JOURNAL OF BALTIC SCIENCE EDUCATION

ISSN 1648-3898 /Print/
ISSN 2538-7138 /Online/

USING A MACHINE LEARNING APPROACH TO EXPLORE NON-COGNITIVE FACTORS AFFECTING READING, MATHEMATICS, AND SCIENCE LITERACY IN CHINA AND THE UNITED STATES

Lu Ye, Yuqing Yuan

Abstract. *Non-cognitive factors are considered critical aspects in shaping students' academic achievement. This study aims to analyze and explore the mechanisms of the influence of non-cognitive factors on 15-year-old students' abilities in China and the United States. Based on the Programme for International Student Assessment (PISA) 2018 education dataset, the Classification and Regression Tree (CART) model identifies and explains the factors. The study finds that there are 11 most influential common features in China and 9 in the United States. The two countries have 5 common features, the meta-cognition assess credibility, summarizing text ability, PISA test difficulty perception, science learning time, and school lessons numbers per week. Family economic status also impacts personal ability. Regarding subject characteristics, attitude towards failure is the determinant of reading and mathematics. Cooperation and competition among students help to improve mathematics and science. Furthermore, the comparison between the two countries concludes that self-awareness, family economic status, and school learning environment are critical to personal ability. The study concludes that it is necessary to foster a sense of healthy competition among students at the school level and provide more attention to students with low family socioeconomic status to improve their abilities.*

Keywords: *machine learning approach; non-cognitive factors; PISA 2018; personal ability*

Lu Ye, Yuqing Yuan
Zhejiang Sci-Tech University, China

Introduction

Personal ability is a student's ability to sift, compare, contrast, and integrate information from multiple sources (Duckworth & Yeager, 2015). The degree of personal ability is significant in determining whether students have acquired the knowledge and skills to adapt to future society. The Programme for International Student Assessment (PISA) regularly evaluates the reading, mathematics, and science literacy of 15-year-old students to assess their personal abilities. Education reflects cultural variations, which lead to various educational ideals, educational views, educational systems, educational techniques, and evaluation standards. As a result, education in China and the United States exhibits glaring contrasts. Based on the viewpoint of cultural values, Zhao et al. (2006) contend that the primary distinction between Chinese and American educational concepts is the distinction between general education and individual education, which is primarily manifested in: innovation and uniformity, independent thought and passive acceptance, equality, looseness and strict authority, and emphasis on knowledge teaching and ability training. In the PISA 2018, in 4 provinces and cities in China (Beijing, Shanghai, Jiangsu, and Zhejiang), students ranked first in the three subjects among the participating countries. The students' basic literacy achievement rate is the first among the participating countries, and the total number of high-level students is the highest. From the vertical point of view, China has made significant progress compared with the 10th ranking in PISA 2015. In a side-by-side comparison, the United States ranked 25th globally, 13th in reading, 18th in mathematics, and 37th in science, ranking low among the 79 countries participating in the test. And with 6% of GDP invested in education, the United States is at the top of the world. The annual investment per student is more than $15,000, more than ten times that of China, ranking second among the Organization for Economic Co-operation and Development (OECD) countries (U. S. Department of Education, 2014). When comparing the two countries we cannot ignore the differences in the reading of Chinese versus English. Chinese is a logographic writing system and is visually denser than English. English words consist of letter strings, while Chinese words are composed of characters. They operate differently in attention, word identification, and eye-movement control (Yu & Reichle,

https://doi.org/10.33225/jbse/22.21.575

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

2017). From the results of input leading and achievement developed countries backward, it is a recognized fact to evaluate the high investment and poor effect of primary and secondary education in the United States. Therefore, studying the factors influencing personal ability can explore why China's primary education is ahead of the United States in three subjects and has great significance in promoting the overall development of the educational environment in developing countries.

This study adopts the Classification and Regression Tree (CART), a machine learning approach to research the complex relationships between non-cognitive factors. The CART model examines the effects of non-cognitive factors on students' reading and mathematics science scores and whether the influencing factors vary from discipline to discipline, which can help us understand the impact of different non-cognitive factors on personal quality and how it affects students' reading, mathematics, and science ability. Further, understand the differences between the two countries and how to improve students' personal ability.

The contributions of the paper are threefold. First, the study investigated the non-cognitive factors in the performance of three subjects and enriched the research results on the factors influencing students' personal ability. Based on PISA 2018 Chinese and American data, all relevant non-cognitive variables were selected. To explore the factors influencing students' personal ability and identify the key variables that distinguish high ability students from low ability students. By improving these variables, students' personal abilities will be enhanced.

Second, the study applied the machine learning method to study this problem and enriched the application of machine learning in pedagogy. Due to PISA data's large and complex relationship, there are more non-response errors. And the initial variables in this study cover a wide range and a large number. So, CART, which implements a decision tree classification algorithm, was chosen to avoid the subjective bias brought by manual selection.

Third, common and different features are derived by comparing and analyzing the performance of non-cognitive factors on students' achievement in three subjects in China and the United States. Education is the nation's foundation. China and the United States are typical countries with different historical, cultural, and psychological backgrounds. The United States is the representative developed country, whereas China is the largest developing country. On the one hand, although national conditions are different, for personal ability itself, the factors that influence personal ability performance identified in this study using machine learning methods are objective. On the other hand, also taking into account the differences in national conditions, factors specific to both countries are further analyzed. This study compares and analyzes the influencing factors of students' personal abilities in the two countries, reformulates the differences between Chinese and American education, and provides feasible suggestions for improving students' personal abilities.

Literature Review

Previous studies have found that many factors affect students' performance in personal ability (Lee & Shute, 2010; Yorulmaz et al., 2009; Eriksson et al., 2021). Scholars have focused on teaching methods and activities, but these instructional factors are likely to interact with students' specific dispositions (Tonga et al., 2019; Bellova et al., 2018). Too much attention given to learners' general and specialized cognitive talents and intellectual tendencies may limit the ability to profit from various types of experience (Heckman et al., 2006). However, little research has been done on the links between personal ability and non-cognitive factors to see how they alter. Duncan and Murnane (2014) stated that non-cognitive factors are related to the quality of teaching and the internal capacity accepted by students. The educational experience transforming content-based knowledge into applied knowledge may be shaped by non-cognitive factors related to students' individual, family, and school. There are non-cognitive factors such as seriousness and learning methods related to reading, mathematics, and science that shape personal ability. For example, learning style (Tseng & Chu, 2018), reading frequency, self-cognition, and learning methods, including meta-cognition and knowledge application ability, are essential predictors of readers' good reading ability. However, other factors not explicitly related to students' study experiences are also associated with personal ability. These factors are collectively called motivational factors, such as mastery or learning goals (Toste & Didion, 2020), learning drive, and goal attractiveness. Hobri et al. (2018) also emphasized that critical thinking is one of the essential competencies individuals should have in modern society and a top priority in higher education. It is an objective and self-regulated judgment aimed at solving problems or making decisions that can be recognized as a predictor of potential study success.

In addition to the factors related to students' personal ability, the research also found environmental factors affecting students' personal ability (Chen et al., 2021; Bailey et al., 2010; Krieken et al., 2015). Environmental influ-

https://doi.org/10.33225/jbse/22.21.575

ences frequently provide resources and support for learning and growth processes connected to improved reading, mathematics, and scientific literacy. Juan and Visser (2017) found that students' most pertinent social contexts are the home and school environments, involving different actors, social interactions, and resources.

In terms of the home environment, Chen et al. (2019) have found that parents' educational attainment, employment position, and home assets are all factors that directly affect students' achievement. Huang et al. (2021) also claimed that these home assets, such as culture and learning resources, tend to assist pupils' motivation and resources for personal ability improvement. Personal ability is also influenced by parents' expectations, educational techniques, and family environment. Tan and Hew (2017) have shown that home use of information and communications technology (ICT) positively impacts students' reading, mathematics, and science performance. Hu and Gong (2018) found that students with higher ICT levels perform better than those with lower ICT levels. Furthermore, these factors in the home environment are linked to families' socio-economic status.

Regarding classroom teaching, with the rise in popularity of quality education in the international community (Aditomo & Koehler, 2020), China's quality education is rising. The essence of teaching and educating people is gradually being highlighted, encouraging people to conduct in-depth examinations and pay attention to the influencing factors of students' personal ability. As a result, the evaluation of teaching levels emerges at a historical juncture and spreads swiftly worldwide (Shao & Sun, 2019). The reform of China's education system has intensified in recent years, and the size of education has been continuously enlarged, increasing the impact of teaching level on human ability. Conducting rigorous, objective, and quantitative assessments of students' learning environments and teaching approach is critical.

