning environment, and found that students with high rates of engaged concentration were more likely to perform well in the test. Harvey and Kumar (2019) used linear regression, decision tree, and Naive Bayes classifiers to develop predictive models to predict math scores for high school students. According to their results, the Naive Bayes technique produced the highest accuracy. Besides, Crossley et al. (2019) conducted a study to explain the relationship between math success and language production by applying a mixed-effect model and NLP techniques in an online tutoring system. They concluded that students who were successful in math produced sophisticated language. Coleman et al. (2019) proposed an approach to circumvent the cold-start problems associated with students' at-risk status using an ensemble model. They showed that the approach produces effective predictive power to address cold-start scenarios.

The current study proceeds in the following manner. We begin with the research framework highlighting hypotheses and research questions followed by reviewing

XGBoost and SHAP methods to analyze the contributions of student approaches to math learning towards their math academic performance . Next, we summarize major findings and suggest education policy implications and R&D directions for the Ed-tech sector in the UAE.

2. Research Framework

2.1. Hypotheses

We propose the following hypotheses to foster our research on the issues discussed:

1. Student approaches to learning are considered as both predictors and outcomes of schooling. Outcomes of schooling should consist not just of academic performance measured by standardized test scores, but also how students approach academic learning, which is the "why" behind what the test score says. We hypothesize that student approaches to learning predict academic performance.

2. Student approaches to math learning within the Instructional Core consist of the interplay of the self-system factors, metacognitive strategies, math cognitive strategies, and instructional language skills to execute math learning tasks and produce learning outcomes, given that the student plays a central self-regulatory role.

2.2. Research Questions

1. What is Grade 5 to Grade 9 students' math academic performance level compared with their international peers who are using the same American Common Core curriculum standards?

2. What are the predictive learning factors that measure student approaches to math learning hypothesized above?

3. Is there a relationship between these factors and students' math academic performance?

4. If yes, what is the relationship? Does the relationship vary between students?

3. Methodology

3.1. Qualitative and quantitative measurements

This section explains the measurements of the hypothesized student approaches to math learning which are highlighted in Table 1. Since this study focuses more on general metacognitive functioning, self-system, and the role of instructional language when the student acts as the self-regulatory agent within the Instructional Core, math content-specific cognitive strategies are not included in this article. Metacognitive skills and self-regulated learning strategies are proven to drive effective independent learning (Zimmerman and Schunk 2001, as cite in OECD 2010, p.42). Moreover, designing a battery of qualitative measurements for math content-specific cognitive strategies in the

UAE context requires extensive research itself. However, the lack of cognitive strategies is acknowledged as a limitation.

Table 1

Predictive factors for student approaches to math learning

Student approaches to math learning	Categories	Predictive factors
Self-system	Domain-specific expectations and judgments about self and task	Math self-efficacy
		Math self-concept
	Affective reactions (key anxiety factors)	Math in-class anxiety
		Math grade anxiety
	Motivation	Math intrinsic motivation
		Math instrumental motivation
	Grit	Grit
Metacognitive regulatory strategies	MARSI (Metacognitive Awareness of Reading Strategies Inventory)	Global reading strategies
		Problem-solving strategies
		Support reading strategies
	PISA Math Learning strategies	Memorization strategies
		Control strategies
		Elaboration strategies
Instructional Language for math	English	Language
		Reading

The definition of the self-system factors is based on understanding the preliminary causal model of the role of the self-system in self-regulated learning (McCombs, 1986). According to research on the causal relations between self-system variables, affect, motivation, and actual performance, student affective reactions of competency are most strongly related to motivation to perform which strongly predicts performance (Harter 1985, as cited in McCombs, 1986, p.9). Once students develop the motivation to perform, the process of activating general metacognitive skills and domain-specific cognitive strategies would start. All survey questions were delivered in both English and Arabic to ensure students understood the questions. Before the implementation of surveys, three rounds of bilingual survey design evaluations were conducted by Arabic and English specialists, considering our research understanding of students' literacy levels in Abu Dhabi public schools.

Among the self-system factors, self-efficacy and self-concept are designed to reflect student expectations and judgments about self and math-learning tasks. Self-efficacy is defined by items such as "I believe I can understand the content in my math lesson" and self-concept by items such as "I learn math quickly." Moreover, five questions were designed to measure self-efficacy in math and five for math self-concept, using scale items such as "very different from me, a little different from me, a little like me, a lot like

me." Among the various affective reactions to math learning, two key anxiety factors, in-class anxiety and grade anxiety, were designed due to prior field research experience and future intervention considerations to improve classroom learning dynamics. Five questions were designed to measure in-class anxiety such as "I am afraid to give wrong answers during my math class," three questions for grade anxiety such as "I worry that I will get poor grades in math tests," using scale items "very different from me, a little different from me, a little like me, a lot like me." Intrinsic motivation defined by items such as "I am interested in the things I learn in math" and instrumental/extrinsic motivation by items such as "Math is an important subject for me because I need it for what I want to study later on" are used as sources of motivation toward math learning. In addition, four questions were designed to measure intrinsic motivation and four to measure instrumental/extrinsic motivation, using scale items "very different from me, a little different from me, a little like me, a lot like me." Among these factors, self-efficacy is the foundation and key because self-efficacy beliefs reflect the first action that the self-system judges about self and the learning task, and then produce diverse effects through four major processes, including affective responses, motivation, metacognitive, and cognitive processes (Bandura, 1997). The surveys were designed based on the understanding of relevant research literature, PISA survey design framework, field research experience, and future intervention considerations. Simple phrasing with a scale of 1–4 was chosen to reduce complexities that students may have in interpreting questionnaires. The weighted averages of survey responses were used for statistical analysis.

Eight original grit survey statements such as "Setbacks do not discourage me. I bounce back from disappointments faster than most people" were applied using a Likert scale of 1–5, with scale items such as "not like me at all, not much like me, somewhat like me, mostly like me, very much like me" (Duckworth et al., 2007). The weighted average of survey responses was used for statistical analysis.

For metacognitive strategies, two batteries were applied. First, Metacognitive Awareness of Reading Strategies Inventory (MARSI) was used to assess students' metacognitive awareness and perceived use of ESL reading strategies while reading school-related materials. Fifteen questions were categorized under global reading, problem-solving, and support reading strategies, using a Likert scale of 1–5. A student's response >3.5, between 2.5 and 3.4, and <2.5 indicates high, medium, and low levels of awareness, respectively (Mokhtari et al., 2018). The primary reason for choosing this tool was that students in Abu Dhabi public school system learn math in their second language, that is, English, and metacognitive processing of English comprehension is the first step to process math concepts. Also, PISA math learning strategies such as memorization, control, and elaboration strategies were applied to measure student metacognitive operations involved in math learning. Memorization strategies involve learning facts or rehearsing examples. These strategies are needed when learners retrieve information for further processing. Elaboration strategies indicate that students relate the understanding of new knowledge to prior learning and knowledge, which deepens students' understanding. Control strategies mean that learners set goals and monitor progress in reaching learning goals. Likert scale of 1–5 was used to understand the frequency

of exercising the above three strategies, such as "never or almost never, occasionally, sometimes (about 50% of the time), usually, always or almost always" (OECD, 2010). The weighted averages of survey responses were used for statistical analysis.

