MODESTUM EURASIA Journal of Mathematics, Science and Technology Education, 2022, 18(10), em2167
ISSN:1305-8223 (online)
OPEN ACCESS Research Paper https://doi.org/10.29333/ejmste/12455

Topic modeling of the student emails sent before and during the birth of COVID-19 in physics and math classes

Danilo Jr Austria Tadeo [1,2*], Junehee Yoo [1]

[1] Seoul National University, Seoul, SOUTH KOREA
[2] Yongsan International School of Seoul, Seoul, SOUTH KOREA

Received 06 February 2022 ▪ Accepted 27 August 2022

Abstract
The COVID-19 pandemic caused physical classes to suddenly transition to online learning all over the globe two years ago, resulting in students becoming more active in online email communication. The emails sent by the students were observed to contain students' concerns and needs for teacher support during the early stages of worldwide online classes due to COVID-19. As such, this study was interested in those email contents that were explored and analyzed through topic modeling, network analysis, and grounded theory. Six hundred twenty-three emails sent by seventy students in physics and math classes were analyzed using InfraNodus. This tool can perform topic modeling and visualize network graphs of verbal text data such as emails. By topic modeling and network graphical analysis, the findings revealed that the main topic clusters of the student emails corpus pertain to class assessments – questions and tests. Moreover, the influential keywords in the network graphs were coded, and the emails representing those keywords were further categorized using grounded theory. Doing so led to the finding that students needed teacher support on the content and supportive pedagogy. Supportive pedagogy needs may include test goals, schedule, content, and procedures, reviewing the test solutions and answers, and providing necessary test accommodations. Further study on teacher support in the online physics class and the effect of delivering teacher support on the student's performance can be a topic of future research.

Keywords: COVID-19, email communication, InfraNodus, text-network analysis, topic modeling

INTRODUCTION

It has been two years since when COVID-19 pandemic changed the norm and challenged the whole education community. The educators, in quick time, had to switch the teaching platform and deliver the lessons online. Students of all ages struggled with the use of technology and had to learn by themselves most of the time asynchronously. The curriculum developers and stakeholders brainstormed on providing quality education to the students. It was two years of embarking on the uncertainty of teaching, assessing, feedbacking, and grading.

During that time, there was research on how the pandemic impacted education, so that teachers have to offer interventions to help students (Sintema, 2020), and another study on what factors students view may support their readiness for a sudden change in synchronized online learning (Naji et al., 2020). These two recent studies influenced the researcher to pursue this work and thought about what concerned students during the year when the COVID-19 pandemic struck.

This study recognized students' representative areas of concern that would have needed teacher support through topic modeling, network graph analysis, and the grounded theory. Initially interested in the students' felt needs when teaching and learning happened online, the researcher was interested in identifying the areas where students also needed support before the pandemic. To determine these concerns, the researcher utilized a year of email communication between the teacher and the students since it was the only thing that consistently occurred before and during the pandemic. By topic modeling, the topic clusters of emails were chosen, then the nodes representing the clusters were categorized. The indicative email communications belonging to the categorizations led to the discovery of the support students needed during the pandemic.

While the results section divides into three sub-topics answering the three main questions of the study, the

✉ sirdhan@snu.ac.kr **(*Correspondence)** ✉ yoo@snu.ac.kr

Contribution to the literature

- The study showed how essential email communication was to the students during online learning due to COVID-19 pandemic.
- Students' email communication accurately represents their words, actions, and behavior, imploring teacher support needed in an online learning environment.
- Topic modeling and network graph visualization provided valuable information about the student's concerns about teacher support revealed in the emails sent during the COVID-19 pandemic.

methodology presents the description of the email communication corpus in the methodology section to aid the discussion of the whole paper.

Literature Review

Email communication

The history of sending emails started in 1971. During that time, the US Department of Defense implemented the Advanced Research Projects Agency Network (ARPANET) to connect numerous computers across the agency for communication purposes. Ray Tomlinson first invented and developed electronic mail by creating ARPANET's networked email system. Exchanging emails between computers became possible when he used the symbol "@," which had become necessary as to where the message would be addressed. In 1976, the idea of sending email to a computer outside the internal network sprung up, connecting people around the globe through Internet Service Providers (ISPs) in 1980. After a decade, the word electronic mail was replaced by email in the general dictionary and became more widely used. Then within the next two decades, AOL, Hotmail, Yahoo, Microsoft, and Google fashioned the internet and email. In 2007, Apple first introduced mobile emails when the first iPhone was released 2007 (Gibbs, 2016). In this period, emails have become a massive part of human's daily activity up to the present time.