Beyond the classroom, there are also essential factors in the school environment that support student learning and achievement (Gimenez & Ciobanu, 2021; Liu et al., 2017; Saeki & Quirk, 2015; Laftman et al., 2017). For example, Gimenez and Ciobanu (2021) concluded that the peer effect is essential for academic performance. The class composition according to gender or race, students' ability, and their socio-economic levels are the most commonly used characteristics to measure peer effect. Furthermore, each school has different learning resources, including basic infrastructure, materials, and teaching resources (Liu et al., 2017). These issues were frequently linked to school financing sources and the overall quantity of resources available. Some school-level factors not related to school financing resources have also influenced the students' academic performance to a large extent. Although not related to school financing resources, students' social and interpersonal experiences are equally important in the school environment. For example, students, social connections (Saeki & Quirk, 2015), and exposure to bullying (Laftman et al., 2017) have also been essential predictors of personal ability.

Research Aim and Research Questions

From the above research, students have high ability, which comes from the support of personal learning, family economic status, school resources, and other aspects. Although a general association between these influence factors and personal ability has been found in prior research, the influences of these factors on personal ability may vary in different subjects (reading, mathematics, and science). In addition, most research articles usually only focus on some specific factors, and few articles systematically and comprehensively study the relationship between all non-cognitive factors and personal ability. The development of non-cognitive factors is necessary for students to acquire good literacy and ability, which will improve students' adaptability to future society.

Therefore, this study aimed to explore the influence of non-cognitive factors on personal ability. The following research questions were determined:

(a) Do students' grades in three subjects make a statistically significant correlation? If so, what non-cognitive factors simultaneously impact all three subjects?
(b) What are the subject characteristics of reading, mathematics, and science?
(c) What are the same and different non-cognitive factors in China and the United States?

Research Methodology

General Background

This study was divided into three parts to explore the critical non-cognitive factors that influence students' personal abilities. In the first part, factors influencing all three subjects of reading, mathematics, and science were

https://doi.org/10.33225/jbse/22.21.575

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

identified and described qualitatively and quantitatively based on machine learning methods and existing theoretical frameworks. Next, for the second part, the single-subject characteristics of the three subjects were explored. In the third part, the results from both countries were comparatively analyzed to obtain constructive suggestions.

Instrument and Procedures

Based on the proposed conceptual framework for the relationship between non-cognitive factors and students' academic performance, this study employed variables from the PISA 2018 database. PISA 2018 mainly assessed students' literacy in reading, mathematics, and science, and all three were used as dependent variables in the present study. The variables analyzed in this study are divided into three categories: personal variables, family variables, and school variables. Personal variables refer to the learning motivation, learning attitude, and learning purpose related to reading, and family variables refer to the characteristics of parents and learning resources at home. Finally, school variables refer to the teachers' teaching methods, classroom learning experience, the school's educational resources, and students' life experiences in school.

In this study, 114 initial variables were set. First, set the missing values of variables to 0, delete the samples with more than half of the missing value, and then use k Nearest Neighbor for interpolation. Finally, 62 variables were considered totally in China, including 36 student-level variables, 10 family-level variables, and 16 school-level variables. There were 73 variables in the United States, including 40 student-level variables, 13 family-level variables, and 20 school-level variables. The differences between the Chinese and American variables were mainly in the ICT questionnaire. Since Chinese students lacked a questionnaire on ICT use, only one variable about ICT was set. The data from the ICT questionnaire of American students were complete and comprehensive, and nine relevant variables were selected in this study.

Some variables in this study were based on the original indices, while some were derived variables that PISA develops. The derived variables were calculated using the Item Response Theory (IRT), which is a general term for a series of psychometric models (Davier et al., 2019). The purpose of IRT was to determine whether underlying characteristics can be reflected by the PISA questions and the interaction between the test questions and the test taker.

Data Sources

The data for this study were from the Chinese and American samples in the OECD PISA 2018 database. PISA examines 15-year-old students' personal ability using a stratified sample approach. 12,058 Chinese students were sampled from 361 schools across Beijing, Shanghai, Jiangsu, and Zhejiang. And 4,838 students were sampled from across the United States. Moreover, we combined the ICT familiarity and students' social background questionnaires completed by students to gather student information. After deleting the invalid questionnaire, it was finally determined that the sample size of 11,977 in China and 4,838 in the United States.

Data Analysis

A data mining approach was often deployed for data analysis when detecting and interpreting a large database with complex relations between numerous variables. Given many possible predictors of different types in this study, CART is adopted. CART belonged to the supervised learning methods of data mining techniques, for specifying the conditional distribution, given a vector of predictor values. Such a model was developed on a random subset of the data (training sample), and then the results were validated on a separate random sample (test sample) (Breiman & Friedman, 2015; Strobl et al., 2009).

CART did not make any assumptions about the probability distribution of the individuals sampled from it and was not affected by the issue of multi-collinearity between predictors. According to statistical criteria, the algorithm divided participants into consecutive binary groups. Each independent variable was assessed by splitting the subjects into two groups (referred to as the child nodes) based on their capacity to reduce the contaminants of the parent node, starting with the entire sample (referred to as the root or parent node). The independent variables might be sorted, continuous, or nominal (Hox, 2010). Non-cognitive factors were entered into the CART: student-, family-, and school level.

For machine learning, a comparable distribution of each group, low and high, was preferable to remove bias in model training. Students were divided into two groups: those with low ability and those with high ability.

https://doi.org/10.33225/jbse/22.21.575

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Low-ability students were those with limited ability at reading levels two and below, mathematics levels three and below, and science levels three and below. High-ability students were those with reading at levels three and better, mathematics at levels four and better, and science at levels four and better. As dependent variables, the study had taken the values of the scores obtained by the school in the general skills evaluated in each competence (PV1READ, PV1MATH, PV1SCIE) (OECD. PISA 2018 Data Base). At the same time, other items in the questionnaire linked to personal ability were used as independent variables.

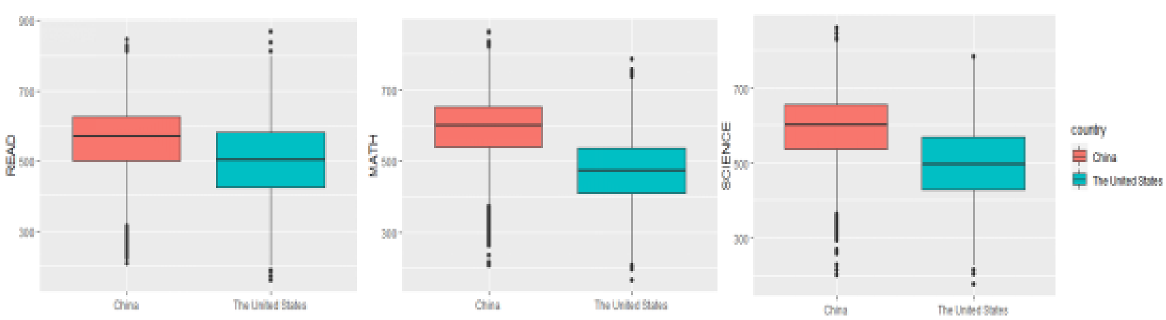
After data processing, first, student, family, and school-level non-cognitive factors are used as independent variables. After being divided into high-low levels, the scores of the three subjects are used as dependent variables to build CART models for different subjects. Second, the top 20 most relevant factors for each of the three subjects are derived separately. Third, in the top 20 factors of the three subjects, find the same features affecting the three subjects. Expressed them as the key influencing factors on students' personal ability, that is, the common features. Fourth, factors affecting only two subjects were identified in addition to the factors common to all three subjects. These factors can reflect the subject's characteristics. Fifth, the above steps are the same in China and the United States, and the results in the two countries will be compared and analyzed.

Research Results

Descriptive Statistics

In PISA 2018, Chinese students scored an average of 555, 591, and 590 in reading, mathematics, and science. American students scored an average of 500, 493, and 497 in reading, mathematics, and science. Mathematics and science were nearly 100 points lower than Chinese students. The distributions of scores in three subjects in China and the United States are summarized in Figure 1.