For the instructional language measurement, the MAP Growth English adaptive test[1] was selected to measure student English performance which consists of language and reading. Language assesses vocabulary and grammar skills while reading assesses reading comprehension only.

To understand the math academic performance of Grade 5 to Grade 9 students in Abu Dhabi public schools, the Renaissance Star Math test[2] was implemented. The Star Math test is designed based on American Common Core standards and more than one-third of US schools use this solution, which gives us a good benchmark to compare Abu Dhabi public school students against their peers who follow the same curriculum standards. The Star Math test provides two different types of test scores: criterion-referenced scores and norm-referenced scores. The criterion-referenced scores measure students' knowledge level and their abilities, such as Scaled Score (SS) and Grade Equivalent (GE) score. Norm-referenced scores provide a relative measure of student achievement compared to the results of other students who have taken the same test as Percentile Rank (PR) and Normal Curve Equivalent (NCE). In this study, we use the criterion-referenced scores: SS and GE.

3.2. Sampling and data collection

Purposive sampling, weighted random sampling, and demographic considerations such as gender and school location are applied to create a sample of 4,107 students from Grade 5 to Grade 9 across 20 Abu Dhabi public schools. These sampling approaches were used to select representative samples for all Abu Dhabi public schools. Of note, Abu Dhabi public schools use the Alef platform as a primary education tool and students are assessed using formative assessments. To ensure that our sample data is a valid representation of the student population in Abu Dhabi public schools, we have used the Kolmogorov–Smirnov two-sample statistical test on students' math formative assessment scores on the Alef platform. The test results in a *p*-value of 0.342, indicating that the sampling distribution follows the real student population distribution.

Moreover, prior to the implementation of questionnaires and diagnostic tests, approvals from the Department of Education and Knowledge (ADEK) and school principals were obtained. Students and teachers were introduced to the diagnostic tests and questionnaires with an explanation of purposes. All data collected follow the government student privacy protocol and do not contain any personally identifiable information.

While the MAP Growth English diagnostic test and the Renaissance Star Math diagnostic test were implemented through online adaptive test systems, the bilingual questionnaires in English and Arabic were implemented through Survey Monkey. The Diagnostic tests and questionnaires were implemented in the same environment and at the same time in November 2019. Measures are taken to ensure the quality of data collection. For instance, teachers and proctors were trained to ensure that students

record the required information properly such as their student ID, avoid cheating, and spend at least 20 minutes on the diagnostic test that covered 30 questions.

3.3. Dataset description

The collected datasets were combined and analyzed for descriptive and exploratory insights. Of the 4,107 students, high-quality data of 1,660 students were obtained post data merging and cleaning. Additionally, data of students who did not record the correct information such as student ID on both questionnaires and diagnostic tests and those who spent <20 minutes as the minimally required time for the diagnostic test were removed. Table 2 shows the distribution of students per grade level.

Table 2

Number of students in each grade level

Grade level	No. of students
5	305
6	226
7	451
8	515
9	163

Further, Cronbach's alpha with item analysis was used to measure the reliability of all survey questions, which yielded excellent internal consistency with a Cronbach alpha of 0.9074. The reliability for self-system and metacognitive regulatory strategies questions were 0.8178 and 0.9217, respectively. Table 3 summarizes all the variables used for this research. Star Math SS ranged from 0 to 1400 and was used as a target value in this study to compare student math performance across grade levels. GE is a norm-referenced score ranging from 0.0 to 12.9+. The weighted average responses for student questionnaires were used for analysis.

3.4. Learning algorithm

Machine learning is used to solve and unfurl complex problems in much of the research today. It is utilized to find valuable patterns within data and turn information into knowledge. These analyzed patterns and knowledge are applied to build the foundation for further prediction.

Challenges in the education domain are complex because there is an interplay of multiple factors from many stakeholders including the student, teacher, and content factors and family. This study only focuses on understanding the contribution of measurable and malleable factors to student math learning within the Instructional Core to identify key issues. To solve this problem, we used an extreme gradient boosting model (XGBoost) technique. XGBoost is a scalable implementation of gradient boosted decision trees (Chen & Guestrin, 2016). This algorithm is immensely popular among data scientists due to its high accuracy and low risk of overfitting. XGBoost is based

Table 3

List of features

Feature	Data type	Description	Range | Values
Language	Numerical	English diagnostic test result	[100, 260]
Reading	Numerical	English diagnostic test result	[100, 260]
SS (Scaled Score)	Numerical	Math diagnostic test results (Target Feature)	[0, 1400]
GE (Grade Equivalent)	Numerical	Math diagnostic test results	[0.0, 12.9+]
Self-concept in math, Self-efficacy in math, Intrinsic motivation, Instrumental motivation, In-class anxiety, Grade anxiety	Numerical	Self-system	[1, 4]
Control strategies, Memorization strategies, Elaboration strategies, Global reading strategies, Support reading strategies, Problem-solving strategies	Numerical	Metacognitive regulatory strategies	[1, 5]
Gender	Categorical (Nominal)	Demographic	Boy and girl
Grade	Categorical (Ordinal)	Grade level of students	[5, 6, 7, 8, 9]

on boosting techniques which iteratively updates the parameters to create a strong predictor (Friedman, 2001).

Boosting techniques are complex ensemble algorithms that provide insights into the system from a global perspective based on features' importance. However, it fails to provide interpretation at individual predictions. To unfold these complex algorithms, we applied Explainable Artificial Intelligence (XAI) methods. These methods have been designed to provide visibility to understand predictions from both global and local level interpretation. One of the methods that comes from cooperative game theory is known as SHAP (Shapley Additive exPlanations).

3.5. Interpretation of complex algorithms (SHAP)

SHAP is a game-theory and a local explanation technique that is used to interpret the contribution of each feature (predictive factor) on the output of the model (Lundberg & Lee, 2017). Shapley values are utilized to compute the significance of a feature by contrasting what a model predicts with and without a feature from every conceivable blend of n features in the dataset S. Given a value of feature $x \in S$, SHAP calculates the prediction p of the model with x. The Shapley value φ is calculated as follows (Shapley, 1953):

$$\phi(p) = \sum_{S \subseteq N/x} \frac{|S|\,!\,(n-|S|-1)!}{n!}(p(S \cup x) - p(S))$$

However, the order of features is relevant and can affect the predictions, so the model does all possible permutations to evaluate features decently. It suppresses a feature of no contribution to zero and provides equal values to features that have the same contribution (Lundberg & Lee, 2017).

3.6. Experimental evaluation

We conducted the XGBoost model training by holding 10% of the data in the test set and used the remaining ones in the training set to predict SS as a target value using the predictive features in Table 3. To evaluate our experiments, we considered two widely popular statistical metrics, Root Mean Square Error (RMSE) and Coefficient of Determination (R^2). RMSE is a measure of the deviation of predicted values from their true ones, while R^2 values represent the magnitude of the relationship between the predicted values and the observed outcomes. We implemented the proposed method in Python with the XGBoost[3] package and SHAP[4] library and executed them on a pc with 2.6 GHz intel@core i7, 16GB DDR4 memory.

After tuning the hyperparameters and employing Monte Carlo simulations with 100 iterations, the final XGBoost model processed the test dataset with R^2 of 0.67 and $RMSE$ of 80.6. Mean and Standard Deviation of the difference between actual and predicted values were –0.87 and 80.85, respectively. We then applied SHAP values to interpret our complex XGBoost model.

