Thai Journal of **Math**ematics
Volume 22 Number 2 (2024)
Pages 411–439

http://thaijmath.in.cmu.ac.th

ISSN 1686-0209

Understanding Characteristics of Thai Students through PISA 2018 Data Using Data Visualization Techniques

Nirattaya Khamsemanan and Cholwich Nattee*

Sirindhorn International Institute of Technology, Thammasat University, Pathum Thani 12120, Thailand
e-mail : nirattaya@siit.tu.ac.th (N. Khamsemanan); cholwich@siit.tu.ac.th (C. Natee)

Abstract Education is the cornerstone of a nation's development, economy, and competitiveness. A clear and deep understanding of the current educational situation is needed. The Programme for International Student Assessment (PISA) has been widely used as an indicator of a nation's school system. This study is conducted using a topological data analysis technique called UMAP to obtain a deeper understanding of Thai students based on PISA 2018 data. For each subject—mathematics, reading, and science—students's performances are preprocessed into high-dimensional vectors. Our proposed method reduces the high-dimensional vectors into two dimensions. The results show deeper insights beyond mere average scores, which are unattainable through other techniques. For example, although the Thai students' average scores are below the OECD averages, their performance across various skills was not uniformly poor. Thai students excelled in the interpretation skill in both mathematics and science compared to others. In reading, Thai students performed well in the cognitive process of locating information but struggled with the process of gaining understanding. Students with higher scores were found to read more books, have higher expectations of completing a bachelor's degree, and experience less bullying. These insights are valuable for policymakers and educators aiming to improve the Thai education system.

MSC: 62R40; 68T99

Keywords: PISA; Thailand; UMAP; data visualization; education; topological data analysis

Submission date: 13.06.2023 / Acceptance date: 09.05.2024

1. Introduction

It is widely accepted that education is one of the core resources for national development, and economic competitiveness [1, 22]. Authorities and policymakers around the globe put much effort into improving and elevating the education qualities of their nations. One of the key aspects needed for improving the educational quality of a nation is understanding the current educational situation. Many international organizations have developed assessments that yield educational qualities and rank the relative educational

*Corresponding author.

Published by The Mathematical Association of Thailand.

performances. Among these assessments, the Programme for International Student Assessment (PISA) (https://www.oecd.org/pisa/) is one of the most prominent tools, widely used worldwide.

PISA is a recurring knowledge assessment that takes place every three years. The Organisation for Economic Co-Operation and Development (OECD) first launched PISA in 1997 to test skills and knowledge in mathematics, reading, and sciences of fifteen years old students in compulsory education. PISA also includes questions related to the social and economic backgrounds of students. PISA cognitive items and questionnaires are written in students' native languages, to eliminate language barriers. In the seventh cycle of PISA, 79 nations participated in PISA 2018, with more than half being non-OECD countries. Samples of students who participated in PISA exam were carefully selected according to strict guidelines. PISA results are publicly available on the OECD's website. Moreover, the objectives of PISA are aligned with the United Nations' Sustainable Development Goals [24].

PISA has become an international standard indicator of the school system and progress [5]. PISA results have influenced education reforms in many countries [1, 6, 24]. Many researchers and policymakers have used PISA results to study global education trends [9], along with various other aspects and relationships.

Many studies, such as [6, 12, 20, 23], use PISA data to study gender differences and gender-equality effects in mathematics and reading performances. In most PISA participating nations, male students perform better in mathematics, but female students perform better in reading [20, 23] and this trend has been steady for over a decade. The gender-different performances correlate with the nation's gender equality policies. However, a study by [12] shows that both male and female students' performances increase with an improvement of gender equality policy but the differences remain. The debate on the issue of gender differences and gender equality is ongoing.

Studies, such as [10, 11, 15, 19, 21], explore the relationships between personal factors, socio-economic background, school and teacher factors, and academic performance from PISA data. Most studies show that students from better socio-economic backgrounds tend to perform better in the PISA exam. Though a study from Neuman [15], using methods from network theory, shows that performances of students from the less privileged socio-economic background from countries such as China, Spain and Portugal are above the worldwide average. A study by Gutiérrez-de-Rozas et al. [10], using classification and regression trees, reveals that variables that are strongly associated with low performances are mainly from students. Family factors have a slightly less explanatory capacity for low performances than personal factors. Variables associated with teachers and schools show low to moderate effects on low performances. A study by Gutiérrez et al. [11] shows a positive correlation between self-concept and the feeling of belonging to a school.

Many studies use PISA as a tool to gain a deeper understanding of a nation's education system and its characteristics. Bernardo et al. [4] use machine learning techniques to extract 20 factors, e.g. socio-economic constraints, learning motivations and mindsets, social experiences in the school environment etc., that have the highest impacts on reading proficiency in the Philippines. A study by Gorostiaga et al. [8] uses Support Vector Machine (SVM) to analyze PISA results in Spain. Their results show that high performances are related to regional variables, computer availability, gender, etc. Dong and Hu [7], using SVM, show that the index of economic, social and cultural status, learning time, school size, home possessions, and family wealth are among the highest factor contributing to high reading scores in Singapore.

From PISA 2018 results [2], average scores of Thai students in all three subjects, reading, mathematics and science, are below OECD average. In reading, 40% of Thai students reached at least level 2 proficiency in reading, whereas the OECD average is 77%. In mathematics, 47% of Thai students reached at least level 2 proficiency in mathematics, whereas the OECD average is 76%. And in science, 56% of Thai students reached at least level 2 proficiency in science, whereas the OECD average is 78%. On average, girls obtained significantly higher scores in all three subjects; 16 points in mathematics, 39 points in reading, and 20 points in science.

The objective of this study is to gain more insights and a deeper understanding of the Thai compulsory educational system from PISA 2018 results using our newly proposed visualization technique based on the Uniform Manifold Approximation and Projection (UMAP) [3, 13, 14]. UMAP is a topological data analysis tool, typically used for dimension reduction and visualizations. UMAP projects high-dimensional space onto low-dimensional space while preserving local structures as well as maintaining global structures.

In this work, we propose a preprocessing technique to represent the performance of a student as a data point in a high-dimensional space. Our proposed representation is based on assessment aspects of each subject; 22-dimensional space for mathematics based on an average score and a standard deviation of 11 assessment aspects shown in Figure 1, 42-dimensional space for reading based on an average score and a standard deviation of 21 assessment aspects shown in Figure 2, and 40-dimensional space for science based on an average score and a standard deviation of 20 assessment aspects shown in Figure 3. These three high-dimensional spaces are mapped onto two-dimensional spaces using UMAP. Each point in these 2D visualizations represents a performance of a Thai student who took the PISA 2018 exam. Various highlights and colors are used to explore the ranges of performance levels of each assessment aspect and the influences of different socio-economic on performances.

Unlike traditional statistical measurements or basic machine learning technique such as SVM, where performances and socio-economic factors of individual students are integrated into the trends of the samples, the proposed technique shows visualizations of individual students while preserving local and global structures. To the best of our knowledge, such analysis has not been done before with PISA data. The proposed technique provides insights and a deeper understanding of Thai education based on PISA 2018 data. We hope that these insights help Thai policymakers and educators in improving Thai education.

2. Materials and Methods

2.1. Data and Data Preprocessing

The data used in this study are from the OECD PISA 2018 database. They are publicly available at https://www.oecd.org/pisa/data/2018database. Since this study aims to investigate Thai students' characteristics from how the students perform in the PISA 2018 test, we extract 8,633 result records of Thai students from the cognitive item data file. Each record represents one student. These students from 290 schools are sampled from 575,713 15-year-old Thai students using a two-stage stratified random selection process

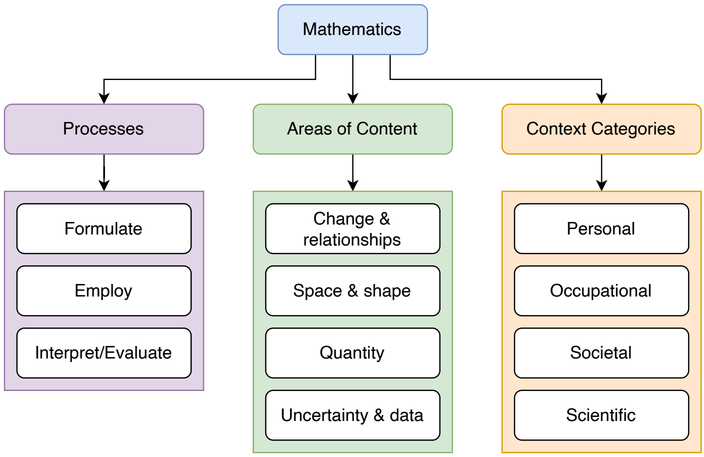


Figure 1. PISA 2018 Mathematics Framework

[2]. After the extraction process, each record provides timing, raw responses, scored responses, and coded responses to the items that each student answered.

According to the PISA 2018 assessment and analytical framework [17], each item is categorized based on several assessment aspects. Figure 1 shows the framework for mathematics items. Each mathematics item is designed to assess students from three interrelated aspects: mathematical processes, content, and contexts. Each mathematics item is therefore labeled by a mathematical process, an area of content, and a context category. This information is provided as the item pool classification, which is Annex A of the PISA 2018 technical report [16]. Table 1 shows examples of mathematics items with the classification. Similarly, the PISA 2018 reading and science frameworks are shown in Figure 2 and Figure 3, respectively. All reading and science items are labeled according to the frameworks.

We aim to apply a data visualization technique to construct graphical representations of the cognitive item data. To conduct the visualization, all data points must be represented by the same set of attributes, enabling direct comparison between any two records. However, students are randomly assigned to do different clusters of items according to the field trial design of PISA 2018 [18]. Consequently, if we randomly select two records from the dataset, they may not be directly comparable due to students being assigned to different clusters.

To overcome this problem, we propose a technique to preprocess the cognitive item data of Thai students by using the assessment aspects of the items and the coded/scored responses of the items to represent the characteristics of a student. When a student receives either a coded or a scored response from an item, the response shows the student's characteristics due to the item's assessment aspects. We then pair the student's received response with each assessment aspect of the item. For example, suppose a student is assigned to the "Cash Withdrawal" mathematics unit composed of two items (CM496Q01S and CM496Q02S), and the student received the scored response of 1 from the CM496Q01S item. Since the CM496Q01S item is categorized under the *formulate* process, we then use this information to represent the student's characteristics in the *formulate* process.