With billions of emails sent daily, email has become a standard mode of communication to share information, send greetings of all types, and exchange files such as photos, videos, and website links. In education, email communication is one of the main factors in learning. For example, students email their teacher when needing assistance–questions about homework assignments, upcoming exams, and problems students think teachers can help them with (Dürscheid et al., 2013). Emails allow the teacher to communicate swiftly the same information to many students. Hence, the communication between the teacher and the students is all-important. Communication between teachers and students link students' learning process (Zhou et al., 2018). Due to the current pandemic, emails and other social media have become the primary ways of communication, even between teachers and students. This communication method has become more prevalent after traditional face-to-face classes have switched to online courses.

Research works that utilized email communications in online courses revealed that emails help promote student cognitive growth and supported the idea that emails can construct online social and academic learning support (Yu et al. 2002, as cited in Miller, 2020) and that sending emails is an ideal channel of communication that can motivate students in online learning (Parte et al., 2021). In addition, Parte et al. (2021) suggest that sending motivational emails helps the students achieve learning objectives based on the students' positive responses to online learning and academic performance. A recent study by Maphosa et al. (2021) showed that students who utilized emails for communicating with the teachers through their mobile phones were said to have experienced academic support. Moreover, when teachers provided feedback on their work, students were satisfied with the teaching and learning process (Maphosa et al., 2021).

Dickinson (2017) worked on a study involving email conversations. He studied the impact of the email tone on students' performance and teacher evaluations by using verbal immediacy–a more friendly manner within bodies of emails sent to the students. Examples were, 'good luck, 'I am happy to help you, 'have a safe weekend,' and other statements encouraging social connection and support. The study results were interesting: the student's success rates increased positively between two consecutive semesters, and a more positive teacher evaluations during the semester when the teacher incorporated verbal immediacy in the emails.

During the recent COVID-19 pandemic, studies have determined the relation between online learning and communication. Alamwaleh et al. (2020) explored the effect of online learning on negative communication between the teacher and students. Their semi-structured online survey revealed that online learning decreased the teacher and students' communication levels through random sampling. They suggested that teachers must build opportunities for interactions and communications during online hours. Another study by Kayumova et al. (2021) indicated that teachers could solve communication problems due to the pandemic by providing the younger students with offline activities through a platform called "ROUND." Their study introduced the possibility of utilizing such an offline interactive platform for students of different

developmental levels of complicated socioeconomic statuses. Although the works of Alamwaleh et al. (2020) and Kayumova et al. (2021) did not directly make use of email, both works implicitly present the importance of communication during online learning. As such, for the online learning environment, sending emails is an appropriate method of communication between the teacher and students.

Understanding emails sent by students during online learning is exciting and challenging. Recently a study suggested investigating the students' email usage and perceptions (Huang & Chen, 2021). Student emails feature texts and social interaction. Like any other text statement, one can analyze email communications. The use of text analysis and network analysis helps understand emails. The most common analysis is topic modeling. In network analysis, 'the objects of interest are the interactions between model elements' (Shaffer et al., 2016).

Topic modeling

A topic modeling that considers the collection of documents to have evolved. Topic modeling is a new technique for organizing, searching, indexing, and browsing patterns of words in an extensive collection of documents (Alghmandi & Alfalqi, 2015). The importance of topic modeling is primarily to discover ways of using the words and the connections between documents sharing the same patterns. Secondly, a generative model can specify a simple probabilistic procedure to generate documents. Many have used topic modeling to analyze images, biological data, survey information, and data (Blei & Lafferty, 2009). This modeling can disclose important information hidden in the pool, allowing the identification of topics appearing and changing with time (Alghamdi & Alfalqi, 2015).

The primary topic modeling technique is the latent Dirichlet allocation (LDA). This kind of topic modeling is an automatic technique for visualizing, analyzing, and summarizing topics in an extensive collection of documents, such as the web, blogs, news articles, and literature. LDA is a generative model that can copy a writing process to generate a document. However, the drawbacks of this modeling are

(1) manual removal of stop words and

(2) inability to represent the relationship among topics discovered.

Blei and Lafferty (2009) aimed to provide minimum descriptions of the parts of a collection that would efficiently enable the processing of a more extensive collection of documents while retaining the essential statistical relationships useful in classifying, detecting, summarizing, and judging similarity and relevance the parts of the pool.

Yun (2020) utilized topic modeling LDA to identify the trends in the topics found in the two primary journals in physics education research–AJP and PRPER. The author collected 2,959 AJP abstracts classified as teaching under the educational field from 1934 to 2019 and selected 745 PRPER abstracts published from 2005 to 2019. The data collection was done electronically through the journals' websites. Python 3 executed text processing and topic modeling (Mallet's LDA model) while topic value determines the appropriate number of topics in the two journals. She found out that " pedagogical content knowledge" was the most common topic in the AJP papers while " assessment" and " students' reasoning process' were joint in the PRPER papers. Both journals have common topics such as " problem-solving," " school program" and introductory physics." The result meant that communities represented by the journals, namely, the science and the physics education communities, differ in the topic of interest. Furthermore, in the recent work of Shaw et al. (2017), topic modeling creates an instructional tool for assessing science discourse in interactive multimedia presentations made, shared, and discussed by students. The teacher might utilize the result to gauge the student discourse and overall participation in real-time, understand the gaps in their learning, and improve the teaching practices.