4. Results and Discussions

4.1. Student math academic performance compared with international peers

As mentioned in the introduction section, the UAE 15-year-old-student PISA math mean scores have been well below OECD since 2012. Moreover, there has not been much change in the math mean score from 2012 to 2018, and the performance gap between high and low achievers has been widening. This implies that students experience learning gaps before reaching the age of 15. Not surprisingly, as shown in Figure 1, students from Grade 5 to Grade 9 in Abu Dhabi public schools do have a grade gap. In this article, we define a parameter, grade gap, representing how far the students are below the grade level. This parameter is achieved as follows:

Grade gap = GE (Grade Equivalent) score – Grade level

The average math grade gap per each grade level falls below zero from Grade 5 to Grade 9, which indicates that on average, students in Abu Dhabi public schools lag behind international peers from Grade 5 to Grade 9.

Further, the average grade gap is observed with a growing trend moving from Grade 5 to Grade 9, despite a small deviation in Grade 7, reaching 3.366 in Grade 9; thus, on

Figure 1

Average student grade gap per grade level
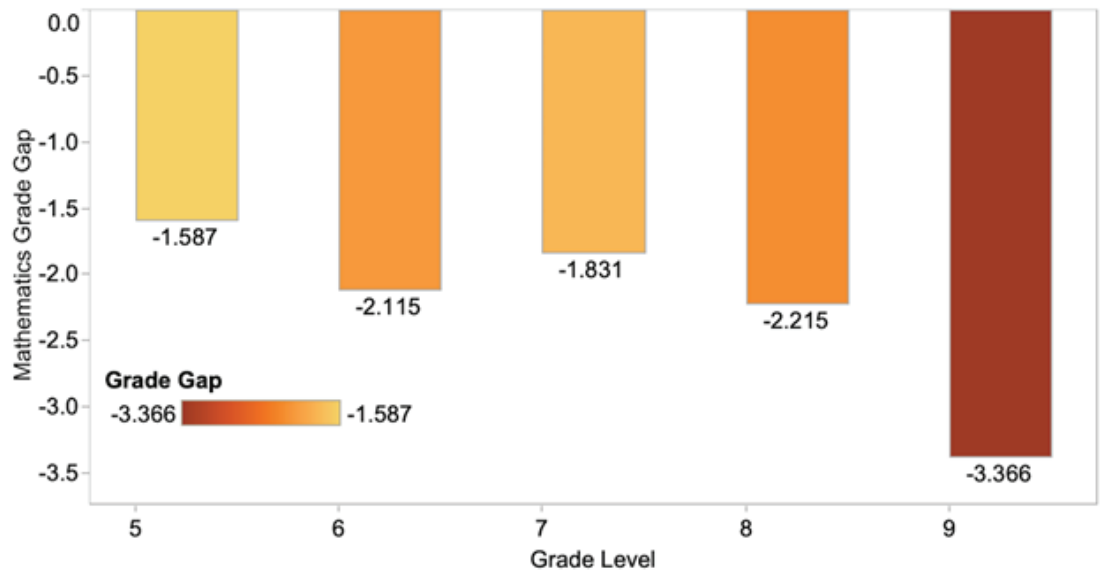
average, the actual math performance level of an average Grade 9 student equals that of a fifth grader reaching the seventh month of the academic year. Possible explanations for the increase in the average grade gap in Grade 9 might be that our sample size for Grade 9 is relatively smaller than that of other grade levels.

The following section focuses on uncovering evidence of the contribution of predictive factors measuring student approaches to math learning to math academic performance.

4.2. Relationship between predictive factors for student approaches to math learning and math academic performance

Figure 2 shows the contribution of predictive factors to an individual student's math Scaled Score (SS = 609.66) which is the output value. Predictive factors that push math SS scores to higher and lower values from the baseline are shown in red and blue, respectively. The length represents the magnitude of the contributions. For instance, language score 171 is the number one factor that negatively contributes to math academic learning for this student, whereas being in Grade 8 is the number one factor that positively contributes to math academic learning for this student. As seen in Figure 2, SHAP provides an interpretation of predictive factor contribution at each student level. If we have decent quality predictive datasets per student, personalized diagnosis is within reach, which is immensely powerful in designing creative, yet, targeted personalized education interventions.

Figure 3 shows a SHAP summary plot aggregating the importance and effects of predictive factors for all students. Each dot on the summary plot represents the SHAP

Figure 2

SHAP predictions for an individual student's math academic performance
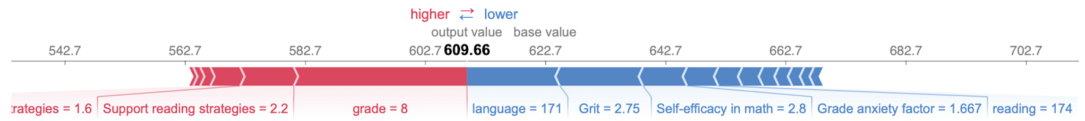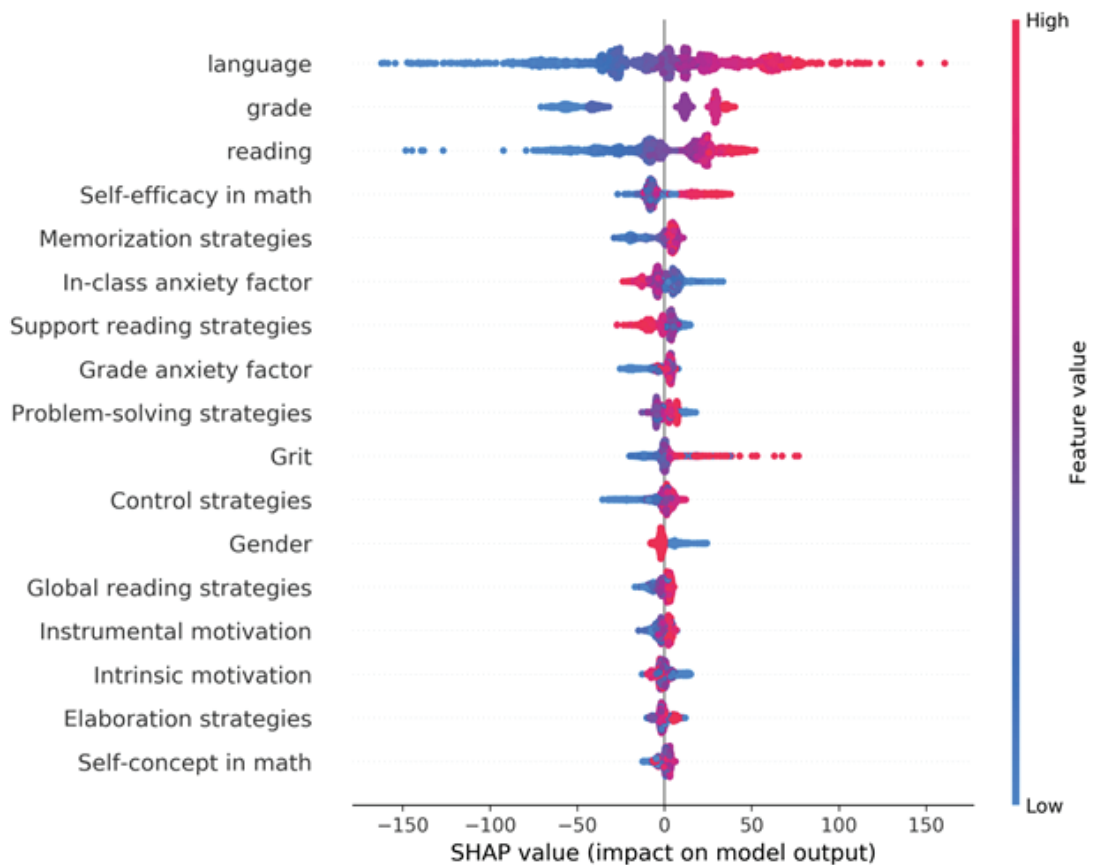