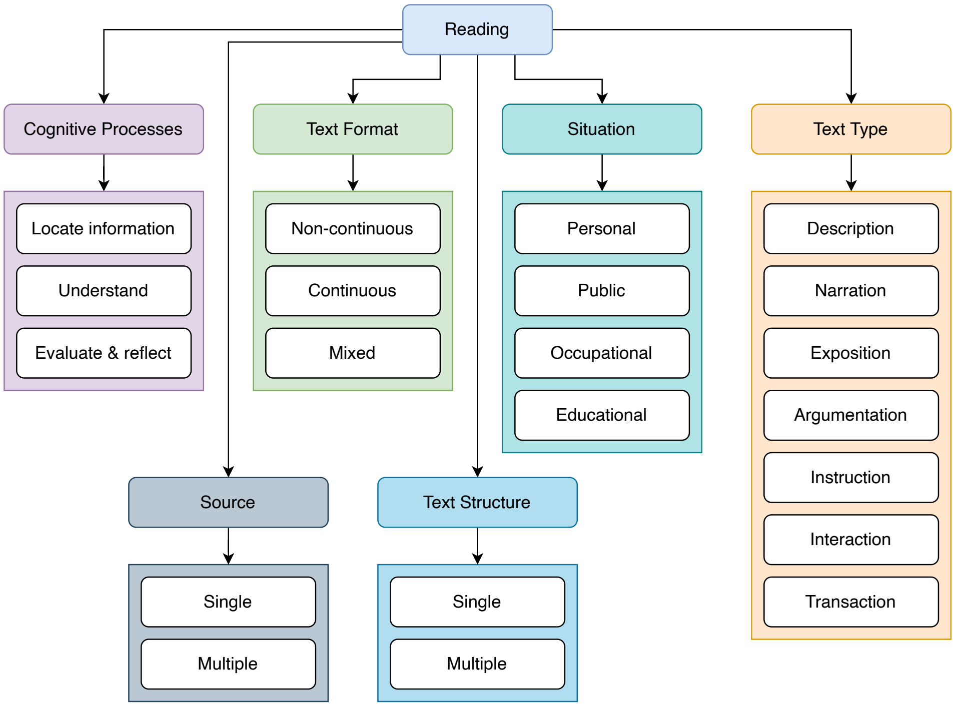


FIGURE 2. PISA 2018 Reading Framework

TABLE 1. Examples of Mathematics Item Classication

Unit Name: Population Pyramids

Item ID	**Process**	**Content**	**Context**
DM155Q01S	Employ	Change & relationships	Scientific
DM155Q02C	Interpret/Evaluate	Change & relationships	Scientific
DM155Q03C	Employ	Change & relationships	Scientific
DM155Q04S	Interpret/Evaluate	Change & relationships	Scientific

Unit Name: Cash Withdrawal

Item ID	**Process**	**Content**	**Context**
CM496Q01S	Formulate	Quantity	Societal
CM496Q02S	Employ	Quantity	Societal

According to the test design, a student is assigned to do multiple items. We pair the assessment aspects of all items to all the responses that the student received. Then, we group all responses of the same assessment aspect and use the average response with the standard deviation of the responses to represent the characteristic according to the assessment aspect of the student. Table 2 shows the part of the table for storing pre-processed students' characteristics based on mathematics items. Each record in the table represents a student's characteristics with the same set of attributes. Therefore, any two students' characteristics can be directly compared. We perform the same preprocessing

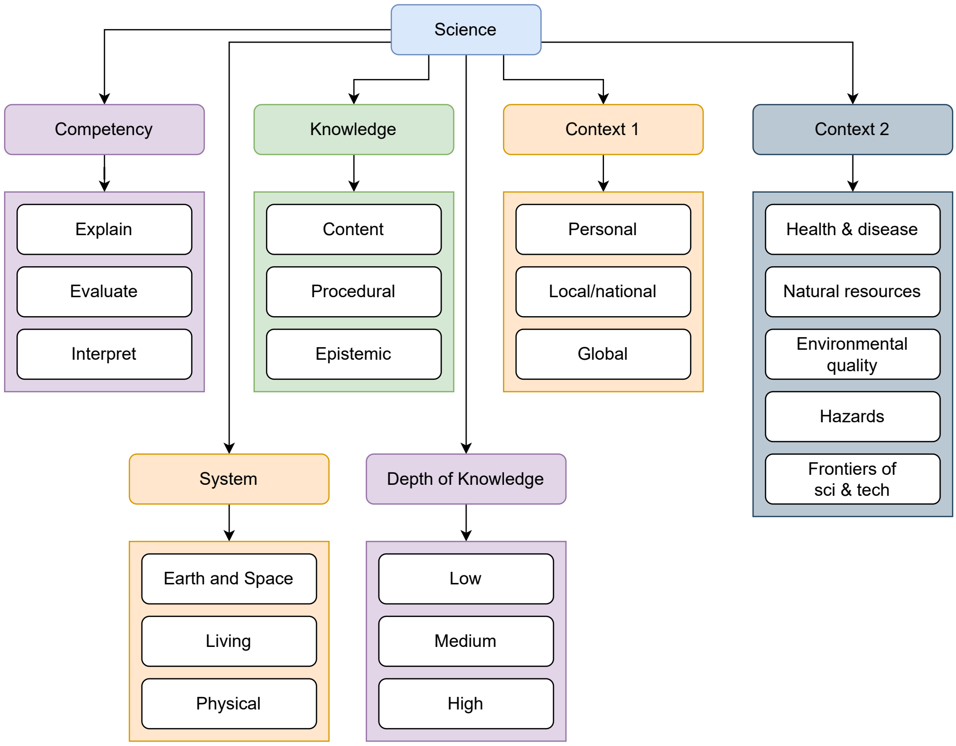


Figure 3. PISA 2018 Science Framework

Table 2. Partial structure of the table for storing preprocessed students' characteristics based on mathematics items

Student	Process						Content				
	Formulate		Employ		Interpret/Eval		Quantity		Space/shape		...
	Avg	SD	Avg	SD	Avg	SD	Avg	SD	Avg	SD	

steps for both reading and science items. Thus, three tables of students' characteristics are obtained after the preprocessing steps.

We summarize our proposed steps to preprocess data as follows.

(1) Extract Thai students' cognitive data from the PISA 2018 cognitive data file.

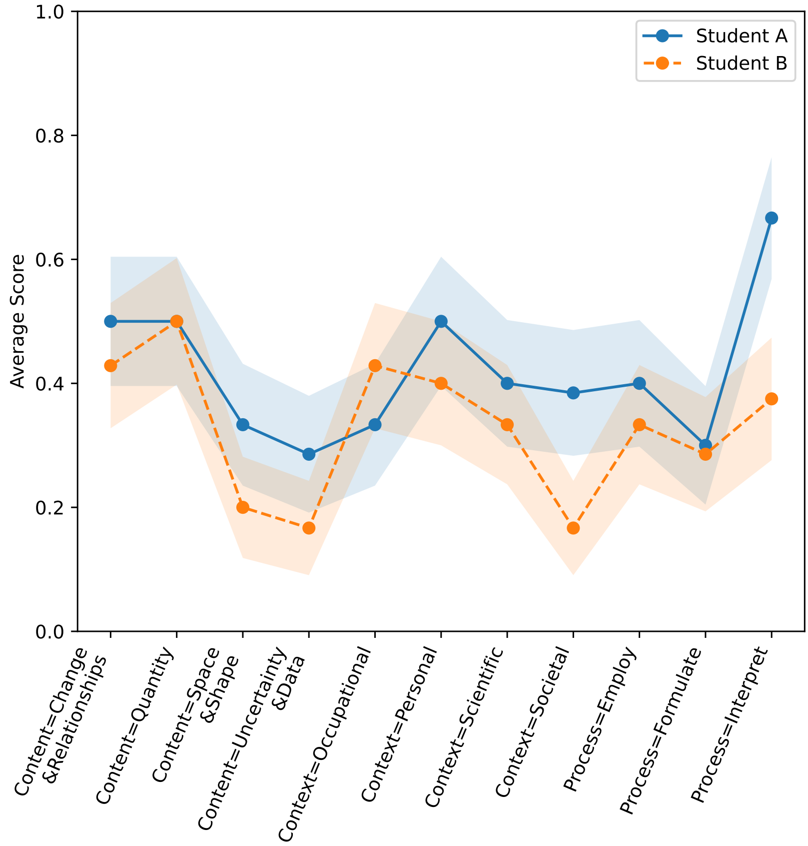


FIGURE 4. Visualization of two preprocessed student records extracted from their performances on mathematics items

(2) Retrieve each student's assigned items and coded/scored responses of those items.
(3) Join each item ID with its assessment aspects from Annex A.
(4) Aggregate the coded/scored responses of each assessment aspect and calculate the average and the standard deviation for each assessment aspect.

Figure 4 visualizes two examples of student records preprocessed according to the proposed steps. The horizontal axis shows the assessment aspects and the vertical axis shows the scores/responses. Each marker shows the average response for each assessment aspect. The shaded area shows the standard deviation. It can be seen that the proposed steps allow us to directly compare two student records based on the assessment aspects. For example, Student A performs much better than Student B in the items requiring the 'interpret' process. But, Student B performs better than Student A in the items in the 'occupational' context.

We conduct the preprocessing steps three times for each subject i.e. mathematics, science, and reading.

2.2. Data Visualization

From the proposed preprocessing steps explained in the previous section, we can represent the performance of each student in a subject as a fixed-size vector. These vectors can be considered a point in a high-dimensional space where the number of dimensions equals twice (average and standard deviation) the number of assessment aspects in each subject (22 dimensions for mathematics, 42 dimensions for reading and 40 dimensions for science). To study the performance-based characteristics of Thai students from the data, we propose to apply data visualization techniques to construct a graphical representation of the dataset. Since each data point is represented in a high-dimensional space, the data cannot easily be visualized as a two-dimensional or three-dimensional plot.

To visualize the data, we first apply a dimensionality reduction technique to map the dataset from a high-dimensional space into a two-dimensional space. Then, we plot the transformed data using a scatter plot where each data point is shown as a point in a two-dimensional space. From the plot, we can study relationships among students by comparing the closeness of plotted points, or study clusters of students with particular characteristics. To properly visualize the dataset, a dimensionality reduction technique has to preserve the shape and structure of the original dataset. Therefore, we propose to apply a non-linear dimensionality reduction technique called Uniform Manifold Approximation and Projection (UMAP) [3, 13, 14]. UMAP is a technique based on topological data analysis and manifold learning. It models the data with a fuzzy topological structure using an assumption that the data are uniformly distributed on Reimannian manifold. The UMAP technique then projects the data into a low-dimensional space that has the closest equivalent fuzzy topological structure. Since the preprocessed data are in $\mathbb{R}^n$ where n is the number of dimensions, it is a Reimannian manifold. UMAP can therefore be applied to project the preprocessed data into a two-dimensional space.