Figure 1
Distribution of Scores in Each Subject

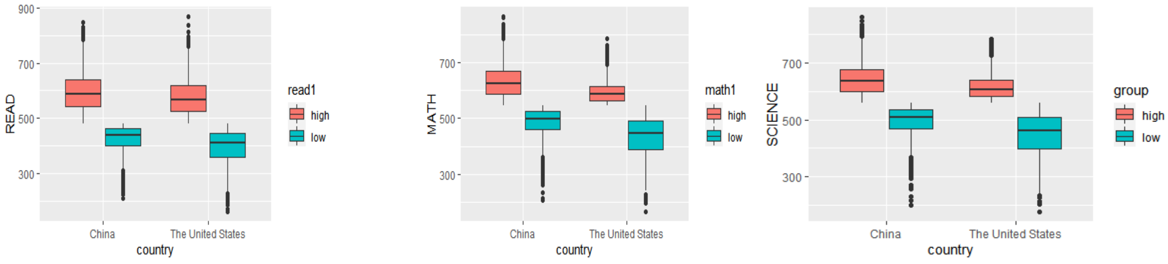


The students (China N=11,977, the United States N=4,838) were divided into low-high groups according to their scores in the three subjects. As a result, 2,252 (19%) and 9,725 (81%) Chinese students belonged to the low and high reading performing groups, respectively. 3,195 (27%) and 8,782 (73%) in low-high mathematics groups, 3,839 (32%) and 8,138 (68%) in low-high science groups. The distributions of the United States showed that 2,040 (42%) and 2798 (58%) of the students belonged to the low and high reading performing groups. Low-ability students dominated mathematics and science. 3,720 (77%) and 3,463 (72%) students belonged to the low mathematics and science performing groups. China's mathematics and science scores were higher than the United States overall and were spread over a wide range of high-level students. Figure 2 showed that China's performance in all three subjects was higher than that of the United States and was concentrated among low-level students.

https://doi.org/10.33225/jbse/22.21.575

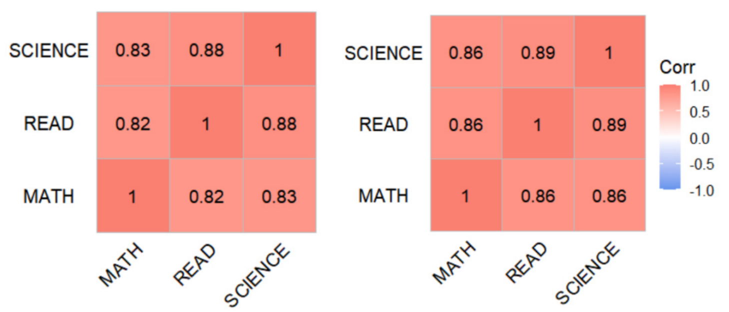
ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Figure 2
Distribution of Scores in Each Subject Using Low-high Groupings



The correlation diagram was shown in Figure 3. The first picture was China, and the second picture was the United States. Of the three subjects, they all had a strong correlation. Scores of the variables (PV1READ, PV1MATH, PV1SCIE) presented high correlations. The results were statistically significant under the 95% confidence interval, and the pairwise correlations were between 0.8 and 0.9. Comparing the two countries, the correlation coefficients of mathematics and science were 0.88 and 0.89. The two countries' correlation between mathematics and science was consistent. Furthermore, the correlation between reading and mathematics, reading and science American data showed a stronger correlation than in China.

Figure 3
The Correlation Diagrams of Three Subjects in China and the United States



Common Features of Three Subjects in China

Table 1
Common Features of the Three Subjects Identified in China

Rank		Feature	Description	Value range (min.~max.)
Student	1	METASPAM	Meta-cognition: assess credibility	-1.41~2.33 (IRT scaling)
	3	METASUM	Meta-cognition: summarizing	-1.72~1.36 (IRT scaling)
	5	UNDREM	Index understanding and remembering	-1.64~1.5 (IRT scaling)
	6	JOYREAD	Index enjoyment of reading	-2.7114~2.6574 (IRT scaling)

https://doi.org/10.33225/jbse/22.21.575

Rank		Feature	Description	Value range (min.~max.)
	8	PISADIFF	Perception of difficulty of the PISA test	-1.272~3.0064 (IRT scaling)
	10	RESILIENCE	Resilience	-3.1675~2.3693 (IRT scaling)
Family	9	ESCS	Index of economic social and cultural status	-5.0771~3.1015 (IRT scaling)
School	2	SMINS	Instructional time: Science courses	0~2210
	4	ST059Q03TA	Number of class periods in science	0~40
	7	ST060Q01NA	number of class periods per week.	10~80
	11	BELONG	Index sense of belonging	-3.2583~2.7562 (IRT scaling)

In China, 11 common features with high scores in three subjects were selected from 67 factors. Among these 11 factors, six were related to individual students, four were at the school level, and only one was related to families. As a result, the 11 common features were identified and are provided in Table 1.

At the student level, the most important variable related to classifying students into low and better-performing students in reading, mathematics, and science was METASPAM. A dependability index displayed people's sincerity in dealing with things and was positively connected with the target variable. Similarly, METASUM, the third most crucial variable, was positively correlated with the target variable. METASUM is an index for the students' meta-cognitive awareness of strategies for summarizing texts. UNDREM is a measure of one's capacity to comprehend and recall information. Students with a good understanding and recall tend to do well in school. JOYREAD is an index of enjoyment of reading, and PISADIFF is an index of subjective difficulty perception of the PISA test. RESILIENCE is resilience to events, also known as self-healing ability. These six variables indicated that students' self-efficacy was closely related to personal ability. Poor readers differed from better readers because they have lower intrinsic sensitivity to learning and weaker meta-cognitive awareness of learning strategies. These results were consistent with the literature on the role of self-efficacy (Li et al., 2020), intrinsic enjoyment of learning (Areepattamannil et al., 2011), and metacognitive awareness of strategies in education (Sun-Lin & Chiou, 2017) in students' achievement.

ESCS was the only family variable among the 11 key factors that affect all three topics at the family level. The ESCS contained a basic summary of each parent's educational level, family property status, and job position. In terms of family social status, both the highest educational background and family property status of parents had positive correlations with the scores of all subjects. Increased family status meant that students had access to better learning resources, positively impacting their reading, mathematics, and science literacy.

4 of the top 20 variables were non-cognitive school variables related to reading, mathematics, and science. There were two variables, a single item was included in some other indexes, and one variable was a single item measurement of a factor. These variables were discussed in more detail. SMINS is an index of students' science courses learning time per week, including the entire learning time of in-class and out-of-class scientific studies. This research considered that curriculum learning belonged to teaching rather than self-study. Thus, it was classed as a school-level variable. Science learning time (SMINS) and scientific-related course numbers per week (ST059Q03TA) demonstrated that science study benefited personal development. ST060Q01NA is an index of school lesson numbers per week, and this variable was positively connected with the target variable. Scientific learning may develop people's cognitive capacity and practical ability. Students who take many school classes are likely to receive good scores, which supports the opinion of Woods-McConney (2013, 2014). In addition to the number of hours and sessions of classroom learning, BELONG, a variable that measures a student's sense of belonging to the school, was also a significant predictor of personal ability. BELONG is an index measuring the students' sense of belonging in their schools, showing students' cognition, emotion, motivation, conduct in school interactions, classroom learning time, and the number of classes.

https://doi.org/10.33225/jbse/22.21.575

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Common Features only Two Subjects in China

Table 2
Common Features only Two Subjects Identified in China

Special	Feature	Description	Value range (min.~max.)
read×math	GFOFAIL	General fear of failure	-1.8939~1.8905 (IRT scaling)
read×science	EUDMO	Eudaemonia: meaning in life	-2.1464~1.7411 (IRT scaling)
math×science	PERFEED	Index perceived feedback	-1.6391~2.0165 (IRT scaling)
	HEDRES	Index home educational resources	-4.4106~1.2099 (IRT scaling)
	PERCOOP	Perception of cooperation at school	-2.1428~1.6762 (IRT scaling)
	PERCOMP	Perception of competitiveness at school	-1.9892~2.0378 (IRT scaling)

More pertinent to the research questions in China, the results of the common features in every two subjects analyzed for reading, mathematics, and science were summarized in Table 2. Both reading and mathematics and reading and science had only one particular shared variable, GFOFAIL, and EUDMO, respectively, and both were at the level of individual student variables. GFOFAIL is the fear of failure Index, which is how students react when faced with failure. Lower scores indicated weaker resistance to failure and a greater tendency to be poor learners. The index ranked 17 in reading predictors and 15 in mathematics, and more scientifically, it had a relatively large impact on reading and mathematics performance. EUDMO is an index of how much students think about the meaning of life. The results showed that students with clear life goals and long-term plans had higher reading and science scores. Thinking about the meaning of life played a decisive role in the effective orientation of one's role. Clear goals guided and promoted one's growth and positively impacted the improvement of one's abilities.