Figure 3

*SHAP summary plot of predictive factors by decreasing importance*

value of a predictive factor for a student. Among the predictive factors, language, grade, reading, self-efficacy in math rank the top influential factors that contribute to student math learning. Grade referring to grade level ranks high in importance because math scaled score is the output value, and it is expected that students from higher grade levels tend to have higher scaled scores, even though the grade is not a predictive factor defined in the previous section.

According to Figure 3, higher values of language and reading (i.e., predictors of instructional language for math) result in higher SHAP values delineating a higher

probability to get higher scores in math. Negative SHAP values for language indicate a negative contribution of language to math scores. Furthermore, we can also see that language SHAP value distribution is very wide, which means that the contribution of language to math learning can be extremely positive or extremely negative. In other words, when a student has extremely low language skills, potentially, math performance of this student is negatively affected to a great extent. Given that language is of the highest importance for math learning, segmentation analysis is carried out to find out the extent to which the contribution of language skills to math academic performance varies from student to student to get a more precise idea to design proper language intervention programs for different students. Figure 4 shows the average contribution of predictive features in terms of SHAP values.

Figure 4

*SHAP feature importance plot of predictive factors*
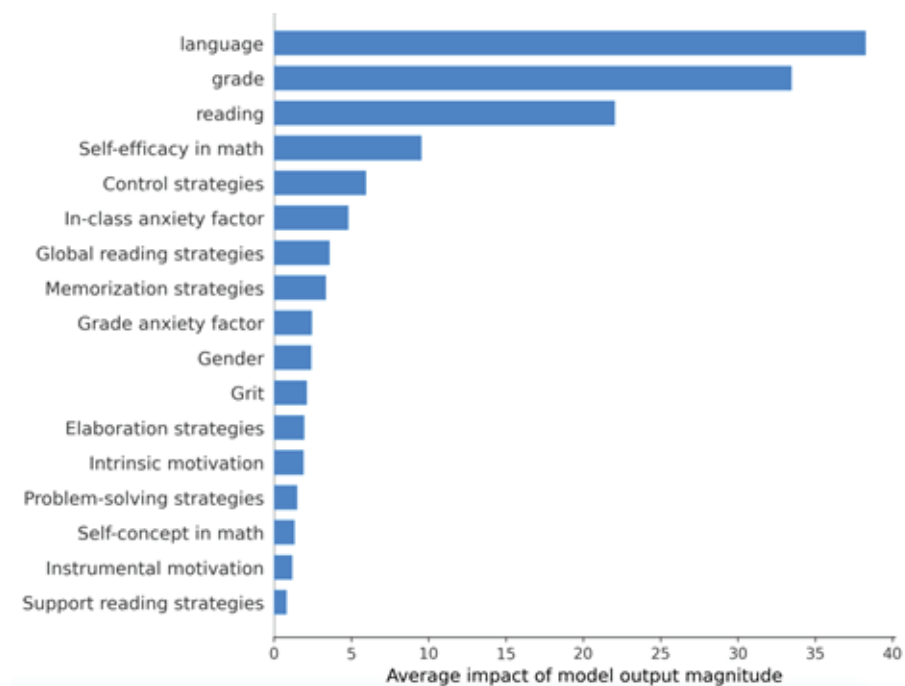
Students were segmented based on the following evidence: if the math grade gap (GE score – grade level) for a student was $\leq$–3, the student was considered very poor; if it was >–3 and $\leq$–1, the student was considered poor; if >–1 and $\leq$1, the student was considered normal; and if the math grade gap was >1, the student was considered outstanding. The segmentation evidence for very poor, poor, normal, and outstanding student groups were applied across Grade 5 to Grade 9. In the following section, the average SHAP values of language per student segmentation/category are analyzed. Note that average SHAP values show whether features are positively or negatively associated with math performance.

As shown in Figure 5, language contributes negatively to math performance for very poor students (i.e., average SHAP value is –22.96), and the average language score for very poor students is 171.86. Poor students stay on the borderline with a

mean score of 180.22, indicating that if the language score decreases from 180.22, math performance will be affected negatively. For normal and outstanding students, language contributes positively to math performance, with average SHAP values of 25.29 and 60.85, respectively. These lead to two main takeaways: (1) when a student scores below 171.86 in language, language starts to contribute negatively to math performance regardless of grade levels and (2) very poor student segmentation group is the intervention focus group whose language is the main factor that contributes negatively to math performance.

Figure 5

Average SHAP values of each student category and language score
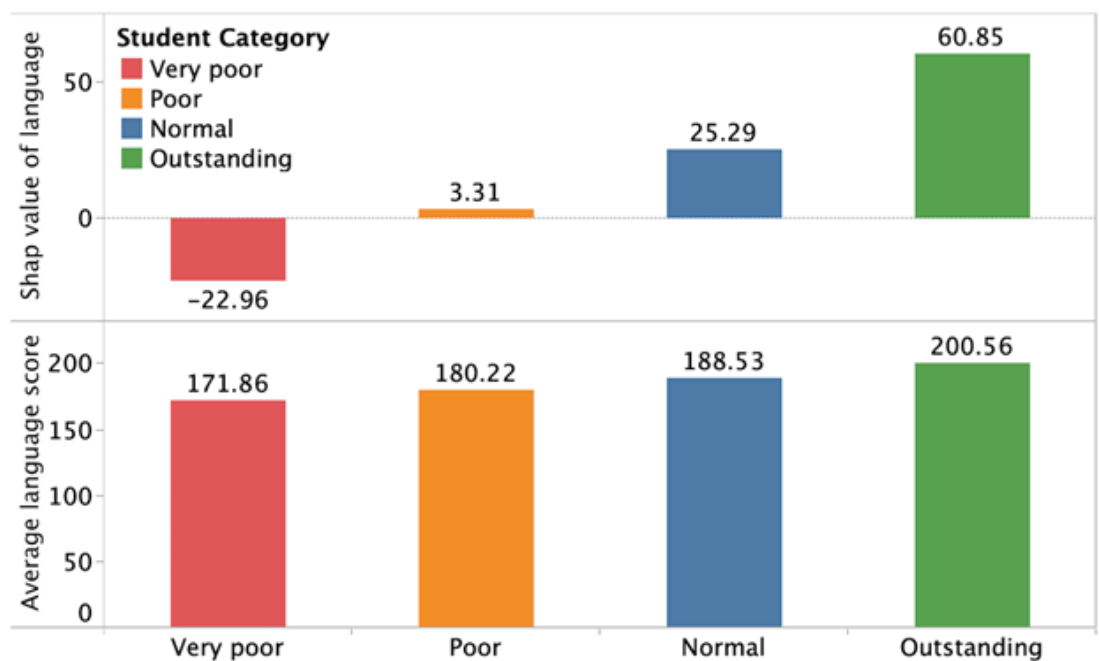
Using the same student segmentation standard above, the following is the relationship between all selected predictive factors and math performance per segmentation group.