As stated above, a scatter plot shows how data points in the dataset are arranged or form clusters. To link the data points to the assessment aspects, each point is colored by an assessment value of the student represented by the point. Since each student was sampled from the entire population, each student, therefore, represents a group of students according to the sampling method. This information is provided via a variable named `W_FSTUWT`. We set the size of each point to make its area reflect the value of `W_FSTUWT`. Figure 5 shows a scatter plot constructed from the mathematics cognitive data of Thai students. Each point is colored by the average scores of all mathematics items where a red point shows a low mathematics score (close to 0%), and a green point shows a high mathematics score (close to 100%). The color bar on the right side of the figure shows a relationship between colors and values.

Since a colored scatter plot can only show one value related to each point, we propose to use multiple scatter plots colored with different values related to the points. We place these plots next to one another.

We also use another color scheme (pink to blue) for students' responses from PISA questionnaires (socio-economic factors).

By doing so, we can compare socio-economic factors, assessment asspects and overall performances related to one point (one student). Figure 6 shows three scatter plots created from the mathematics cognitive data. Three different values based on the mathematics process used in the items are used to color the plots. It can be easily seen from the figure that Thai students perform not so well in the items requiring *formulate* process since most of the points are red. The visualization technique based on multiple scatter plots

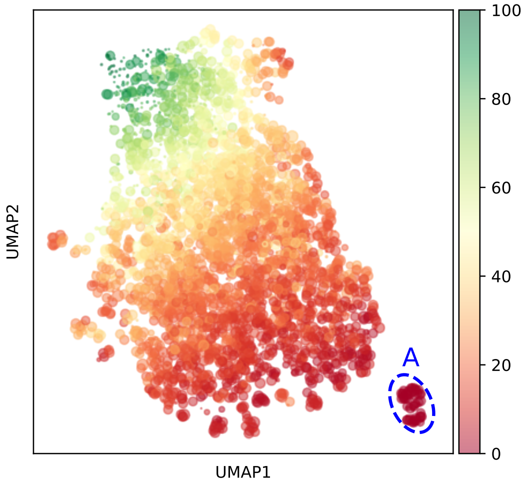


FIGURE 5. A scatter plot showing the overall performance of Thai students based on mathematics items. A point at the same location in Figure 5–8 represents the same student.

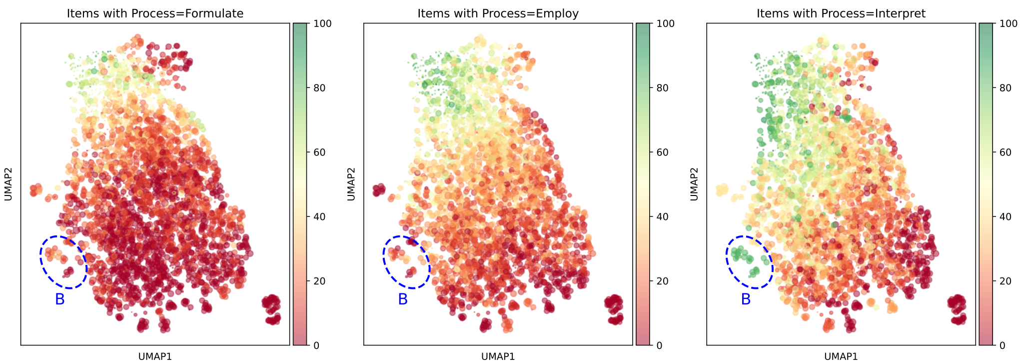


FIGURE 6. A group of three scatter plots shows the average responses of each mathematical process. A point at the same location in Figure 5–8 represents the same student.

allows us to easily compare different values related to the same student. We will discuss the results in detail in the next section.

3. Results and Discussions

We create a web-based application that shows all the visualization results. It is available at `https://gaittech-siit-tu.github.io/pisa2018-visualization/`.

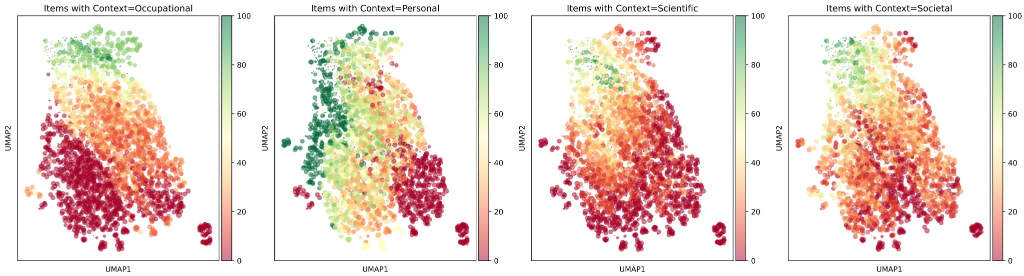


Figure 7. A group of scatter plots shows the average responses of each type of item context. A point at the same location in Figure 5–8 represents the same student.

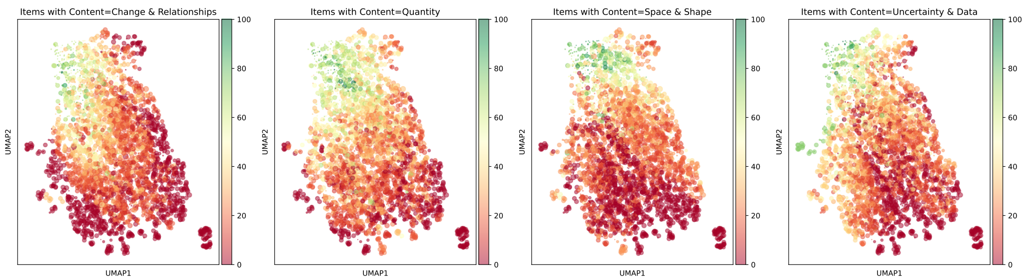


Figure 8. A group of scatter plots shows the average responses of each type of item content. A point at the same location in Figure 5–8 represents the same student.

3.1. Visualizing Mathematics Cognitive Data

Figure 5 shows a graphical visualization, from the proposed method, generated from the mathematics cognitive data of Thai students. Each point representing a student is highlighted by the overall average response of its corresponding student. The result shows that most students are grouped into one main cluster and a few small clusters. The cluster labeled 'A' shows a group of students who answered all mathematics items wrong. This cluster is clearly separated from the main cluster. Figure 6, 7 and 8 show graphical visualizations highlighted by student performances based on assessment aspects, i.e. mathematical process, item context, and item content, respectively.

Here are some performance-based characteristics that we can observe from the results:

- The top-left corner of the main cluster in Figure 5 shows that a small number of Thai students performed well in mathematics items with an average response of at least 70%. However, a large number of Thai students received an average response of below 50%.
- Figure 6 shows that Thai students performed better for problems that required *interpret* skill even students who did not receive overall high scores. Thai students did not do well with items that required *formulate* skill even students who received moderate overall scores. Students who did well on the items with the *employ* process tends to do well on the items with the *interpret* process. This can be

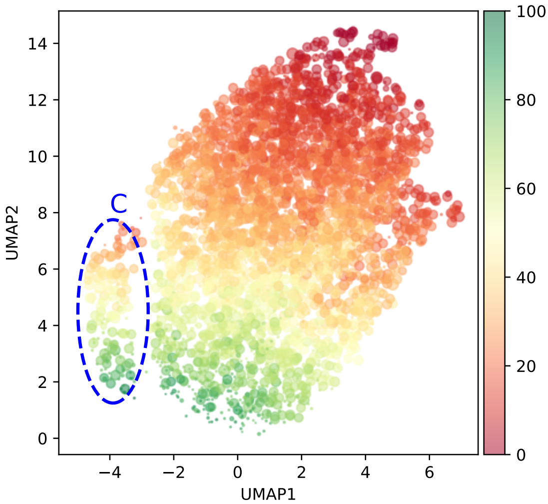


FIGURE 9. A scatter plot showing the overall performance of Thai students based on science items. A point at the same location in Figure 9–13 represents the same student.

observed from the green points. An exception to this is the student in Cluster B in Figure 6. Cluster B shows a group of students who performed well in the items requiring *interpret* process, but not well in the other two processes.

- From Figure 7, it can be seen that different groups of students performed well in different contexts. A large number of students, who did not recieve high overall scores, performed well on the items with *personal* context. This can be seen from the left side of the main cluster highlighted with dark green. However, a few students performed well on the items with scientific context. It can be observed from large numbers of dark red, red, and orange points and a small number of green points.
- From Figure 8, the performance differences among different item contents cannot be easily observed. However, Thai students performed slightly better on the items with quantity content.

3.2. Visualizing Science Cognitive Data

Figure 9, 10, 11, 12 and 13 show graphical visualizations, from the proposed method, generated from the science cognitive data of Thai students highlighted by various assessment aspects. The result shows two clusters of students. The largest or the main cluster includes most of the students. Cluster C on the left side of the plot shows a small group of students with unique characteristics. Due to the variety of assessment aspects, some students were not assigned to the items covering all assessment aspects. Each gray point represents a student who was not assigned to a particular aspect.

Here are some performance-based characteristics that we can observe from the results:

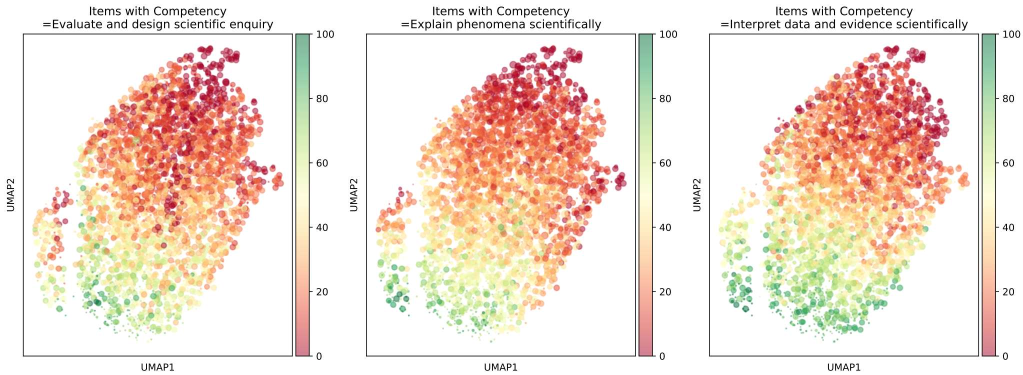


FIGURE 10. A group of scatter plots shows the average responses of each type of scientific competency. A point at the same location in Figure 9–13 represents the same student.