As shown in Table 2, mathematics and science contained four special common features. Three belonged to the school level. PERFEED is an index to measure the quality of teachers' instructional practice. PERCOOP and PERCOMP belonged to the 'competitive and cooperative relationship of students in learning' factor. In addition, HERDES belonged to the family category, indicating the family education resource index. The comprehensive literature review also revealed the consistency between the results of this study and those previously reported. For example, teaching quality and family resources have been recognized as important factors in mathematics and science literacy (Lau & Lam, 2017).

Common Features of Three Subjects in the United States

Table 3
Common Features of the Three Subjects Identified in the United States

Rank		Feature	Description	Value range (min.~max.)
Student	1	PISADIFF	Perception of difficulty of the PISA test	-1.272~3.0064 (IRT scaling)
	2	METASPAM	Meta-cognition: assess credibility	-1.41~2.33 (IRT scaling)
	4	METASUM	Meta-cognition: summarizing	-1.72~1.36 (IRT scaling)
	5	AUTICT	Perceived autonomy related to ICT Use	-2.5144~2.0258 (IRT scaling)

https://doi.org/10.33225/jbse/22.21.575

Rank		Feature	Description	Value range (min.~max.)
	7	COMPICT	Index ICT: Perceived ICT competence	-2.6033~1.9885 (IRT scaling)
	9	COMPETE	Competitiveness	-2.345~2.0054 (IRT scaling)
Family	6	WEALTH	Family wealth PISA 2006	-6.8518~4.2599 (IRT scaling)
School	3	ST060Q01NA	number of class periods per week.	10~80
	8	SMINS	Instructional time: Science courses	0~2210

Among the 9 variables of these three subjects selected from 73 factors in the United States, 6 were related to individual students, 2 were at the school level, and only 1 was related to families. As a result, the 9 common features were identified and are provided in Table 3.

At the student level, as illustrated in Table 3, the most important variable related to classifying students into low and better-performing students in reading, mathematics, and science was PISADIFF. PISADIFF is an index of subjective difficulty perception of the PISA test. This indicator determined students' ability to perceive the difficulty of the PISA test in 3 aspects: storage of word statements, overall comprehension of the text, and logical thinking when doing the questions. Ranking 2nd and 4th are METASPAM, an index of how seriously students take things, and METASUM, an index of their ability to summarize texts. They were indexes of metacognitive awareness and strongly predicted Chinese students' academic performance, positively correlated with the target variable. AUTICT is an index of perceived autonomy related to ICT use. The question "How often do you use digital devices for the following activities outside of school?" reflected the autonomy of ICT use by investigating the frequency of active use of electronic devices or information technology outside of school. COMPICT is an individual's self-perception of the ability to use information communication. Unlike autonomy, this index measures students' perceived ability to use ICT. Students with solid perceptions learned efficiently when using ICT, leading to good learning outcomes, and thus helping improve academic performance. The last of the nine common features in the United States was COMPETE. It ranked 17th in reading and mathematics and 18th in science among the top 20 variables extracted for a single subject. It was clear that although students' sense of competition did not have a decisive effect on personal ability, it did affect all three subjects, and developing a sense of competition to improve personal ability was an area of concern.

WEALTH was the only family variable among the nine key factors affecting all three topics at the family level. WEALTH is an index of a family's wealth, which provides a picture of the family's wealth status. The family's economic situation influenced the learning environment and the educational resources provided to the child, and these resources were closely linked to the child's success. A more robust family economy brought more social capital and human resources to invest in their children. On the other hand, families with weaker economic power will have a disproportionate amount of resources invested in meeting the family's basic needs and then will have an insufficient investment in education. Families of different economic levels also have different levels of parenting for their children, which can affect their children's academic performance.

2 of the top 20 variables were non-cognitive school variables related to reading, mathematics, and science. One single variable item was included as part of some other index, and one variable was a single-item measure of a factor. School lesson numbers per week (ST060Q01NA) and science learning time (SMINS) were prescribed hours of study. ST060Q01NA was ranked 3rd among the total variables, which showed that the higher the number of courses of study established by the school, the more it contributes to students' academic performance. In-school education played a vital role in developing the personal abilities of students.

https://doi.org/10.33225/jbse/22.21.575

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Common Features only Two Subjects in the United States

Table 4
Common Features only Two Subjects Identified in the United States

Special	Feature	Description	Value range (min.~max.)
read×math	TMINS	Total Learning time (minutes per week)	170~3000
	SCREADCOMP	Self-concept of reading: Perception of competence	-2.4403~1.8839 (IRT scaling)
	ST059Q04HA	Number of class periods per week in foreign language	0~34
	ST185Q01HA	Clear meaning or purpose.	1~4
	GFOFAIL	General fear of failure	-1.8939~1.8905 (IRT scaling)
read×science	HOMEPOS	Index home possessions	-7.8077~5.3188 (IRT scaling)
	ICTOUTSIDE	Subject-related ICT use outside of lessons (WLE)	-1.3048~2.4969 (IRT scaling)
math×science	ESCS	Index of economic social and cultural status	-4.0953~3.2545 (IRT scaling)

In the United States, the results of the common features in every two subjects' analyses were summarized in Table 4. Common features were specific to Americans in reading and mathematics, 4 from the student level and 1 from the school level. Total time spent by students studying (TMINS) and the number of foreign language courses (ST059Q04HA) were important indexes of high or low student achievement in reading and mathematics. Students' study time can be extended to homework and other extracurricular activities (Xiao & Hu, 2019). In this study, students with high literacy tended to spend more time studying each week. In addition to study time, the factor of metacognitive awareness was also important. Research has shown that students who read with an explicit self-concept can grasp the focus of a text and effectively grasp the entire meaning (Salchegger, 2016). As a result, these students achieve better academic performance. Clarity of personal planning (ST185Q01HA) was also an important indicator of students with high literacy. GFOFAIL is the fear of failure index, the only variable affecting Chinese students' reading and mathematics performance.

In this study, other home possessions like a desk that students can use and a quiet place to study (included in HOMEPOS) were associated with students' reading and science literacy, consistent with the previous study (Ngorosho, 2010). What's more, ICT available out of lessons (ICTOUTSIDE), including test language lessons, science, music, and sports, was also a strong index of students with high reading and science literacy (Hu et al., 2018). The frequency of using electronic devices outside of the classroom was a better index of students' ability to use ICT. As digital reading became increasingly important, PISA 2018 utilized a new e-reading model. Students' ability to use ICT subjectively became particularly important.

ESCS is an index of family socioeconomic status. Numerous studies have demonstrated that socioeconomic status significantly affects students' academic performance (Flores-Mendoza et al., 2021; Muelle, 2020). And unlike previous studies, the index positively affected all three subject scores in China, but only mathematics and science were associated with it in the United States. It ranked 4th in mathematics and 3rd in science. This may be since the United States invests 6% of its GDP in education resources, which is at the top of the world and is more prevalent at home. Regardless of the family's economic status, books, laptops, and other learning materials are available at home, less impacting children's reading motivation.

https://doi.org/10.33225/jbse/22.21.575

A Comparative Analysis of China and the United States

Table 5
Comparison of Common Features between China and the United States

	Rank in China	Feature	Rank in US	Feature
Student	1	METASPAM	1	PISADIFF
	3	METASUM	2	METASPAM
	5	UNDREM	4	METASUM
	6	JOYREAD	5	AUTICT
	8	PISADIFF	7	COMPICT
	10	RESILIENCE	9	COMPETE
Family	9	ESCS	6	WEALTH
School	2	SMINS	3	ST060Q01NA
	4	ST059Q03TA	8	SMINS
	7	ST060Q01NA		
	11	BELONG		

The section comparatively analyzed the factors between China and the United States. There were nine common features in the United States, 5 of which were non-cognitive factors identical to China's. 3 came from the individual student level, PISA test difficulty perception (PISADIFF), meta-cognition assess credibility (METASPAM), and summarizing texts ability (METASUM). 2 came from the school level, science learning time (SMINS), and school lessons numbers per week (ST060Q01NA). All 3 variables at the student level involved cognitive abilities, perceived difficulty of the PISA test, and the student's attention and summarizing and analyzing ability when treating things. The 2 variables on metacognition (METASPAM, METASUM) were both ranked high. However, PISADIFF, the perceived difficulty of the PISA tests, was ranked very differently. In China, PISADIFF ranked 9th out of 11 common features at the bottom of the list and had a relatively small effect on performance in the three subjects. This indicated that students' perceptions of the difficulty of the test were highly influential in their final performance. The two variables at the school level were the number of weekly lessons (ST060Q01NA) and science study time (SMINS). These two variables also differed in importance for the two countries. Science learning time (SMINS) was more important for China than for the United States, ranking 2nd in China and 8th in the United States. However, the number of weekly lessons (ST060Q01NA) was highly relevant to American student achievement. However, it ranked 7th in China, which showed that science learning was more effective in promoting student achievement in China than in the United States. At the same time, the number of school courses was more effective in the United States.