As shown in Table 4, moving from very poor to poor, to normal, and to outstanding math performance segmentation groups, the average SHAP values that reflect the relationship between predictive factors and math performance increase from negative to positive, representing a relationship between these factors and student math academic performance. For instance, outstanding students have better language and reading skills, hence better math performance, whereas very poor students do not perform well in math, mainly due to the lack of proficient language and reading skills.

The following sections delve further into a detailed analysis of the relationship between math academic performance and instructional language, self-system, metacognitive regulatory strategies, respectively, with segmentation evidence.

Table 4

Average SHAP value of features per student category

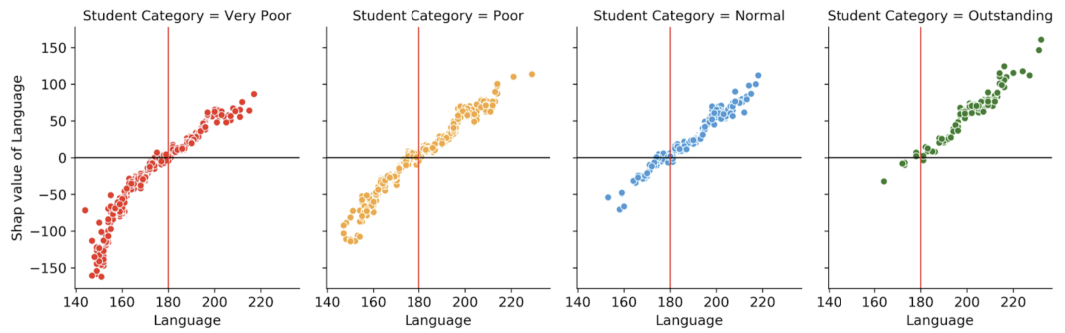| Features	Very poor	Poor	Normal	Outstanding
Avg. SHAP Language	−22.96	3.31	25.29	60.85
Avg. SHAP Reading	−10.08	−1.02	12.88	24.52
Avg. SHAP Control strategies	−1.50	0.03	1.31	2.35
Avg. SHAP Elaboration strategies	−0.54	0.27	0.83	1.08
Avg. SHAP Global reading strategies	−1.03	0.08	1.10	1.58
Avg. SHAP Grade anxiety factor	−0.57	0.07	0.54	0.41
Avg. SHAP Grit	−0.73	−0.37	0.92	3.55
Avg. SHAP In-class anxiety factor	−0.88	−0.09	0.95	2.53
Avg. SHAP Instrumental motivation	−1.37	−0.41	0.50	1.14
Avg. SHAP Intrinsic motivation	−1.28	−0.88	−0.35	0.01
Avg. SHAP Memorization strategies	−0.99	0.45	1.65	2.23
Avg. SHAP Problem-solving strategies	−0.37	0.18	0.44	0.66
Avg. SHAP Self-concept in math	−0.05	0.43	1.03	1.85
Avg. SHAP Self-efficacy in math	−2.47	0.02	3.97	7.55
Avg. SHAP Support reading strategies	0.17	0.42	0.54	0.79

4.3. Relationship between instructional language and math performance segmentation groups

As shown in Figure 6, when students score >180 (±1.94) in language, their language skills start making a positive contribution to math academic performance, given other predictive factors, regardless of student segmentation groups. However, the number of students whose language scores fall <180 increases from outstanding to normal, to poor, and to very poor categories. A language score of 180 was considered as the cut-off value, from which the SHAP value is either positive or negative. The vertical red lines in Figures 6 and 7 show the cut-off values separating positive and negative SHAP values.

In terms of the contribution of reading factor, as shown in Figure 7, when students score >170 (±1.29) in reading, their reading skills start making a positive contribution to math academic performance, given other predictive factors, regardless of student segmentation groups. A reading score of 170 was considered as the cut-off value, from
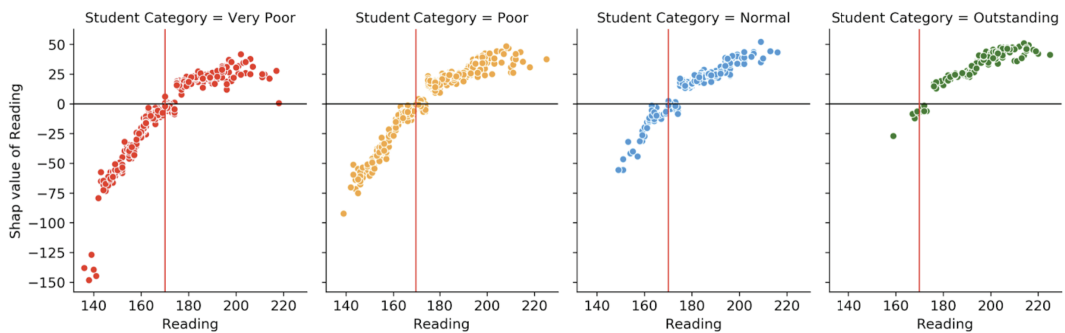
Figure 6

Scatter plot between language score and SHAP value of language feature for each student category

which the number of students whose reading scores fall below the value increases from outstanding to normal, to poor, and to very poor categories.

Figure 7

Scatter plot between reading score and SHAP value of reading feature for each student category

Further, Table 5 focuses on the relationship between language and reading per math performance segmentation category.

Table 5

Pairwise correlations between language and reading in each student category

Student categories	Variables	Language
Very poor	Reading	0.334
Poor	Reading	0.562
Normal	Reading	0.518
Outstanding	Reading	0.566

As Table 5 illustrates, there is a moderately strong positive correlation between language and reading per outstanding, normal, and poor student segmentation groups. The correlation between language and reading is moderate (0.334) per very poor student group.

Based on Figures 6 and 7, and Table 5, determining the language and reading skills evaluated behind the cut-off values (for language, 180, and for reading, 170) is very important to design targeted second foreign language learning intervention programs for those who are struggling with math because of their poor English language and reading skills. Learning grammar and vocabulary is a starting point to help very poor math learners to read in English so they are able to better process math content comprehension.

4.4. Relationship between self-efficacy and math performance segmentation groups

Figure 8 shows that the contribution of self-efficacy to math academic performance starts to be positive for those who responded >3.5 in math self-efficacy. This is compelling evidence that regardless of math performance level, when any student has higher expectations and judgements about self and math-learning tasks, math academic performance benefits.