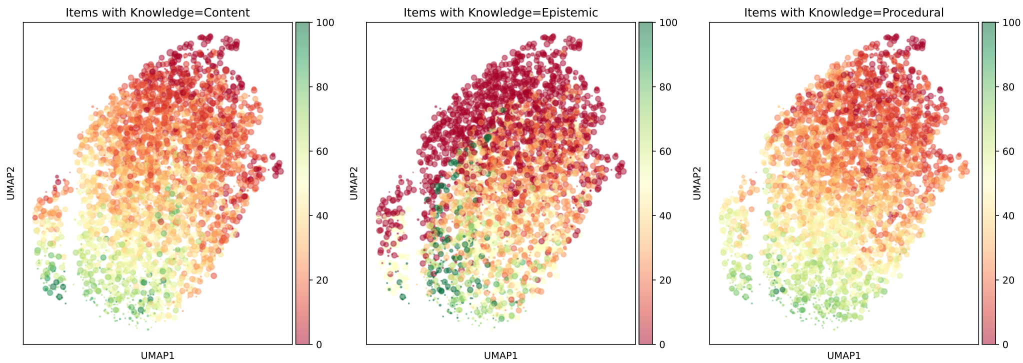


FIGURE 11. A group of scatter plots shows the average responses of each type of scientific knowledge. A point at the same location in Figure 9–13 represents the same student.

- From Figure 10, Thai students performed better on the items requiring *interpret* competency than the two other competencies. This is similar to the mathematics result where Thai students performed well on the items requiring *interpret* process.
- From Figure 11, Thai students performed worse on the items requiring *epsitemic* knowledge. However, there is a small number of students who performed very well on items with *epsitemic* knowledge (dark green points) even though their over all scores were low.
- From Figure 12 and 13, the students in cluster C performed very well with a response of greater than 80% on the items with *living* system.
- From Figure 13, many Thai students performed well on the items with *personal* context even though their overall scores are low.

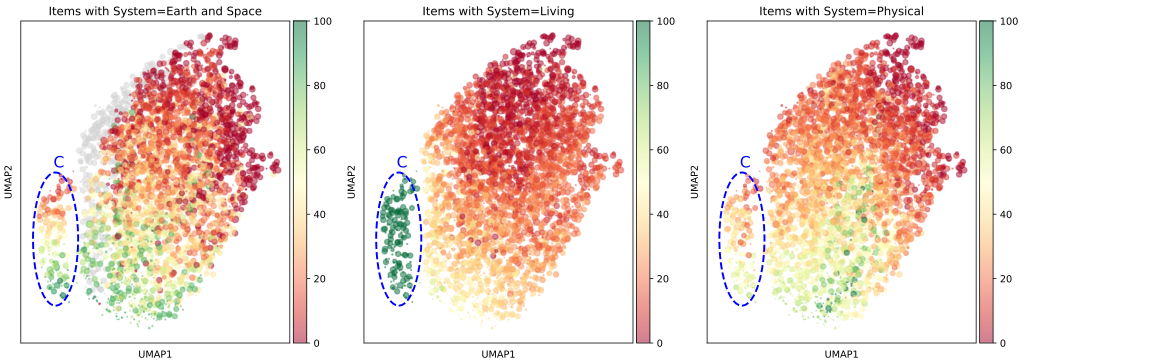


Figure 12. A group of scatter plots shows the average responses of each type of scientific system. A point at the same location in Figure 9–13 represents the same student.

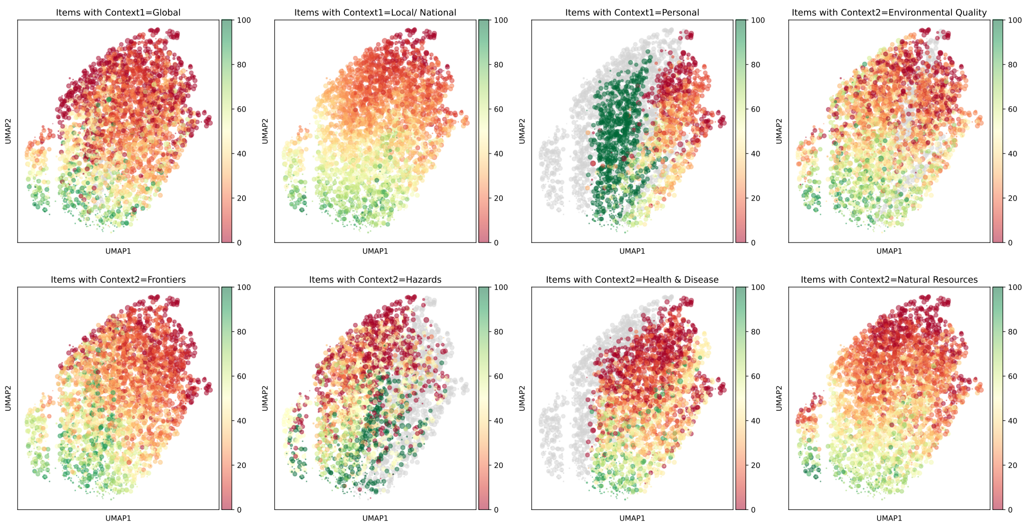


Figure 13. A group of scatter plots shows the average responses of each type of item context 1 and 2. A point at the same location in Figure 9–13 represents the same student.

3.3. Visualizing Reading Cognitive Data

Figure 14, 15, 16, and 17 show graphical visualization, from the proposed method, constructed from the reading cognitive data of Thai students highlighted by assessment aspects. Different from mathematics and science results, the reading result separates into many small clusters with one main cluster. Moreover, the numbers of students who perform well on reading items is the highest among the three subjects. This can be seen from the number of green points in Figure 14. The students who did well in the reading items are separated into two groups located at the bottom and the right side of the plot.

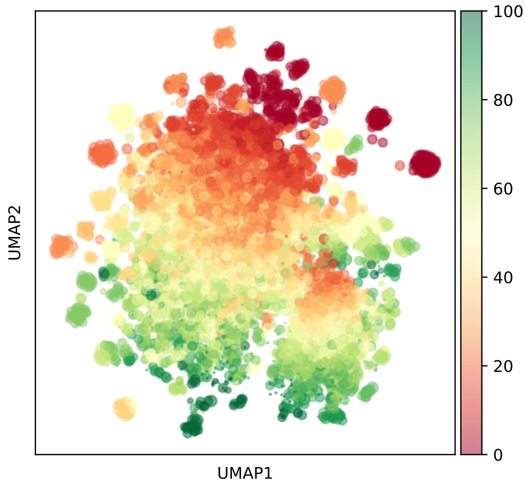


Figure 14. A scatter plot showing the overall performance of Thai students based on reading items. A point at the same location in Figure 14–17 represents the same student.

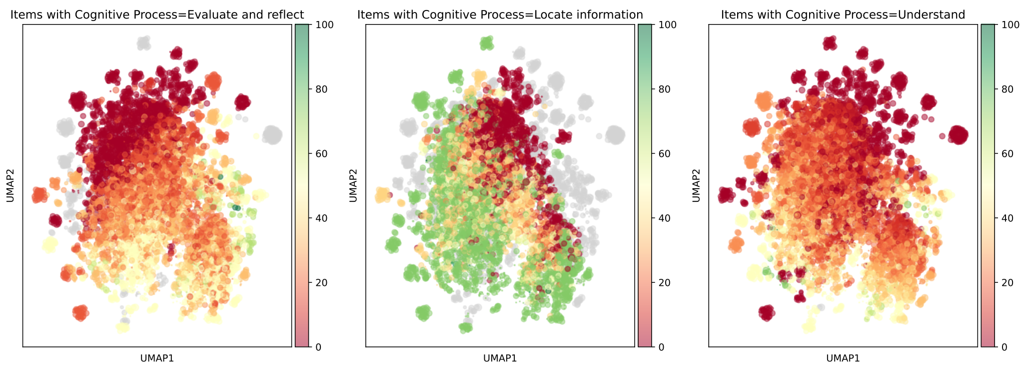


Figure 15. A group of scatter plots shows the average responses of each cognitive process. A point at the same location in Figure 14–17 represents the same student.

Similar to the science result, some students were not assigned to the items covering all assessment aspects. We use gray points to denote these students.

Here are some performance-based characteristics that we can observe from the results:

- From Figure 15, a large portion of Thai students performed very well on the items requiring the cognitive process of *locating information* even though some of them did not perform well overall. We can see from the result that many green points are scattered around half of the main cluster. Only a few students performed well on the items requiring the two other cognitive processes. Even students with high overall scores (green on overall) did poorly (red or orange)

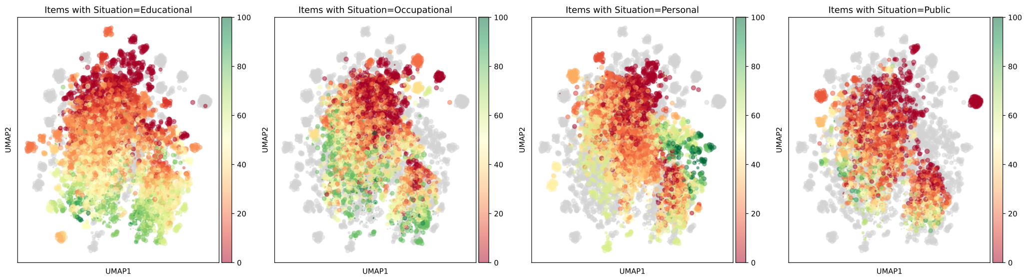


FIGURE 16. A group of scatter plots shows the average responses of each situation. A point at the same location in Figure 14–17 represents the same student.