2 of the 4 non-cognitive factors specific to the United States were related to ICT, which was significantly different from China. AUTICT reflected the autonomy of ICT use by investigating the frequency of active use of electronic devices or information technology outside school. COMPICT measured students' perceived ability to use ICT. The other two variables specific to the family level were WEALTH and COMPETE, a student-level measure of student competitiveness. Similar to the Chinese family indicator ESCS, both were the educational resources or life assistance the family can provide to the student, but the difference was that WEALTH measured the family's property status. At the same time, ESCS was more focused on the family's social situation. ESCS was not a common feature in all three subjects in the United States. However, it was at a high level in the ranking of importance in mathematics (4th in a single subject) and science (3rd in a single subject). It was not a significant factor in reading, as it was located after the 20th place.

https://doi.org/10.33225/jbse/22.21.575

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Discussion

To explore the non-cognitive factors influencing students' personal abilities, this study applied the CART algorithm to analyze the reading, mathematics, and science subjects separately, classifying students as low or high ability and identifying the top 20 variables that had the most significant impact on the model. The study found five factors in the three subjects in China and the United States. The factors were meta-cognition assess credibility (METASPAM), summarizing texts ability (METASUM), PISA test difficulty perception (PISADIFF), science learning time (SMINS), and school lessons numbers per week (ST060Q01NA). PISADIFF belonged to the student level, and SMINS, ST060Q01NA belonged to the school level. The student-level variables had the most significant influence, and the family-level variables had a relatively small impact.

At the student level, METASPAM, METASUM, and PISADIFF were non-cognitive factors that jointly affect the two countries, illustrating the importance of the learner's self-awareness to personal ability. The findings suggested that persons who are more cautious and serious about their job have a higher personal ability level, implying that work attitude is vital to academic progress (METASPAM). Different summary styles (METASUM) had a significant influence on the three subjects' outcomes. In addition to self-awareness, reading style and interest positively impact student achievement. When reading an article, read it rapidly twice more or only look at the main lines, and then explain it in your own words to help you grasp it better. Discussing the article's topic with others based on a basic grasp is beneficial to a deeper understanding and recall. It may be inferred that a student's capacity to comprehend and recall information (UNDREM) is an important factor in determining their grades. Reading enjoyment (JOYREAD) is an important factor in measuring motivation to read, which is consistent with the outcomes found by Cho and Toste (2018). Individuals who have loads of fun in reading will take the initiative to peruse, while those with a low index think reading is an exercise in futility or just when there are rigid requirements in school. The outcomes show that students with high records have higher scholarly execution at the student level. You can begin by developing the interest and learning motivation to improve students' performance.

At the family level, family socioeconomic status (ESCS) was the only family variable among the 11 key factors that affect all three subjects in China. Family property status (WEALTH) was the only family variable in the United States. Family socioeconomic status has long been a significant focus of educational study over the last 50 years as a critical environmental element impacting kids' learning. Consistent with Chiu and Ming (2017), their family's socioeconomic condition influences students' academic success. According to relevant meta-analysis research, family socioeconomic status (ESCS) was favorably connected with kids' overall academic achievement (White et al., 1993). Children's socioeconomic situation offers them some crucial social impact, which will help them obtain good academic achievements and pave the route for their success. Conger and Donnellan (2007) and Longo and Lombardi (2017) reported that parental investment in resources (such as financial, human, and social capital) aids children's development. A parent's social and economic circumstances influence their emotional relationship and parenting conduct, substantially impacting their children's academic performance and accomplishment level (Chen, 2018; Kriegbaum & Spinath, 2016). In general, parents with a higher socioeconomic standing are more willing and able to spend more on their children materially and emotionally. Emotional investment is quite significant. Their academic performance will increase equally and favorably if parent-child contact within the family is more harmonic and regular (Fan & Chen, 2001). Furthermore, if a student's family's general socioeconomic situation is poor, parents will exhibit some behavioral and emotional abnormalities, preventing or reducing parents' emotional attention and educational investment in their children. The result is a decline in the same level of children's academic performance and achievement (April et al., 2017).

At the school level, the students' school as a social environment also gave rise to specific experiences that positively relate to students' personal ability. 2 of these factors were included in the PISA 2018 survey as part of student well-being assessment (BELONG BEINGBULLIED). The findings reveal that school belonging (BELONG) is also linked to students' academic achievement. Students who belong to a high school are more likely to have excellent grades, and students who belong to a low school are more likely to have poor grades. This is in line with prior research findings. Students' time spent on homework and learning motivation is positively connected with their sense of belonging at school (Hagborg, 1994). Even when ideal learning environments and other characteristics are considered, school belonging is still the most important predictor of academic accomplishment (Gonzalez & Padilla, 1997). School membership is positively connected with academic success (Smerdon, 2002).

The above analysis revealed that students' beliefs, motivation, resources, and environment distinguish the strength of their overall personal ability. Besides, the characteristics of single subjects should not be neglected.

https://doi.org/10.33225/jbse/22.21.575

Many factors do not affect all three subjects but have significant predictive power for the performance of two subjects.

Reading and mathematics are more responsive to failure. When people feel the humiliation and shame generated by failure in an achievement circumstance (GFOFAIL), they tend to avoid self-expression and endeavor. This concept captures the basic emotion of failure anxiety and its behavioral repercussions (Conroy, 2004; McGregor & Elliot, 2005). Failure causes dread because it deprives people of their worth (Conroy, 2001). Fear of failure may have various negative consequences (Martin & Marsh, 2011), including increased anxiety and decreased motivation, and a detrimental influence on academic performance (Martin, 2002). The reason that science subjects are not sensitive to failure may be because learning science encounters setbacks. Learning science subject is a process of repeatedly experiencing failure and continually overcoming it, so the psychological tolerance in the face of failure is large and does not have too serious consequences.

Mathematics is a rational subject that focuses more on theory and thinking, and reading and science learning focus more on practice and thinking about the meaning of life. Individuals' understanding of the connotation of life is referred to as their sense of the meaning of life (EUDMO). One of its purposes is to improve self-control. Individuals with a sense of purpose in life can better think about long-term objectives based on cultural norms and then change their emotions and actions to deal with scholastic assignments, which helps students improve their grades. The findings showed a strong link between learning goals and academic success (Bouffard et al., 1995).

The results revealed that top performers' mathematics and science teachers generally spent a larger proportion of class time demonstrating, explaining, and extending scientific ideas, which involved interaction and whole-class discussion with students. High achievers in mathematics and science literacy generally perceive their teachers as providing considerable feedback or adapting lessons to students' individual needs (PERFEED) (Han & Xu, 2020). However, as the PISA report (OECD 2016) claimed, students typically receive teachers' feedback as a remedial measure to help them become academically stronger. Like Singapore, it outperformed all other participating countries in science literacy on PISA 2015. It is reported to be a country where teachers commonly provide individual feedback to students in the class (Lau & Lam, 2017). Accordingly, a stronger association between this form of teaching strategy and student competency might improve students' mathematics and science scores.

The results of the study showed that the stronger the cooperative and competitive (PERCOOP, PERCOMP) relationships, the better the students' academic achievement. Cooperation and competition among students helped improve individual mathematics and science achievement. Many studies have been conducted abroad and found that cooperative learning, both in problem-solving situations and real-world knowledge, significantly enhances problem-solving skills and academic achievement (Roseth & Johnson, 2008). Consistent with Tauer and Harackiewicz (2004), competition and cooperation have positive aspects. They have the same effect on students' learning by setting up two different cooperative and competitive teaching environments (Lawrence & Sherman, 2006). Cooperative learning helps cultivate students' awareness of equal cooperation and enables students to cooperate, encourage and complement each other as feedback information at different levels to form a multi-directional information exchange. Finally, Clhen and Sapon-shevin (2004) confirm that "individual wisdom" affects "collective wisdom" so that cooperative learning can improve students' learning and social skills. It is also an assistant to teachers in classroom management and teaching. While in today's society, adolescents have an increasingly prominent desire to express themselves. According to psychological studies, adolescent students gradually recognize the relationship between personal and social development under new living and educational conditions. The 'I' and others, the 'main self' and the 'guest self' changed, began thinking of themselves, and really 'discovered' themselves. This has led to a new development of self-awareness and expressive consciousness. This desire for expression is the psychological basis for participation in the competition, making individuals more willing to learn, more interested in learning, more motivated to learn, and more time for education, which positively contributes to academic achievement.