Figure 8

Scatter plot between self-efficacy in math and SHAP value of self-efficacy in math feature for each student category
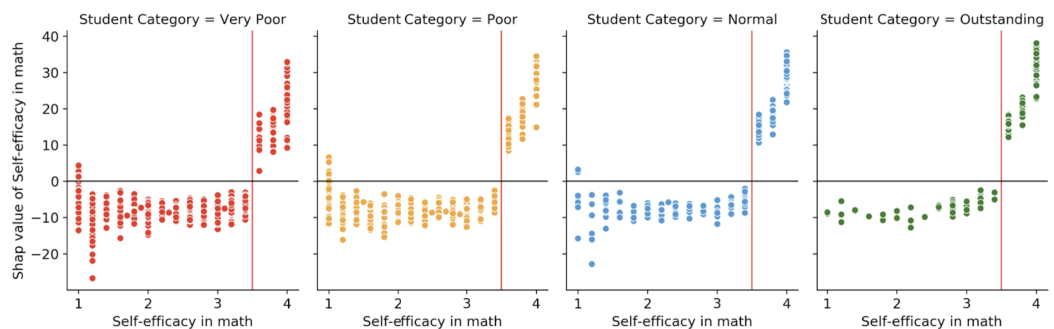

Furthermore, when a student has higher expectations and judgements about self and math learning tasks, it is also expected that self-efficacy beliefs motivate students to behave differently. The correlation between self-efficacy and intrinsic motivation is moderately strong for poor, normal, and outstanding student segmentation groups, with correlation coefficient values of 0.637, 0.667, and 0.60, respectively. For very poor students, the correlation between self-efficacy and intrinsic motivation is moderate ($r =$ 0.462). A strong positive correlation exists between self-efficacy and self-concept per very poor, poor, normal, and outstanding students, with correlation coefficient values of 0.645, 0.709, 0.714, and 0.719, respectively. Among the self-system factors, self-efficacy in math acts as the base filter when a student interacts with math-learning tasks. To boost student intrinsic motivation to learn math, the starting point is to build stronger self-efficacy in math.

4.5. Relationship between MARSI and math performance segmentation groups

As Figure 9 demonstrates, any student, regardless of being outstanding, normal, poor, or very poor, when they respond about 3.7 or above in global reading strategies (i.e., high-level awareness), their metacognitive awareness in using ESL global reading strategies contributes positively to math academic performance. However, the number of students who do not have a high level of awareness in using ESL global reading strategies is increasing from outstanding to normal, to poor, and to very poor categories.

Figure 9

Scatter plot between global reading strategies and SHAP value of global reading strategies feature for each student category
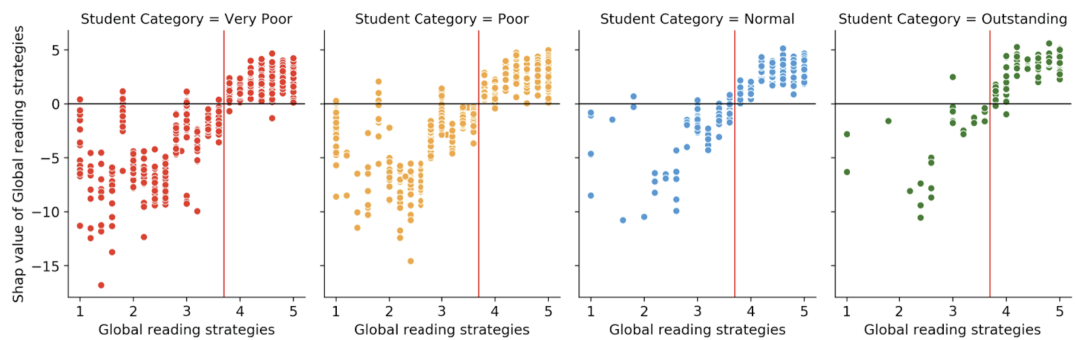
The same findings were found for problem-solving strategies under the MARSI framework, with a cut-off point of 4.1. Thus, having a high level of metacognitive awareness in using global reading strategies and problem-solving strategies benefits math academic performance for all learners.

4.6. Relationship between PISA math regulatory learning strategies and math performance segmentation groups

The cut-off point for math control strategies (i.e., usually or always) is 3.7, beyond which the control strategies contribute positively to student's math performance (see Figure 10). In addition, the number of students who exercise control strategies not frequently (i.e., sometimes, never, or occasionally) is increasing from outstanding to normal, to poor, and to very poor categories. The same findings were found for elaboration and memorization strategies with a cut-off point of 4.9 and 3.0, respectively.

Based on the above findings for control, memorization, and elaboration strategies, exercising PISA math metacognitive regulatory strategies more frequently (usually or always) benefits learners regardless of their math performance levels.

Figure 10

Scatter plot between control strategies and SHAP value of control strategies feature for each student category
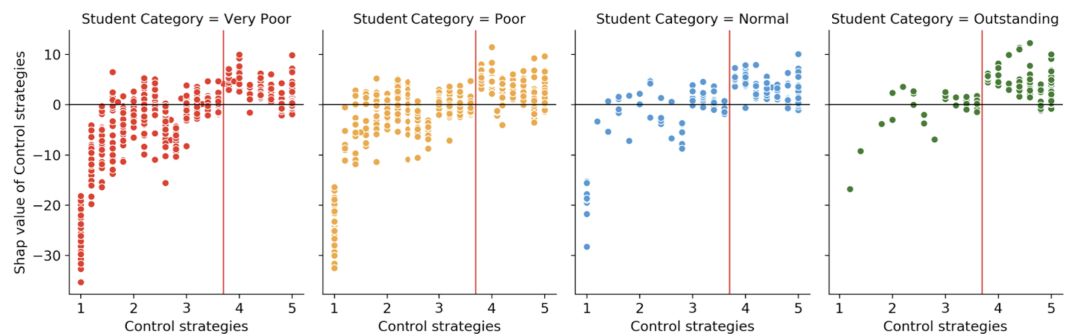
4.7. Predictive factors for very poor students

As shown in Table 6, the importance of predictive factors is ranked based on average SHAP values. Almost all predictive factors except for support reading strategies negatively contribute to math academic performance for very poor students. Amongst these factors, instructional language factors such as language and reading levels have the strongest negative contribution to math academic performance. It is highly likely that because very poor students struggle with English language and reading that could result in low self-efficacy in judgments about self and math learning tasks in English. As a result, students might have a lack of intrinsic motivation to learn math. In addition, based on the evidence from the previous sections, many very poor students do not practice PISA math learning strategies frequently and they do not have high-level awareness of using global reading and problem-solving strategies under MARSI, which worsens the learning experience and outcomes for them. In short, the very poor student segmentation group needs urgent and targeted intervention from all ends, especially from language and reading intervention programs, plus metacognitive skill training.

5. Conclusions, Limitations, and Future Steps

This article presents the results of a large-scale investigation of how Grade 5 to Grade 9 students approach math learning and how their approaches contribute to math academic performance in Abu Dhabi public school system. The main purpose of the study was to use research to uncover evidence-based findings to inform research and policy development, research and development initiatives in the EdTech sector to improve student learning experience and outcomes with systematic data and a rigorous research methodology. Although the analysis of additional variables can further improve this study, the findings suggest a number of solid possibilities for policy refinement, future research, and R&D initiatives in the EdTech sector in the UAE.