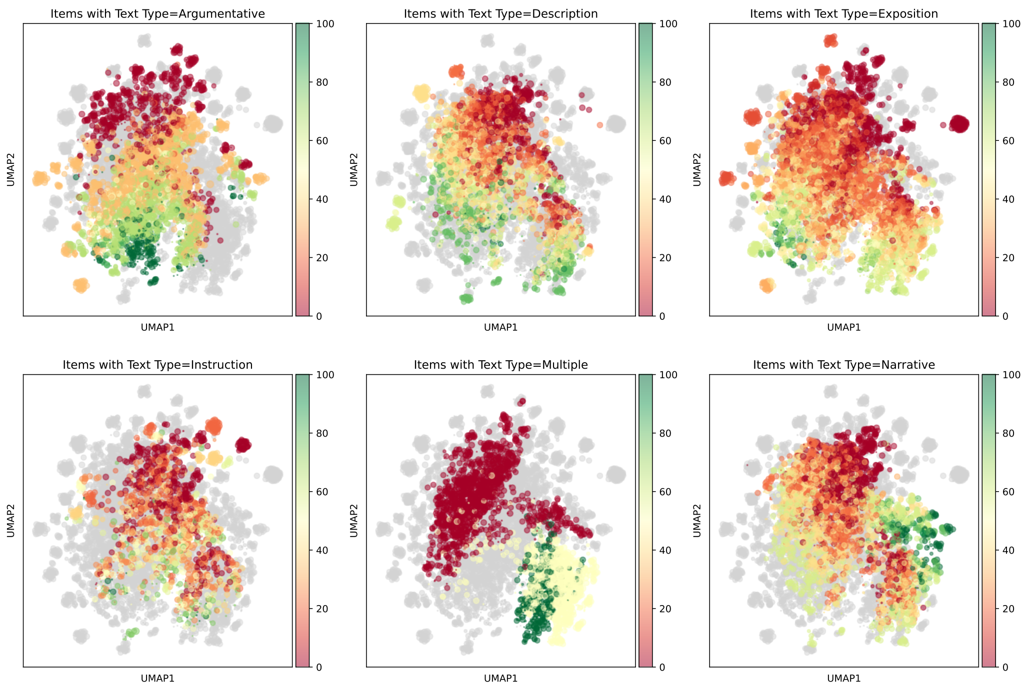


FIGURE 17. A group of scatter plots shows the average responses of each text type. A point at the same location in Figure 14–17 represents the same student.

with problems that require *evaluate and reflect* and *understand* cognitive process skills.

- From Figure 16, Thai students performed well in *personal* and *occupational* situations. This can be seen from the number of green points in the two situations. However, the groups of students who did well in different situations are different.

- From Figure 17, the performance in different text types are quite different. The result on the items with *multiple* text types has the highest number of red points which means that the students did not perform well on the items with *multiple* text types.

3.4. Relationship with Questionnaire Results

Figure 18, 19, 20, 21, 22, 23, and 24 show visualizations of selected questionnaire responses (left) versus the overalnoperformance of each subject (right). These visualizations also reveal some interesting characteristics of Thai students. Here are some highlights:

- From the OECD report [2] the average scores of girls are higher than boys' average scores in all three subjects (mathematics, reading and science). But if we look closer at the plots of genders versus overall scores in Figure 18, we can see that boys and girls who received high and moderate scores are distributed somewhat evenly in sciences and mathematics. There are visibly more boys (more blue points) than girls who received low scores in both mathematics and science. On the other hand, the distributions of boys and girls are different in the visualization of reading. There are more girls (more pink points) for students who have high and moderate scores but there are more boys for students who have low scores. The visualization of gender vs. scores in Figure 18 shows that even though Thai girls are performing better than Thai boys in all three subjects, the distributions of boys and girls in reading are different from those in mathematics and science.
- From Figure 19, students who received high and moderate scores (green and yellow points in the right image) mostly have expectations to complete a bachelor's degree level or higher (pink points in the left image) in all three subjects. A good portion of students with low scores did not expect to complete a bachelor's degree in all three subjects.
- From Figure 20, students who received high scores (green points on the right image) mostly agreed that their intelligence can be changed (blue points). Students with moderate and low scores (yellow to red points) have mixed feeling about changing intelligence. Visually, it seems like more Thai students believe that intelligence cannot be changed (more pink points than blue points).
- Figure 21 shows that more Thai students did not agree that reading is a waste of time (more blue points than pink points in the left image). However, the portion of students with low scores who think that reading is a waste of time is bigger than that of students with high scores. This behavior is consistent throughout all three subjects.
- Figure 22 shows that most Thai students have never or rarely been hit or pushed around by other students. However, the portion of students with low scores that have been hit or pushed around is bigger than those of students with moderate or high scores.
- From Figure 23, most Thai students read non-fiction books (visually more pinks points than blue points in the left images). Students with high scores read non-fiction books slightly more often than students with moderate and low scores.
- Figure 24 shows that many Thai students read digital texts (left images are bluer). Students with moderate to high scores read the digital text more often than students with low scores. This behavior is aligned with reading non-fiction habits from Figure 23.

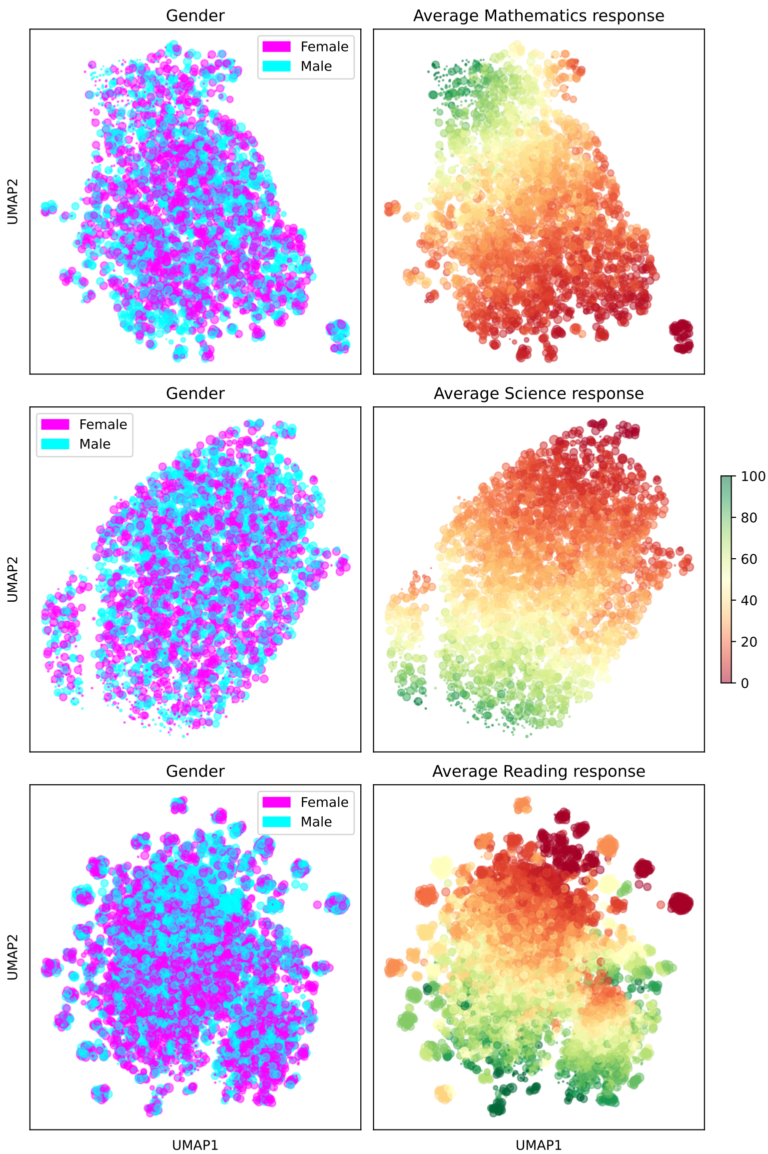


FIGURE 18. Scatter plots showing the relations between students' genders and the average response. A point at the same location on the left image and the right image of the same row represents the same student.

3.5. Comparison to Other Visualization Techniques

To show the advantages of the proposed visualization technique using the UMAP to visualize PISA 2018 data, we apply a number of commonly used data visualization techniques on the preprocessed PISA 2018 and compare the results with the UMAP results.

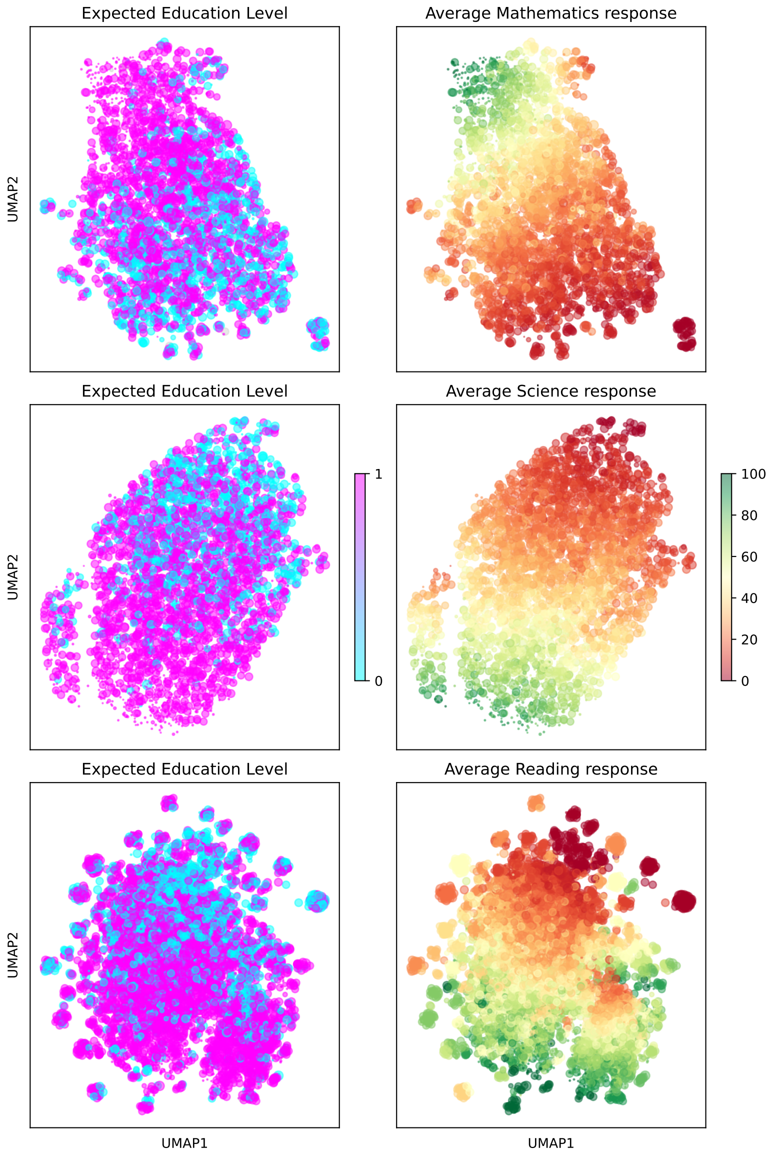


FIGURE 19. Scatter plots showing the relations between students' expectations to complete bachelor's degree level or higher and their average response. A point at the same location on the left image and the right image of the same row represents the same student.