The previous discussion highlighted how specific variables seem to arise in the individual, the home, and the school and the characteristics of the three subjects of reading, mathematics, and science. Due to the lack of some answers in the questionnaire, the variables involved in this study may not be comprehensive. For example, there were fewer responses on ICT in the questionnaire for Chinese students. Therefore, this study further explored the predictor variables that influenced the personal abilities of United States students and analyzed them in comparison with China, resulting in 3 findings worth considering.

First, competitive relationships were critical in today's society. The two predictors of competition (COMPETE and PERCOMP) in the United States and China were similar. Competitive relationships among students were vital

https://doi.org/10.33225/jbse/22.21.575

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

for students' academic performance. However, in contrast to the views of some scholars, who argued that highly competitive learning environments were detrimental to student development, Asian countries and regions tended to "score first," and many educational administrators prioritize competitive management models to achieve higher scores (Hau et al., 2008). The OECD Social and Emotional Competence Assessment 2021 (OECD, 2021) also found that when students are in highly competitive school environments, their ability to communicate with and trust others decreases. Compared to students in OECD countries, students in China perceive higher levels of competition and cooperation in the learning environment and hold higher achievement motivations to pursue success. With the acceleration of the information age in our society and the prevalence of high technology and the Internet, today's society has gradually entered a period where it likes to argue, explore, and debate. Nowadays, students tend to prefer to think independently to get things done. The fierce competition in the general environment of Chinese society has a subtle impact on students. It is highlighted that it drives them to be significantly more motivated to achieve in their pursuit of success. Thus, although a highly competitive learning environment is not conducive to students' psychological well-being, it can significantly motivate students to learn and better prepare them for future social competition. At the same time, moderate "comparison" can help foster students' achievement motivation in pursuit of success and enhance their sense of meaning in life, which seems to be overlooked by most studies.

Second, unlike China, the United States ESCS indicator was only related to mathematics and science and does not significantly influence reading literacy. However, the household property status indicator (WEALTH) impacted all three subjects. Families with wealth accumulation and income levels above the social meant pay more attention and invest more in students' education. It has a certain utility on students' subjective perceptions and judgments of their learning abilities. Then it will significantly impact students' academic performance and achievement levels. Therefore, society and schools should pay more attention to students from low socioeconomic status families, as they are more likely to have low self-efficacy than other students. Improving the quality of parental involvement can compensate for the negative effects of family background disadvantages on students' early development. Families with low socioeconomic status should pay more attention to acquiring parent-child interaction skills and improving the quality of parent-child interaction, which helps children improve their academic self-efficacy and promote students' academic achievement.

Third, Chinese students' academic performance was less correlated with ICT use than in the United States. Only reading literacy was correlated with ICT. The factor is called ICTRES, an indicator of communication information technology resources, ranking 10th in the single subject of reading. The study results showed that for the United States, the more opportunities to use information technology outside of school (AUTICT) to support subject learning or school activities, the better the students' academic performance. However, in China, this is not the case. The reasons for this are as follows. First, for most areas in China, integrating information technology with the curriculum in the school setting has not worked as well as it should. Such as insufficient attention to the information technology courses offered, and insufficient facilities and resources to offer such courses. Second, the design of using information technology to assist students' learning in the out-of-school setting is not enough to truly promote student learning autonomy and diversification.

To this end, integrating ICT with curriculum and teaching is needed to enhance Chinese students' capabilities. First, the quality of IT-integrated courses should be improved. Make subject knowledge relevant to the real world and allow students to explore ICT-based real-life problems. Second, it is essential to enhance ICT use outside of school. For example, as Hu and Gong (2018) observed, assigning relevant homework based on real-life situations and tasks and interacting through feedback on results allow students to gain self-esteem. Finally, creating contexts further enhances students' interest in ICT. Students' interest in ICT significantly predicts academic achievement. However, students' interest in ICT should focus on academically relevant content rather than leisure and entertainment.

Conclusions and Implications

This study analyzed and explored how individual, family, and school-related non-cognitive factors influence secondary school students' personal abilities in 4 Chinese provinces and cities, based on the OECD publicly available PISA 2018 large education data set. It also compares and analyzes the non-cognitive factors affecting Chinese and American students in reading, mathematics, and science, filling the research gap on the influence of non-cognitive factors on different subjects and systematically and comprehensively exploring the importance of non-cognitive factors. The study examined a broader range of non-cognitive factors that may help describe the experiences and attributes of students assessed as having higher reading, mathematics, and science proficiency.

https://doi.org/10.33225/jbse/22.21.575

The study found that three subjects were closely associated with one another, and personal ability depended much on their self-awareness and family income. Regarding subject characteristics, attitude towards failure was the determinant of reading and mathematics. Mathematics and science were improved by student cooperation and competition. The comparison of the two countries revealed the significance of rivalry in contemporary culture. Additionally, family status had little effect on personal ability in the United States. The findings have implications for students, families, and schools regarding the most important factors for students' personal ability and future studies regarding the effect of non-cognitive factors.

Acknowledgments

This study is supported by the National Natural Science Foundation of China (Reference NO.71901195), Hangzhou Philosophy and Social Science Planning Project (Reference NO.M22JC104).

Declaration of Interest

The authors declare no competing interest.

References

Aditomo, A., & Köhler, C. (2020). Do student ratings provide reliable and valid information about teaching quality at the school level? Evaluating measures of science teaching in PISA 2015. *Educational Assessment, Evaluation and Accountability*, *32*(3), 275-310. https://doi.org/10.1007/s11092-020-09328-6

Bailey, M. F., & Borooah, V. K. (2010). What enhances mathematical ability? A cross-country analysis based on test scores of 15-year-olds. *Applied Economics, 42*(29), 3723-3733. https://doi.org/10.1080/00036840802314598

Bellova, R., Melichercikova, D., & Tomcik, P. (2018). Possible reasons for low scientific literacy of Slovak students in some natural science subjects. *Research in Science and Technological Education, 36*(2), 226-242. https://doi.org/10.1080/02635143.2017.1367656

Breiman, L., Friedman, J., Olshen, R., & Stone, C. (2015). Classification and regression trees. Belmont, ca: Wadsworth international group. *Encyclopedia of Ecology, 57*(1), 582-588. https://doi.org/10.1007/978-3-642-57292-0_10

Bouffard, T., Boisvert, J., Vezeau, C., & Larouche, C. (1995). The impact of goal orientation on self-regulation and performance among college students. *British Journal of Educational Psychology, 65*(3), 317-329. https://doi.org/10.1111/j.2044-8279.1995.tb01152.x

Chen, J., Zhang, Y., Wei, Y., & Hu, J. (2021). Discrimination of the contextual features of top performers in scientific literacy using a machine learning approach. *Research in Science Education, 51*, 129-158. https://doi.org/10.1007/s11165-019-9835-y

Conger, R. D., & Donnellan, M. B. (2007). An interactionist perspective on the socioeconomic context of human development. *Social Science Electronic Publishing, 58*(1), 175-199. https://doi.org/10.1146/annurev.psych.58.110405.085551

Conroy, D. E. (2001). Progress in the development of a multidimensional measure of fear of failure: The performance failure appraisal inventory (PFAI). *Anxiety Stress and Coping, 14*(4), 431-452. https://doi.org/10.1080/10615800108248365

Conroy, David, E. (2004). The unique psychological meanings of multidimensional fears of failing. *Journal of Sport and Exercise Psychology, 26*(3), 484-491. https://doi.org/10.1123/jsep.26.3.484

Davier, M. V., Yamamoto, K., Shin, H. J., Chen, H., & Kandathil, M. (2019). Evaluating item response theory linking and model fit for data from PISA 2000-2012. *Assessment in Education Principles Policy and Practice, 26*(4), 466-488. https://doi.org/10.1080/0969594X.2019.1586642

Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher, 44* (4), 237-251. https://doi.org/10.3102/0013189X15584327

Eriksson, K., Lindvall, J., Helenius, O., & Ryve, A. (2021). Socioeconomic status as a multidimensional predictor of student achievement in 77 societies. *Frontiers in Education, 6*. https://doi.org/10.3389/feduc.2021.731634

Flores-Mendoza, C., Ardila, R., Gallegos, M., & Reategui-Colareta, N. (2021). General intelligence and socioeconomic status as strong predictors of student performance in Latin American schools: Evidence from PISA items. *Frontiers in Education, 6*. https://doi.org/10.3389/feduc.2021.632289