First, the results show that student math academic performance benefits when students: (a) have higher self-efficacy in math; (b) have a high level of awareness in using

Table 6

SHAP value of features for very poor students

Predictive factor	Avg. SHAP value
Avg. SHAP Language	−22.96
Avg. SHAP Reading	−10.08
Avg. SHAP self-efficacy in math	−2.47
Avg. SHAP control strategies	−1.50
Avg. SHAP instrumental motivation	−1.37
Avg. SHAP intrinsic motivation	−1.28
Avg. SHAP global reading strategies	−1.03
Avg. SHAP memorization strategies	−0.99
Avg. SHAP in-class anxiety factor	−0.88
Avg. SHAP grit	−0.73
Avg. SHAP grade anxiety factor	−0.57
Avg. SHAP elaboration strategies	−0.54
Avg. SHAP problem-solving strategies	−0.37
Avg. SHAP self-concept in math	−0.05
Avg. SHAP support reading strategies	0.16

global reading strategies and problem-solving strategies under MARSI; (c) exercise PISA math metacognitive strategies on a more frequent basis such as usually or always; and (d) score >170 in language or 180 in reading. These are solid pieces of evidence that our hypothesized model of student approaches to math learning, if encouraged and executed in the right way within the Instructional Core, improves student academic performance in math. For students to be engaged and motivated for math-learning tasks, the right level of math content has to be provided and students have to be equipped with the right level of instructional language skill sets (i.e., language, reading). In the meantime, when students interact with learning tasks, metacognitive knowledge and metacognitive regulatory strategies must be taught and encouraged by the teacher to be executed by the student. As principle 2 of the Instructional Core says, "If you change any element of the instructional core, you have to change the other two" (City et al., 2009). Student engagement, the right math content delivery, and teacher facilitation in how students exercise metacognitive strategies should work in sync to improve student learning. Thus, given these results, national policymakers are highly recommended to have the following two K12 schooling outcomes in the education system: (1) academic performance measured by well-designed standardized tests and (2) measurable and malleable student approaches to academic learning. Standardized test scores tell what is happening, and student approaches to academic learning inform the reasons behind it. These reasons can be modified if predictive factors chosen to measure approaches to learning are malleable in the first place, and targeted interventions could be designed to address problems that result in student learning gaps. By having these two key performance indicators go hand in hand, the quality of schooling is highly likely to be improved. Also, by having the aforementioned two schooling outcomes embedded in

research-based education system monitoring, it is more scientific and fairer to assign proper accountability to different stakeholders within the education system to ensure they deliver key performance indicators required, be it students, teachers, or content providers.

Second, among all the predictive factors selected, instructional language factors (i.e., language and reading) are the most influential learning factors for math academic performance, and the contribution could be extremely positive or extremely negative. For those who scored >170 in language (i.e., vocabulary and grammar) and 180 in reading, math learning benefits; while for those who scored below the cut-off values in language and reading, math learning suffers. The percentage of students who scored >170 in language and 180 in reading accounts for 36.96% of the original sample of 4,107, which indicates that about 63% of the sampled students face challenges learning math in English. From a second language acquisition point of view, language skills including vocabulary comprehension and grammar contribute to reading skill development. Our findings show evidence of a correlation between language and reading. Hence, it is recommended that policymakers and education practitioners understand the language and reading skills embodied behind the cut-off values (for language 170, and for reading 180) to seek targeted intervention strategies to remediate English language proficiency. For students who have serious English literacy challenges, remediation efforts have to start with vocabulary and grammar foundation, meanwhile making sure math and science content is readable for these students is crucially important. For students who are excelling in English, advanced English language programs should be provided. Learner needs must be identified, differentiated, and met.

Taking one step further, students enrolled in Abu Dhabi public schools are first-language speakers of Arabic. Abu Dhabi Education Council (ADEC) introduced English as the instructional language for math and science in 2010 (Gallagher, 2011). However, the reality is more complex and expectations for different student segmentation groups have to be realistic. Many factors need to be considered in bilingual STEM math and science learning in UAE policymaking. For instance, teachers' English proficiency to deliver math instructions, students' English literacy level, degree of English program immersion, math, and science literacy should all be considered. Further research needs to be done in this area.

Apart from education policy refinement recommendations, promises that AI predic-tion, education technology, and digital content together can help address some of the daunting challenges in a creative way, on condition that positioning of the role of technology and digital content, the role of the teacher, and the role of the student follow key principles of the Instructional Core. First, the AI prediction model serves as a powerful tool to diagnose and predict learning behavior and learning outcomes for each student. The results in Figure 4 already show the promises. The main challenge with this model is to have measurable quality data that serve pedagogical purposes in K12 education. On the other hand, the main limitations of this study are twofold. First, data on self-system and metacognitive factors are collected from student self-report questionnaires; second, a lack of data on math cognitive strategies, teachers' English

instructional language skills and teachers' math domain knowledge exacerbates the limitations.

Despite these challenges and limitations, future research and development work in the EdTech sector would involve combining education data mining and other existing qualitative education measurements to serve the purposes of diagnosis, behavior prediction, and to guide differentiated education interventions. Second, AI prediction, education technology, and digital content solutions together could potentially provide differentiated interventions for an extremely heterogeneous student population like the one in Abu Dhabi public school system. For instance, by injecting cooperative gamification mechanism such as shared goal setting, peer learning, and immediate feedback into the technology platform design and by leveling digital content for differentiated learners, a low-achieving math learner with English literacy challenges could set their own learning goals on not just math academic learning but also on the use of metacognitive strategies to read or on how to collaborate with others. Then, this learner could receive readable digital content including scaffolded bilingual instructional language support for math, simulations, interactives, visual and audio support, while collaborating with peers to tackle a shared goal. Eventually, the learner's performance is evaluated based on multiple standards, keeping motivation and positive affective emotions intact throughout the learning processes. Apart from what technology and digital content tools can help within the Instructional Core, students still need teacher support and interaction with peers to foster social-emotional skill development. Therefore, it is important to ground platform feature design, learning analytics data dashboard design, and teacher professional development activities on both self-regulated learning pedagogical framework and collaborative peer learning to drive better academic learning outcomes and 21st-century skill development.

Acknowledgments

Funding Information

Competing Interests

The authors have no competing interests to declare.

Biography

Xin Miao is the lead researcher at Alef Education. She holds a master's degree in International Education Policy from Harvard University (HGSE) and a graduate certificate in International Development from Johns Hopkins SAIS Nanjing Center. Xin Miao has both bottom–up K12 practice experience and top–down education research and policy training. She has worked in China's highly professionalized K12 education system, explored North America and international organizations (such as UNESCO, WBG), and has been working in R&D in the EdTech industry in the UAE, focusing on impact evaluation, policy analysis, education research using AI methodologies, agile R&D in interaction design for children, public–private collaboration in K12 education reforms.

Mr. Pawan Kumar Mishra is a data scientist at Alef Education. He holds a master's degree in Mathematics and scientific computing from MNNIT Allahabad, India and is currently pursuing Ph.D. from IIT (ISM) Dhanbad, India. He has nine years of data science and machine learning experience with a proven record of building large-scale algorithms for cancer research, Industrial IOT, and education domain. He has experience of working for Indian Government and General Electric. His current research interests are aligned with the vision of Alef Education to build large-scale AI solutions to understand students' performance and learning behavior.

Dr. Ali Nadaf is the Head of Data Science at Alef Education. He holds a Ph.D. in Mathematics from Simon Fraser University and held postdoctoral fellowships in Machine Learning. His research spans a wide range of fields within the mathematical sciences and Machine Learning, and has received numerous academic awards and demonstrated record of scholarly accomplishments with published papers. He has more than 15 years of working experience in the United Nations and PayPal. His current interest is to develop an AI solution to identify and measure students' motivational and meta-cognitive behavior using students' interaction data within an educational platform.