3.5.1. Scatter Plot Matrix

Scatter plot matrix is a visualization technique to generate a scatter plot for each pair of variables and rearrange the plots in a matrix. The diagonal of the matrix shows a histogram of each corresponding variable. Figure 25 shows a scatter plot matrix constructed

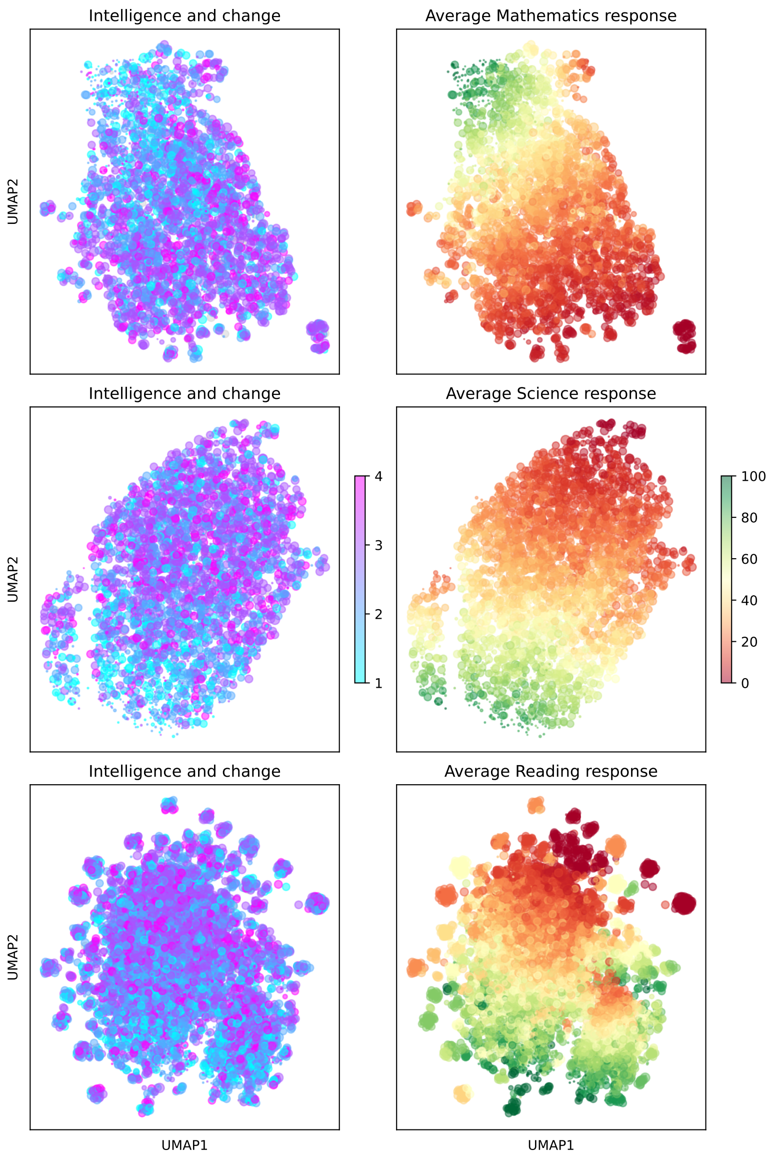


FIGURE 20. Scatter plots showing the relations between students' agreement to a statement '*Your intelligence is something about you that you cannot change very much*' (4 = strongly agree, 3 = agree, 2 = disagree, 1 = strongly disagree) to their average response. A point at the same location on the left image and the right image of the same row represents the same student.

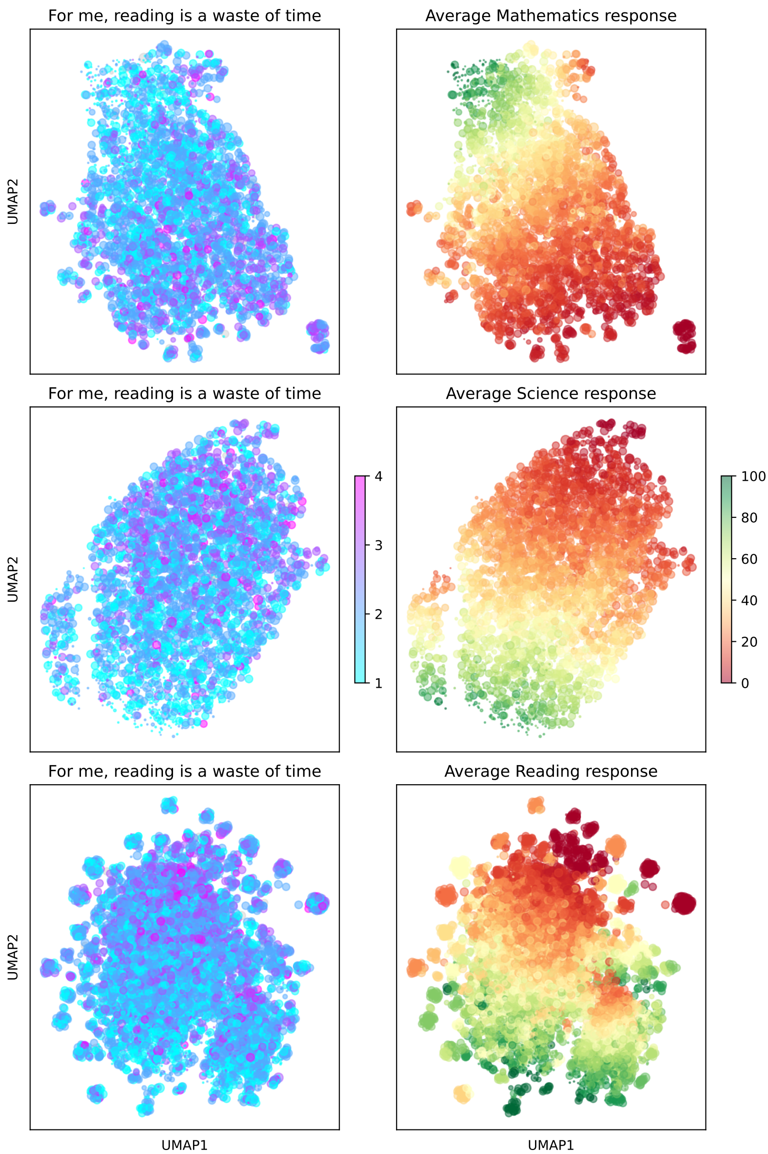


FIGURE 21. Scatter plots showing the relations between students' agreement to a statement '*For me, reading is a waste of time*' (4 = strongly agree, 3 = agree, 2 = disagree, 1 = strongly disagree) to their average response. A point at the same location on the left image and the right image of the same row represents the same student.

from average responses of the three mathematical processes. We can observe the correlations between the mathematical processes, and the distributions of average responses of the three mathematical processes.

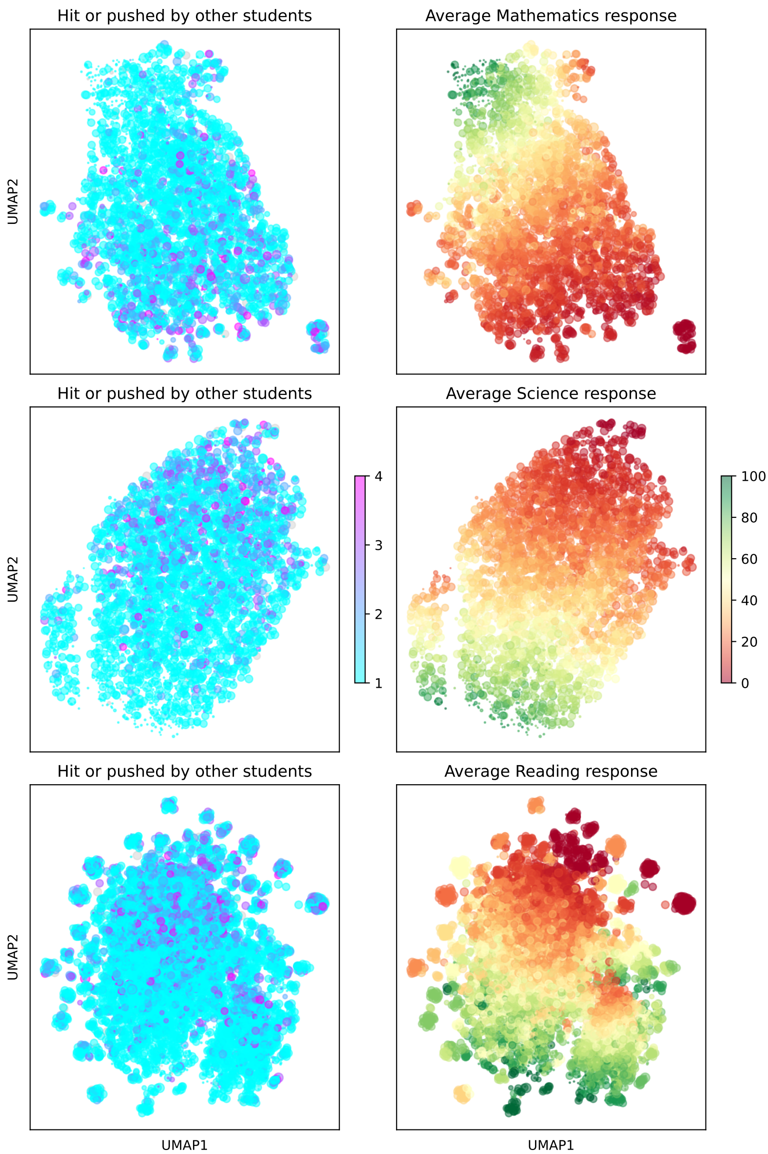


FIGURE 22. Scatter plots showing the relations between students' experiences to a statement '*I got hit or pushed around by other students*' (4 = once a week or more, 3 = a few times a month, 2 = a few times a year, 1 = never or almost never) to their average response. A point at the same location on the left image and the right image of the same row represents the same student.