Gimenez, G., Ciobanu, D., & Barrado, B. (2021). A proposal of spatial measurement of peer effect through socioeconomic indices and unsatisfied basic needs. *Economies, 9*, 72. https://doi.org/10.3390/economies9020072

Gonzalez, R., & Padilla, A., M. (1997). The academic resilience of Mexican American high school students. *Hispanic Journal of Behavioral Sciences, 19*(3), 1-317. https://doi.org/10.1177/07399863970193004

Hagborg, W., J. (1994). An exploration of school membership among middle- and high-school students. *Journal of Psychoeducational Assessment, 12*(4), 312-323. https://doi.org/10.1177/073428299401200401

Han, Y., Xu, Y. (2020). The development of student feedback literacy: The influences of teacher feedback on peer feedback. *Assessment and Evaluation in Higher Education, 45*(5), 680-696. https://doi.org/10.1080/02602938.2019.1689545

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and non-cognitive abilities on labor market outcomes and social behavior. *NBER Working Papers, 24*(3), 411-482. https://doi.org/10.1086/504455

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Hobri, H., Murtikusuma, R. P., Fatahillah, A., Susanto, & Rini, S. M. (2018). The analysis on critical thinking ability in solving PISA question, and its scaffolding. *Advanced Science Letters, 24*(11), 8215-8218. https://doi.org/10.1166/asl.2018.12526
Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York: Routledge.
Hu, X., Gong, Y., Lai, C., & Leung, F. K. (2018). The relationship between ICT and student literacy in mathematics, reading, and science across 44 countries: A multilevel analysis. *Computers and Education, 125*, 1-13. https://doi.org/10.1016/j.compedu.2018.05.021
Huang, X. H., Hua, L., Zhou, X. Y., Zhang, H., Zhang, M., Wang, S., Qin, S. R., Chen, J., & Wang, X. H. (2021). The association between home environment and quality of life in children and adolescents in Hangzhou city, China. *Journal of Child and Family Studies, 30*(6), 1416-1427. https://doi.org/10.1007/s10826-021-01951-1
Juan, A., & Visser, M. (2017). Home and school environmental determinants of science achievement of South African students. *South African Journal of Education, 37*(1), 1-10. https://doi.org/10.15700/saje.v37n1a1292
Kriegbaum, K., & Spinath, B. (2016). Explaining social disparities in mathematical achievement: The role of motivation. *European Journal of Personality, 30*(1), 45-63. https://doi.org/10.1002/per.2042
Krieken, K. V., Sanders, J., & Hoeken, H. (2015). Viewpoint representation in journalistic crime narratives: An analysis of grammatical roles and referential expressions. *Journal of Pragmatics*, 220-230. https://doi.org/10.1016/j.pragma.2014.07.012
Laftman, S. B., Ostberg, V., & Modin, B. (2017). School climate and exposure to bullying: A multilevel study. *School Effectiveness and School Improvement, 28*(1), 153-164. https://doi.org/10.1080/09243453.2016.1253591
Lau, K. C., & Lam, T. Y. P. (2017). Instructional practices and science performance of 10 top-performing regions in PISA 2015. *International Journal of Science Education, 39*(15), 2128-2149. https://doi.org/10.1080/09500693.2017.1387947
Lee, J., & Shute, V. J. (2010). Personal and social-contextual factors in K-12 academic performance: An integrative perspective on student learning. *Educational Psychologist, 45*(3), 185-202. https://doi.org/10.1080/00461520.2010.493471
Li, L., Peng, Z., Lu, L., Liao, H., & Li, H. (2020). Peer relationships, self-efficacy, academic motivation, and mathematics achievement in Zhuang adolescents: A moderated mediation model. *Children and Youth Services Review*, 105358. https://doi.org/10.1016/j.childyouth.2020.105358
Liu, R., Sun, L., Jing, Y., & Bradley, K. (2017). Using the 2006 PISA questionnaire to evaluate the measure of educational resources: A Rasch measurement approach. *International Journal of Assessment Tools in Education, 4*(2), 211-211. https://doi.org/10.21449/ijate.319486
Longo, F., Lombardi, M. P., & Dearing, E. (2017). Family investments in low-income children's achievement and socioemotional functioning. *Developmental Psychology, 53*(12), 2273-2289. https://doi.org/10.1037/dev0000366
Martin, A. J., & Marsh, H. W. (2011). Fear of failure: Friend or foe? *Australian Psychologist, 38*(1), 31-38. https://doi.org/10.1080/00050060310001706997
Martin, A. J., & Dowson, M. (2009). Interpersonal relationships, motivation, engagement, and achievement: Yields for theory, current issues, and educational practice. *Review of Educational Research, 79*, 327-365. https://doi.org/10.3102/0034654308325583
Martin, A. (2002). Motivation and academic resilience: Developing a model for student enhancement. *Australian Journal of Education, 46*(1), 34-49. https://doi.org/10.1177/000494410204600104
Masarik, A. S., & Conger, R. D. (2017). Stress and child development: A review of the family stress model. *Current Opinion in Psychology, 13*, 85-90. https://doi.org/10.1016/j.copsyc.2016.05.008
McGregor, & H., A. (2005). The shame of failure: Examining the link between fear of failure and shame. *Pers Soc Psychol Bull, 31*(2), 218-231. https://doi.org/10.1177/0146167204271420
Muelle, L. (2020). Socioeconomic and contextual factors associated with low academic performance of Peruvian students in PISA 2015. *Apuntes-Revista De Ciencias Sociales, 47*(86), 111-146. https://doi.org/10.21678/apuntes.86.943
Ngorosho, D. (2010). Reading and writing ability in relation to home environment: A study in primary education in rural Tanzania. *Child Indicators Research, 4*(3), 369-388. https://doi.org//10.1007/s12187-010-9089-8
OECD. PISA 2018 DataBase. Available online: https://www. Oecd. Org/PISA/data/2018database/ (accessed on 30 June 2021)
Roseth, C. J., Johnson, D. W., & Johnson, R. T. (2008). Promoting early adolescents' achievement and peer relationships. *Psychological Bulletin, 134*(2), 223-246. https://doi.org/10.1037/0033-2909.134.2.223
Saeki, E., & Quirk, M. (2015). Getting students engaged might not be enough: The importance of psychological needs satisfaction on social-emotional and behavioral functioning among early adolescents. *Social Psychology of Education, 18*(2), 355-371. https://doi.org/10.1007/s11218-014-9283-5
Salchegger, S. (2016). Selective school systems and academic self-concept: How explicit and implicit school-level tracking relate to the big-fish-little-pond effect across cultures. *Journal of Educational Psychology, 108*(3), 405-423. https://doi.org/10.1037/edu0000063
Sherman, L. W. (2010). A comparative study of cooperative and competitive achievement in two secondary biology classrooms: The group investigation model versus an individually competitive goal structure. *Journal of Research in Science Teaching, 26*(1), 55-64. https://doi.org/10.1002/tea.3660260106
Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods, 14*(4), 323-348. https://doi.org/10.1037/a0016973
Sun-Lin, H. Z., & Chiou, G. F. (2017). Effects of comparison and game-challenge on sixth graders' algebra variable learning achievement, learning attitude, and meta-cognitive awareness. *Eurasia Journal of Mathematics Science and Technology Education, 13*(6), 2627-644. https://doi.org/10.12973/eurasia.2017.01244a
Tauer, J. M., & Harackiewicz, J. M. (2004). The effects of cooperation and competition on intrinsic motivation and performance. *Journal of Personality and Social Psychology, 86*(6), 849-861. https://doi.org/10.1037/0022-3514.86.6.849

https://doi.org/10.33225/jbse/22.21.575

Tseng, J. C. R., Chu, H. C., Hwang, G. H., & Tsai, C. C. (2008). Development of an adaptive learning system with two sources of personalization information. *Computers and Education, 51* (2), 776-786. https://doi.org/10.1016/j.compedu.2007.08.002

Tonga, F. E., Eryiğit, S., Yalçın, F. A., & Erden, F. T. (2019). Professional development of teachers in PISA achiever countries: Finland, Estonia, Japan, Singapore and China. *Professional Development in Education, 48*(1), 88-104. https://doi.org/10.1080/19415257.2019.1689521

Toste, J. R., Didion, L., Peng, P., Filderman, M. J., & Mcclelland, A. M. (2020). A meta-analytic review of the relations between motivation and reading achievement for K-12 students. *Review of Educational Research, 90*(3), 420-456. https://doi.org/10.3102/0034654320919352

United States Department of Education. (2014). 10 facts about K-12 education funding. Us Department of Education. http://www.ed.gov/about/overview/fed/10facts/index.html