Notes

[1] https://www.nwea.org/about/

[2] https://www.renaissance.com/about-us/

[3] https://xgboost.readthedocs.io/en/latest/

[4] https://github.com/slundberg/shap

References

[1] Bandura, A. (1997). Self-efficacy: The exercise of control. *Choice Reviews Online*, *35*(3). https://doi:10.5860/choice.35-1826

[2] Chen, T., & Guestrin, C. (2016). XGBoost. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Retrieved from https://doi.org/10.1145/2939672.2939785

[3] City E.A., Elmore R.F., Fiarman S.E., & Teitel, L. (2009). *A Network Approach To Improving Teaching And Learning*. Harvard Education Press.

[4] Coleman, C., Baker, R. S., & Stephenson, S. (2019, July 2–5). *A better cold-start for early prediction of student at-risk status in new school districts (ED599170)*. Paper presented at the International

Conference on Educational Data Mining (EDM), Montreal, Canada Retrieved from https://eric.ed.gov/?id=ED599170

[5] Crossley, S., Karumbaiah, S., Ocumpaugh, J., Labrum, M. J., & Baker, R. S. (2019). Predicting math success in an online tutoring system using language data and click-stream variables: A longitudinal analysis. *2nd Conference on Language, Data and Knowledge* (LDK 2019). Retrieved from https://doi.org/10.4230/OASIcs.LDK.2019.25

[6] Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, *92*(6), 1087–1101. https://doi.org/10.1037/0022-3514.92.6.1087

[7] Efklides, A., Schwartz, B. L., & Brown, V. (2018). Motivation and affect in self-regulated learning: Does metacognition play a role? In D. H. Schunk & J. A. Greene (Eds.), *Educational Psychology Handbook Series*. *Handbook of self-regulation of learning and performance* (pp. 64–82). Routledge.

[8] Farah, S. (2012). Education Quality & Competitiveness in the UAE (Sheikh Saud bin Saqr Al Qasimi Foundation for Policy Research Policy Paper No. 2). doi:10.18502/aqf.0011

[9] Friedman, J.H. (2001) Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, *29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

[10] Gallagher, K. (2011). Bilingual education in the UAE: Factors, variables and critical questions. *Education, Business and Society: Contemporary Middle Eastern Issues*, *4*(1), 62–79. https://doi.org/10.1108/17537981111111274

[11] The government of Abu Dhabi. (2008, November). *The Abu Dhabi Economic Vision 2030*. Retrieved from https://www.actvet.gov.ae/en/Media/Lists/ELibraryLD/economic-vision-2030-full-versionEn.pdf

[12] Hanushek, E. A., & Kimko, D. D. (2000). Schooling, labor-force quality, and the growth of nations. *American Economic Review*, *90*(5), 1184–1208. https://doi.org/10.1257/aer.90.5.1184

[13] Harvey, J. L., & Kumar, S. A. (2019). A practical model for educators to predict student performance in K-12 education using machine learning. *2019 IEEE Symposium Series on Computational Intelligence* (SSCI). Retrieved from https://doi.org/10.1109/ssci44817.2019.9003147

[14] Kippels, S., & Ridge, N. (2019). The growth and transformation of K–12 education in the UAE. In Gallagher K. (Eds.), *Education in the United Arab Emirates* (pp. 37–55). Singapore: Springer. https://doi.org/10.1007/978-981-13-7736-5_3

[15] Kostyuk, V., Almeda, M. V., & Baker, R. S. (2018). Correlating affect and behavior in reasoning mind with state test achievement. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, 26–30. Retrieved from https://doi.org/10.1145/3170358.3170378

[16] Kumari, P., Jain, P. K., & Pamula, R. (2018). An efficient use of ensemble methods to predict students' academic performance. *2018 4th International Conference on Recent Advances in Information Technology* (RAIT). Retrieved from https://doi.org/10.1109/rait.2018.8389056

[17] Lundberg, S. M., & Lee, S. I. (2017). A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems*. Retrieved from https://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions

[18] McCombs, B. L. (1986). The role of the self-system in self-regulated learning. *Contemporary Educational Psychology*, *11*(4), 314–332. https://doi.org/10.1016/0361-476x(86)90028-7

[19] Mokhtari, K., Dimitrov, D. M., & Reichard, C. A. (2018). Revising the Metacognitive Awareness of Reading Strategies Inventory (MARSI) and testing for factorial invariance. *Studies in Second Language Learning and Teaching*, *8*(2), 219–246. https://doi.org/10.14746/ssllt.2018.8.2.3

[20] OECD. (2010). *Mathematics Teaching and Learning Strategies in PISA*. Retrieved from https://www.oecd-ilibrary.org/docserver/9789264039520-en.pdf?expires=1586113622&id=id&accname=guest&checksum=81842A52D802FE8FEBB18E513DB02268

[21] OECD. (2018). *PISA Results 2018: UAE Country Note*. Retrieved from http://www.oecd.org/pisa/publications/PISA2018_CN_ARE.pdf

[22] Pardos, Z. A., Baker, R. S., San Pedro, M., Gowda, S. M., & Gowda, S. M. (2014). Affective states and state tests: Investigating how affect and engagement during the school year predict end-of-Year learning outcomes. *Journal of Learning Analytics*, *1*(1), 107–128. https://doi.org/10.18608/jla.2014.11.6

[23] Shapley, L. S. (1953). A value for n-person games. In H. W. Kuhn & A. W. Tucker (Eds.), *Contributions To the Theory of Games* (AM-28), Volume II (pp. 307–318). Princeton University Press. https://doi.org/10.1515/9781400881970-018

[24] UAE Ministry of Education. (2010). *The Ministry of Education Strategy 2010–2020: Aiming in Accomplishing a Score of 10/10 in All Its Initiatives*. Retrieved from https://www.moe.gov.ae/Arabic/Docs/MOE%20_Strategy.pdf

[25] UAE Ministry of Finance. (2018). *Federal Budget 2018*. Retrieved from https://www.mof.gov.ae/en/resourcesAndBudget/fedralBudget/Pages/budget2018.aspx

[26] UAE Ministry of Finance. (2019). *Federal Budget 2019*. Retrieved from https://www.mof.gov.ae/en/resourcesAndBudget/fedralBudget/Pages/budget2019.aspx

[27] Vision 2021. (2005, December). *UAE Vision 2021*. Retrieved March 16, 2020 from http://www.vision2021.ae

[28] UNESCO. (n.d.). *Sustainable Goals for UAE*. Retrieved from http://uis.unesco.org/en/country/ae

[29] Winne, P. H. (1985). Steps toward promoting cognitive achievements. *The Elementary School Journal*, *85(5),* 673–693. https://doi.org/10.1086/461429

[30] Winne, P. H. (1997). Experimenting to bootstrap self-regulated learning. *Journal of Educational Psychology*, *89*(3), 397–410. https://doi.org/10.1037/0022-0663.89.3.397

[31] Winne, P. H. (2010). Bootstrapping learner's self-regulated learning. *Psychological Test and Assessment Modeling*, *52*, 472–490. Germany: Pabst Science Publishers