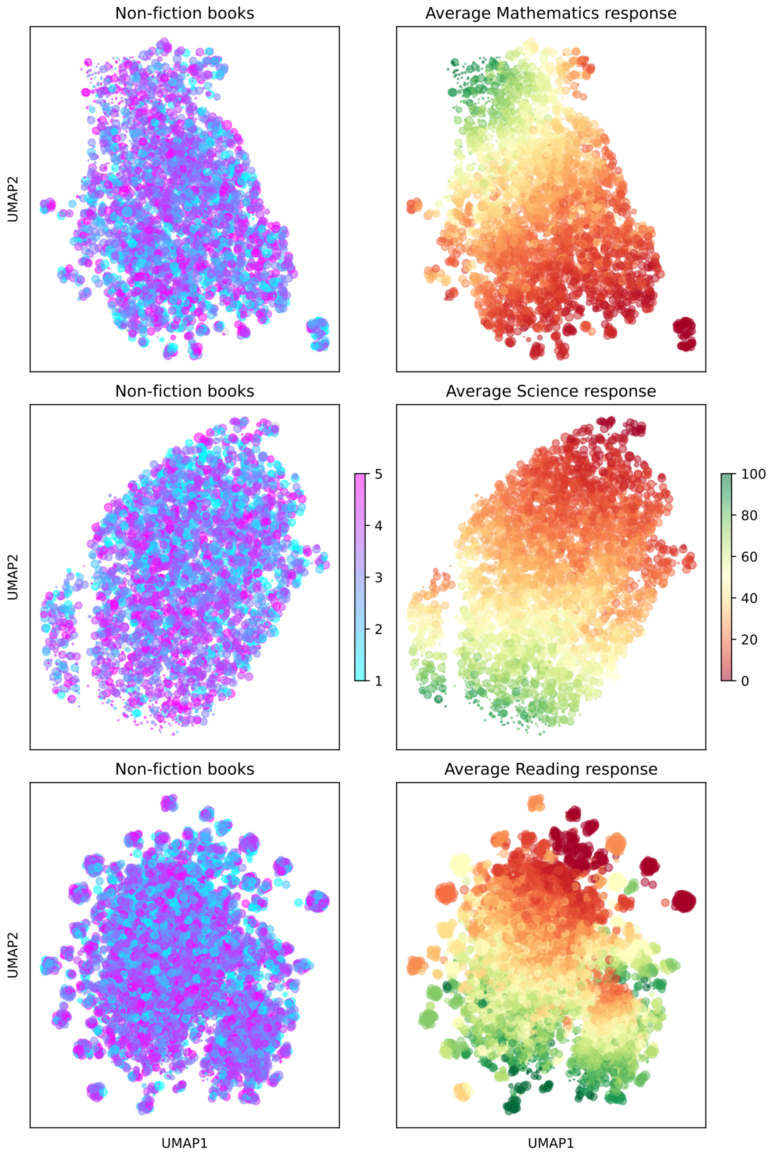


FIGURE 23. Scatter plots showing the relations between students' frequency to read '*non-fiction books (informational, documentary)*' (5 = several times a week, 4 = several times a month, 3 = about once a month, 2 = a few times a year, 1 = never or almost never) to their average response. A point at the same location on the left image and the right image of the same row represents the same student.

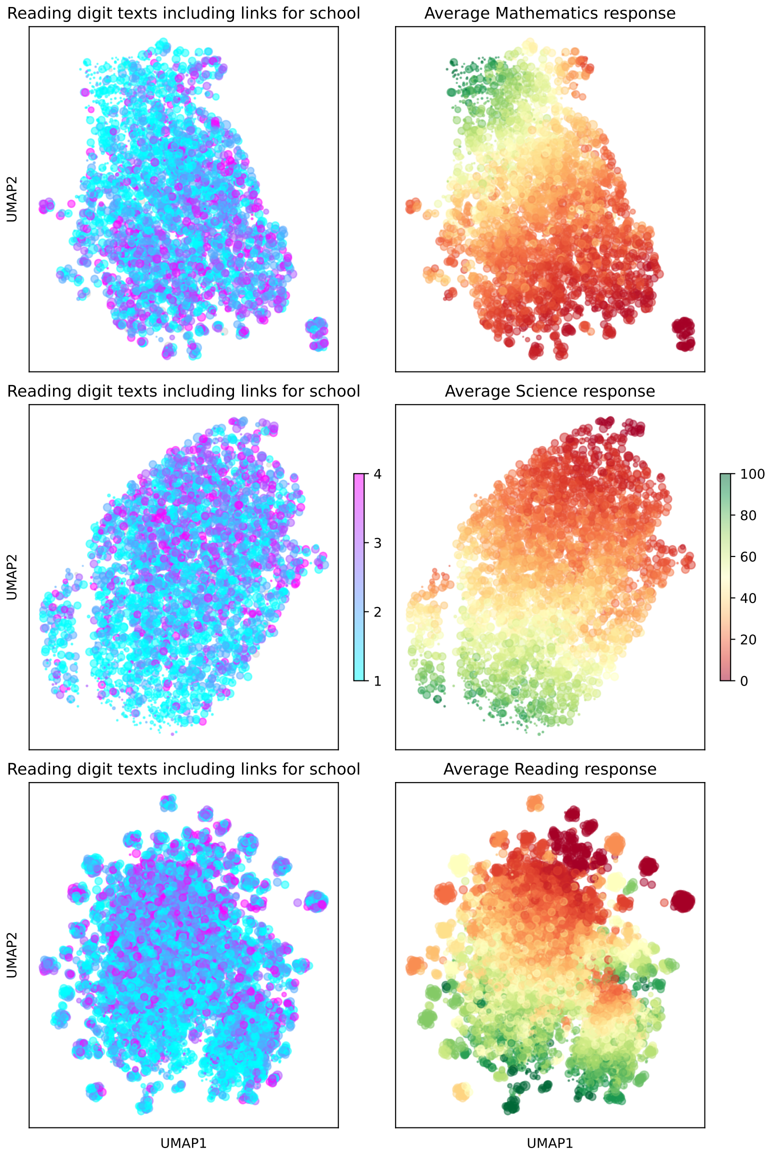


FIGURE 24. Scatter plots showing the relations between students' frequency to read '*digital texts including links in the classroom or for homework*' (1 = many times, 2 = two or three times, 3 = once, 4 = not at all) to their average response. A point at the same location on the left image and the right image of the same row represents the same student.

Compared to the UMAP results in Figure 5 and 6, the scatter plot matrix can only provide the overall relation between a pair of variables. Local relations between subranges cannot be visualized by the scatter plot matrix. For example, the scatter plot cannot

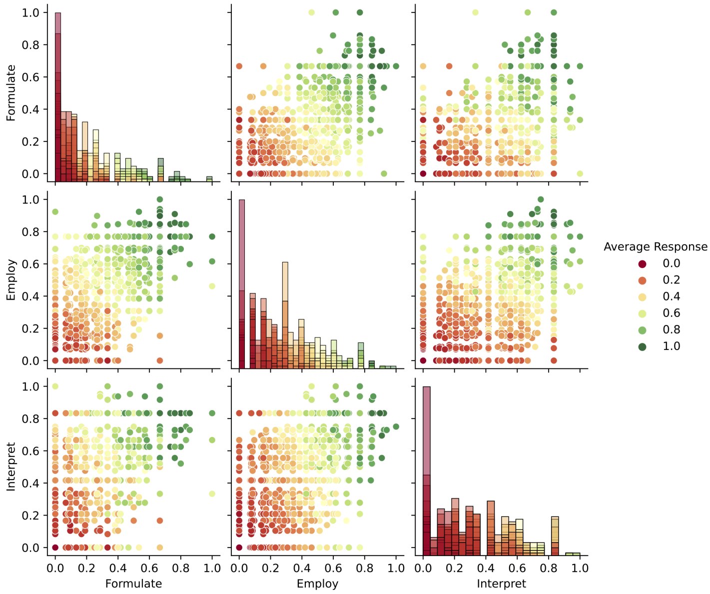


FIGURE 25. A scatter plot matrix showing mathematical process responses highlighted by average response

reveal the information that most of the students performs worse on the items requiring the *formulate* process than those requiring the *interpret*. Moreover, the scatter plot matrix can only be used when the number of variables is small enough to show the entire matrix.

3.5.2. Parallel Coordinates Plot

Parallel coordinates plot is a technique to visualize high-dimensional data using parallel axes instead of orthogonal axes. Thus, each data point is represented as a sequence of line segments. Figure 26 shows a parallel coordinate plot from the three mathematical processes. Each data point is represented two line segments connecting from the *formulate* axis to the *employ* axis and continuing to the *interpret* axis.

Compared to the UMAP results, it is difficult to study the data from the parallel coordinate plot especially when there are a lot of data points and they do not follow the same trend. The parallel coodinate may show the relation between consecutive axes, but it is difficult to figure out the relation between non-consecutive axes. For example, we can learn from the parallel coordiante plot that some students with high average response denoted by green line segments perform better on the items requiring the *employ* process than those requiring the other two processes. This is because we can notice many line segments go up and down at the top center of the plot. Similar to the scatter plot matrix, the parallel coordinate plot has a limitation when the number of variables is large.

3.5.3. Other Dimensionality Reduction Techniques

We compare the UMAP results to the other wide-used dimension reduction techniques i.e. Principal Component Analysis (PCA) and t-distributed Stochastic Neighbor Embedding (t-SNE) as shown in Figure 27. The results of the three techniques have similar

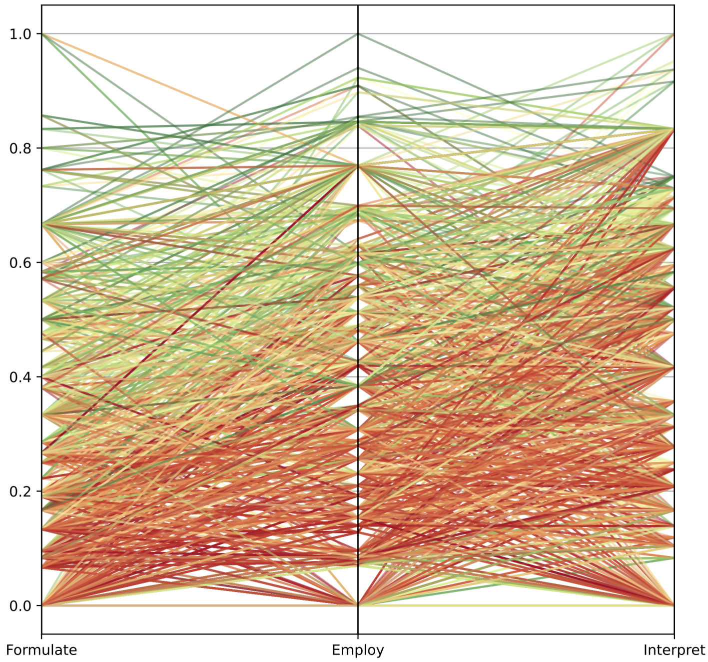
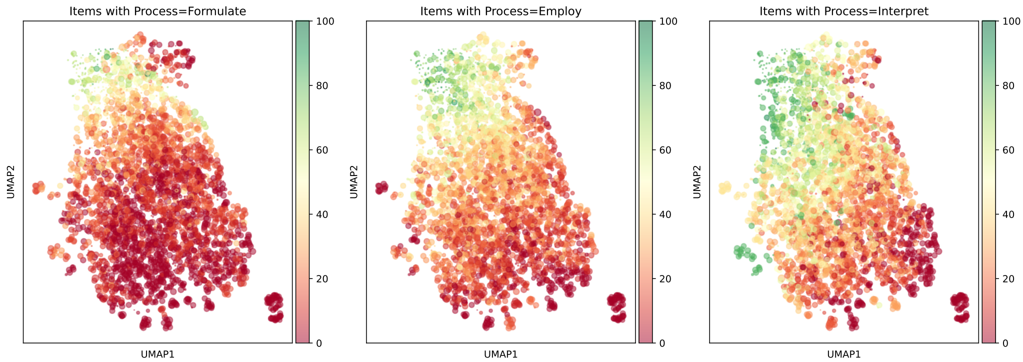