White, S. B., Reynolds, P. D., Thomas, M. M., & Gitzlaff, N. J. (1993). Socioeconomic status and achievement revisited. *Urban Education, 28*(3), 328-343. https://doi.org/10.1177/0042085993028003007

Woods-McConney, A., Oliver, M. C., McConney, A., Maor, D., & Schibeci, R. (2013). Science engagement and literacy: A retrospective analysis for indigenous and non-indigenous students in Aotearoa New Zealand and Australia. *Research in Science Education, 43*(1), 233-252. https://doi.org/10.1007/s11165-011-9265-y

Woods-McConney, A., Oliver, M. C., McConney, A., Schibeci, R., & Maor, D. (2014). Science engagement and literacy: A retrospective analysis for students in Canada and Australia. *International Journal of Science Education, 36*(10), 1588-1608. https://doi.org/10.1080/09500693.2013.871658

Xiao, Y., & Hu, J. (2019). Assessment of optimal pedagogical factors for Canadian ESL learners' reading literacy through artificial intelligence algorithms. *International Journal of English Linguistics, 9*(4), 1-14. https://doi.org/10.5539/IJEL.V9N4P1

Yu, L., & Reichle, E. D. (2017). Chinese versus English: Insights on cognition during reading. *Trends in Cognitive Sciences, 21*(10), 721-724. https://doi.org/10.1016/j.tics.2017.06.004

Zhao, Y., Hoge, J. D., Choi, J., & Lee, S. Y. (2006). Comparison of social studies education in the United States, China, and South Korea. *International Journal of Social Education, 21*(2), 91-122. http://www.bsu.edu/classes/cantu/journal.htm

Appendix

Table 6
Features with Questions Included and Value Scale

Feature	Question	Value scale
METASPAM		
METASUM	1. I write a summary. Then I check that each paragraph is covered in the summary, because the content of each paragraph should be included. 2. I try to copy out accurately as many sentences as possible. 3. Before writing the summary, I read the text as many times as possible. 4. I carefully check whether the most important facts in the text are represented in the summary. 5. I read through the text, underlining the most important sentences. Then I write them in my own words as a summary.	1=Not useful at all 6=Very useful
UNDREM	1. I concentrate on the parts of the text that are easy to understand. 2. I quickly read through the text twice. 3. After reading the text, I discuss its content with others. 4. I underline important parts of the text. 5. I summarize the text in my own words. 6. I read the text aloud to another person.	1=Not useful at all 6=Very useful
JOYREAD	1. I read only if I have to. 2. Reading is one of my favorite hobbies. 3. I like talking about books with other people. 4. For me, reading is a waste of time. 5. I read only to get information that I need.	1=Strongly Disagree 2=Disagree 3=Agree 4=Strongly Agree
PISADIFF	1. There were many words I could not understand. 2. Many texts were too difficult for me. 3. I was lost when I had to navigate between different pages.	1=Strongly disagree 2=Disagree 3=Agree 4=Strongly Agree
ST059Q03TA	How many class periods per week are you typically required to attend for the following subjects?	
ST060Q01NA	In a normal, full week at school, how many class periods are you required to attend in total?	

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Feature	Question	Value scale
BELONG	1. I feel like an outsider (or left out of things) at school. 2. I make friends easily at school. 3. I feel like I belong at school. 4. I feel awkward and out of place in my school. 5. Other students seem to like me. 6. I feel lonely at school.	1=Strongly disagree 2=Disagree 3=Agree 4=Strongly Agree
AUTICT	1. Browsing the Internet for schoolwork. 2. Browsing the Internet to follow up lessons. 3. Using email for communication with other students about schoolwork. 4. Using email for communication with teachers and submission of homework or other schoolwork. 5. Using social networks for communication with other students about schoolwork or teachers. 6. Downloading, uploading, or browsing material from my school's website (e.g., timetable or course materials). 7. Checking the school's website for announcements. 8. Doing homework/using learning apps/learning websites on a computer or a mobile device.	1=Never or hardly ever 2=Once or twice a month 3=Once or twice a week 4=Almost every day 5=Every day
COMPICT	1. I forget about time when I'm using digital devices. 2. The Internet is a great resource for obtaining information. 3. It is very useful to have social networks on the Internet. 4. I am really excited discovering new digital devices or applications. 5. I really feel bad if no Internet connection is possible. 6. I like using digital devices.	1=Strongly disagree 2=Disagree 3=Agree 4=Strongly Agree
COMPETE	1. I enjoy working in situations involving competition with others. 2. It is important to perform better than others on a task. 3. I try harder when I'm in competition with other people.	1=Strongly disagree 2=Disagree 3=Agree 4=Strongly Agree

Table 7
Features Reflecting Subject Characteristics with Questions Included and Value Scale

Feature	Question	Value scale
GFOFAIL	1. When I am failing, I worry about what others think of me. 2. When I am failing, I am afraid that I might not have enough talent. 3. When I am failing, this makes me doubt my plans for the future.	1=Strongly disagree 2=Disagree 3=Agree 4=Strongly Agree
EUDMO	1. My life has clear meaning or purpose. 2. I have discovered a satisfactory meaning in life. 3. I have a clear sense of what gives meaning to my life.	1=Strongly disagree 2=Disagree 3=Agree 4=Strongly Agree
PERFEED	1. The teacher gives me feedback on my strengths in this subject. 2. The teacher tells me in which areas I can still improve. 3. The teacher tells me how I can improve my performance.	1=Never or almost never 2=Some lessons 3=Many lessons 4=Every lesson or almost every lesson
HEDRES	1. A desk to study at 2. A room of your own 3. A quiet place to study 4. A computer you can use for schoolwork 5. Educational software 6. A link to the Internet 7. Classic literature (e.g., Shakespeare) 8. Books of poetry 9. Works of art (e.g., paintings) 10. Books to help with your schoolwork 11. Technical reference books 12. A dictionary 13. Books on art, music, or design	1=Yes 2=No

https://doi.org/10.33225/jbse/22.21.575

Feature	Question	Value scale
PERCOOP	1. Students seem to value cooperation. 2. It seems that students are cooperating with each other. 3. Students seem to share the feeling that cooperating with each other is important. 4. Students feel that they are encouraged to cooperate with others.	1=Not at all true 2=Slightly true 3=Very true 4=Extremely true
PERCOMP	1. Students seem to value competition. 2. It seems that students are competing with each other. 3. Students seem to share the feeling that competing with each other is important. 4. Students feel that they are being compared with others.	1=Not at all true 2=Slightly true 3=Very true 4=Extremely true
SCREADCOMP	1. I am a good reader. 2. I am able to understand difficult texts. 3. I read fluently. 4. I have always had difficulty with reading. 5. I have to read a text several times before completely understanding it. 6. I find it difficult to answer questions about a text.	1=Strongly disagree 2=Disagree 3=Agree 4=Strongly Agree
ST059Q04HA	How many class periods per week are you typically required to attend for the following subjects?	
ST185Q01HA	My life has clear meaning or purpose.	
HOMEPOS	1. Televisions 2. Cars 3. Rooms with a bath or shower 4. Cell phones with Internet access (e.g., smartphones) 5. Computers (e.g., desktop computer, portable laptop, or notebook) 6. Tablet computers (e.g., iPad, BlackBerry, PlayBookTM) 7. E-book readers (e.g., KindleTM, Kobo, Bookeen) 8. Musical instruments (e.g., guitar, piano)	1=None 2=One 3=Two 4=Three or more
ICTOUTSIDE	1. Test language lessons 2. Mathematics 3. Science 4. Foreign language 5. Social sciences 6. Music 7. Sports 8. Performing arts 9. Visual arts	1=No time 2=1-30 minutes a week 3=31-60 minutes a week 4=More than 60 minutes a week 5=I do not study this subject

Received: *June 14, 2022* Revised: *July 20, 2022* Accepted: *August 15, 2022*

Cite as: Ye, L., & Yuan, Y. (2022). Using a machine learning approach to explore non-cognitive factors affecting reading, mathematics, and science literacy in China and the United States. *Journal of Baltic Science Education*, *21*(4), 575-593. https://doi.org/10.33225/jbse/22.21.575

Lu Ye Research Center for Ecological Civilization, College of Economics and Management, Zhejiang Sci-Tech University, Hangzhou, 310018, China.
E-mail: zjwzajyl@hotmail.com
ORCID: https://orcid.org/0000-0001-7681-2708

Yuqing Yuan *(Corresponding author)* Research Center for Ecological Civilization, College of Economics and Management, Zhejiang Sci-Tech University, Hangzhou, 310018, China.
E-mail: eureka346@hotmail.com
ORCID: https://orcid.org/0000-0003-1701-4698