FIGURE 26. A parallel coordinates plot showing mathematical process responses highlighted by average response

properties. For example, they easily reveal that the students performs better on the *interpret* items than the *formulate* items. However, the UMAP technique based on topological data analysis is more capable to preserve both global and local structures of the data. The result from t-SNE contain many small clusters with similar characteristics, and the result from PCA does not show any cluster of data points.

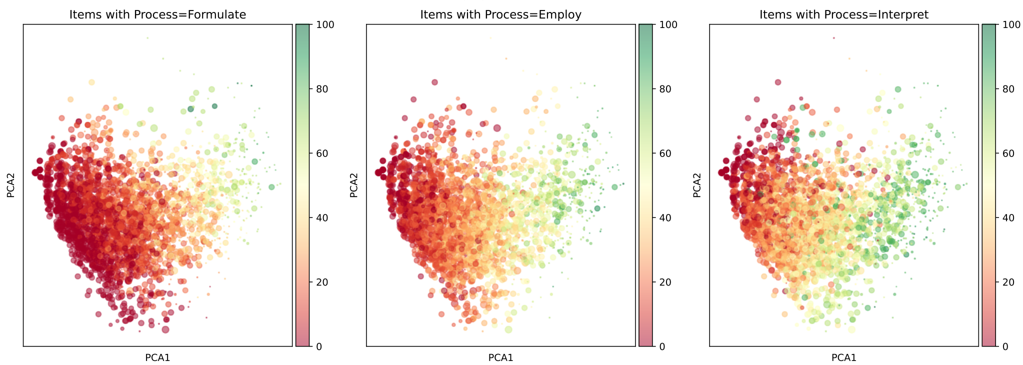
4. Conclusion

This work is conducted to extract more insights and gain a deeper understanding of Thai students and Thai compulsory school systems from PISA 2018 data using data visualization techniques. We proposed a technique to preprocess PISA cognitive item data so that we can apply a visualization technique to display each student's performances on 2-dimension plots. There were 8,633 Thai students participating in PISA 2018. Each student took tests on mathematics, reading and sciences. Moreover, one issue is that different students may have been taking different sets of questions. To overcome this issue, a student is represented by a vector in a high-dimensional space based on assessment aspects of each subject (22 dimensions for mathematics, 42 dimensions for reading and 40 dimensions for science). Each value in a vector that represents a student is an average score or a standard deviation that the student received for all questions with a particular assessment aspect. And to overcome a curse of dimensionality, we apply the UMAP technique to display PISA 2018 results. The size of each point in the visualizations reflects the value of the student's final weight (a size of a population that one particular student



(a) UMAP

(b) t-SNE



(c) PCA

Figure 27. A parallel coordinates plot showing mathematical process responses highlighted by average response

represents). We also apply a color scheme to reflect (red to green; low to high) how a student performs in particular assessment aspects or overall scores. We also apply another color scheme (pink to blue) to illustrate students' responses to PISA questionnaires. With the proposed technique, we are able to gain a deeper understanding of the characteristics of Thai students beyond just average scores.

Some of our findings related to assessment aspects are as follows. In both mathematics and science, Thai students were better with interpret skills than other types of skills. In science, Thai students performed better with questions related to personal context but did quite poorly with questions related to Epistemic knowledge. In reading, Thai students were better with questions related to the process to locate information but did poorly with questions related to the process to evaluate and reflect and the process to gain understanding.

Some highlights of our findings related to socio-economy factors are as follows. we found that students with high and moderate scores were both boys and girls with visually the same portions in mathematics and sciences. Only students with low scores in mathematics and sciences were more boys than girls. On the other hand, in reading, students with high and moderate scores are more girls and students with low scores are more boys. This shows that even though, on average, Thai girls performed better than Thai boys in all three subjects, the distributions of boys and girls, based on their performances, have different characteristics in different subjects. We also found that students with high scores read both books and digital materials more often than students with moderate and low scores. Almost all students with high scores in all subjects expect to complete at least a bachelor's degree. And students with low scores were bullied more often than students with high and moderate scores in all subjects.

As mentioned above, our findings from the proposed method go beyond the fact that Thai students performed below the OECD average. We are able to see more characteristics of Thai students with the proposed technique that other techniques cannot. With these new insights, educators and policymakers can use these new insights to improve and reform the Thai compulsory education system.

Acknowledgments

This work was supported by the Thammasat University Research Unit in Gait Analysis and Intelligent Technology (GaitTech).

References

[1] E. Auld and P. Morris. PISA, policy and persuasion: Translating complex conditions into education 'best practice'. *Comp. Educ.*, 52:202–229, 2016.

[2] F. Avvisati, A. Echazarra, P. Givord, and M. Schwabe. *Country Note: Programme For International Student Assessment (PISA) Results From PISA 2018*. OECD Publishing, Paris, France, 2019. https://www.oecd.org/pisa/publications/PISA2018_CN_THA.pdf [accessed on 17 January 2023].

[3] E. Becht, L. McInnes, J. Healy, C.-A. Dutertre, I. W. H. Kwok, and L. Guan. Dimensionality reduction for visualizing single-cell data using UMAP. *Nature Biotechnology*, pages 38–44, 2019.

[4] A. B. I. Bernardo, M. O. Cordel, Lucas R. I. G., J. M. M. Teves, S. A. Yap, and U. C. Chua. Using machine learning approaches to explore non-cognitive variables influencing reading proficiency in english among filipino learners. *Educ. Sci.*, 11(628), 2021.

[5] S. Breakspear. *How Does PISA Shape Education Policymaking? Why How We Measure Learning Determines What Counts in Education.* Center for Strategic Education: Melbourne, Australia, 2014.

[6] J. A. Campbell. The moderating effect of gender equality and other factors on pisa and education policy. *Educ. Sci.*, 11(10):503–521, 2021.

[7] X. Don and J. Hu. Using machine an exploration of impact factors influencing students' reading literacy in singapore with machine learning approaches. *International Journal of English Linguistics*, 9(5), 2019.

[8] Arantza Gorostiaga and José Luis Rojo-Álvarez. On the use of conventional and statistical-learning techniques for the analysis of pisa results in spain. *Neurocomputing*, 171:625–637, 2016.

[9] J. Gustafsson. Effects of international comparative studies on educational quality on the quality of educational research. *Eur. Educ.Res. J.*, 7:1–17, 2008.

[10] B. Gutiérrez-de Rozas, E. López-Martín, and E. C. Molina. Defining the profile of students with low academic achievement: A cross-country analysis through pisa 2018 data. *Front. Educ.*, 7(910039), 2022.

[11] N. Gutiérrez, I. Mercader, J. J. Carrión, and R. Trigueros. Self-concept and feeling of belonging as a predictor variable of the attitude towards the study from the pisa 2018 report. *Educ. Sci.*, 12(91), 2022.

[12] J. M. Kane and J. E. Mertz. Debunking myths about gender and mathematics performance. *Notices of the American Mathematical Society*, 59(1):10–21, 2012.

[13] Leland McInnes, John Healy, and James Melville. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv*, 2018. https://arxiv.org/abs/1802.03426 [accesed 17 January 2023].

[14] Leland McInnes, John Healy, Nathaniel Saul, and Lukas Großberger. UMAP: Uniform Manifold Approximation and Projection. *Journal of Open Source Software*, 3(29):861, 2018.

[15] M. Neuman. Pisa dtat clusters reveal student and school inequality that affects results. *PLOS One*, 17(5):e0267040, 2022.

[16] OECD. Item pool classification. In *PISA 2018 Technical Report*, chapter Annex A, pages 1–1. OECD Publishing, Paris, France, 2019. https://www.oecd.org/pisa/data/pisa2018technicalreport/PISA2018%20TechRep_Final-AnnexA.xlsx [accessed on 17 January 2023].

[17] OECD. *PISA 2018 Assessment and Analytical Framework*. OECD Publishing, Paris, France, 2019. https://www.oecd-ilibrary.org/content/publication/b25efab8-en [accessed on 17 January 2023].

[18] OECD. Test design and development. In *PISA 2018 Technical Report*, chapter 2, pages 1–11. OECD Publishing, Paris, France, 2019. https://www.oecd.org/pisa/data/pisa2018technicalreport/PISA2018%20TecReport-Ch-02-Test-Design.pdf [accessed on 17 January 2023].

[19] L. Perry and A. McConney. Does the ses of the school matter? an examination of socioeconomic status and student achievement using pisa 2003. *Teachers College Record*, 112(4), 2010.

[20] D. Reilly. Gender, culture, and sex-typed cognitive abilities. *PLOS One*, 7(7):e39904, 2012.

[21] W. Schulz. Measuring the socio-economic background of students and its effect on achievement on pisa 2000 and pisa 2003. *Annual Meeting of the American Educational*

Research Association, 2005.

[22] G. Steiner-Khamsi. Cross-national policy borrowing: Understanding reception and translation. *Asia Pac. J. Educ.*, 34:153–167, 2014.

[23] G. Stoet and D. C. Geary. Sex differences in mathematics and reading achievement are inversely related: Within- and across-nation assessment of 10 years of pisa data. *PLOS ONE*, 8(3):1–10, 2013.

[24] L. Xiaomin and E. Auld. A historical perspective on the oecd's 'humanitarian turn': Pisa for development and the learning framework 2030. *Comp. Educ.*, 56:503–521, 2020.