Topic modeling has been widely used in text analysis to find what topics are in a text and what proportion of each topic is present in a corpus. The elegance of topic modeling is that the top influential words usually give a meaningful interpretation of a topic. In general, the visualization or the network graphing help make these meaningful interpretations.

Network graph analysis

Network graphs show the connections among elements in coded data and represent them in a dynamic network structure. These network graphs model the cognitive network patterns of association between knowledge, skills, values, habits of mind, and other elements characterizing complex thinking derived from the text corpus. In a network graph analysis, the interactions between model components are the objects of interest. One tool that can visualize graphs is the InfraNodus.

InfraNodus is a stand-alone software for organizing and understanding notes, speeches, and parts with the potential for novel ideas and insights. It is a web-based open-source tool that utilizes the text network analysis algorithm that signifies a network graph of any text and determines the most potent words based on their co-occurrence in a discourse (Paranyushkin, 2019). The co-occurrence of concepts within the conversations provides more insight into the content of the emails to examine the relationships and associations of these topics in the student's discourse to develop a model (Cai et al., 2017).

Teacher support

There have been a few articles that suggest how teachers can support their students' learning online. One online article by Maphosa et al. (2021), an educational psychologist, proposed five key considerations that educators must consider when supporting the students' online learning: *instruction*, *content*, *motivation*, *relationships*, and *mental health*. These five considerations can be how schools and teachers can develop and deliver online learning to optimize students' learning.

According to Martin (2020), the five considerations are, as follows:

1. *Instructions* refer to transferring knowledge to the students online or in the classroom. Instructions should be explicit, orderly, and well-organized, especially when delivering new or challenging subject matter.
2. *Content* is the subject matter shown through detailed instructions. The contents presented in the textbooks are made available online, written, and reviewed by experts based on the course syllabus.
3. *Motivation* refers to the student's energy and effort as they learn self and task-management, planning, and persistence.
4. *Interpersonal relationships* between the teacher and the learners and between the learners are integral to learning.
5. *Mental health* is a vital outcome, and a means to learn outcomes. Hence, good mental health means good learning, and inadequate mental means lousy learning.

Yates et al. (2020) have found that *supportive pedagogy*, about the teacher, personally checking with their students, giving feedback on student's work, and communicating with students and parents, through the effective use of technology, enable motivation, collaboration, and authentic learning activities to enhance the students' online learning experience. Another study showed that online learning negatively affects communication, suggesting that teachers create more opportunities for *interaction* and *communication* with the students, especially if there is an abrupt decrease in performance (Alawamleh et al., 2020). Furthermore, *technical support* is an essential factor vital to determining online class satisfaction (Nambiar, 2020) and that students are dissatisfied when they have limited skills (Zhang & Perris, 2004) and dissatisfied with online instructions when instructors are unable to provide technical support (Yang & Cornelius, 2004).

Purpose of the Research

This study explores and understands the student emails sent before and during COVID-19 using topic modeling and network graph analysis to identify the students' concerns about assessment in physics and math classes. Especially, to address following questions:

1. By topic modeling, what topic clusters represent the student emails, and how do these topic clusters differ before and during COVID-19?
2. Using the network graphs, how can the topic nodes in the main topic clusters be categorized, and how do these categorizations differ before and during COVID-19?
3. What concerns about teacher support do students have learning physics and math identified in the network graphs of the topic clusters?

METHODOLOGY

InfraNodus: Topic Modeling and Network Graphs Analysis Tool

The online software InfraNodus performed topic modeling to determine closely related concepts or persuasive topics co-occurring in the emails, identify the most significant nodes or top keywords that appear, and describe the network graph of the discourses by clustering. A network graph is a connection structure of the words in the student emails. Thus, for this study, the researcher explored students' email communication sent to the teachers before and during COVID-19, examined, compared, and analyzed using topic modeling and network graph analysis through InfraNodus.

Grounded Theory

In data analysis, coding is a way for the researcher to determine and describe the main ideas in each set of data (Glesne, 1999). A common form of determining these codes is through grounded theory, wherein the concepts grounded in a text are inductively identified within that text by the researcher (Glaser & Strauss, 2017).

Glaser and Strauss (2017) first proposed the grounded theory for social research in 1967, aiming to construct and scrutinize ideas from a data set. Initially, the grounded theory asserts that ideas naturally emerge from the data set, but Strauss and Corbin (1997) challenged the assertion and broadened the scope of the grounded theory beyond the restriction of only exploring naturally emerging ideas. Charmaz and Bryant (2010) further widened the theory's scope by allowing the researchers actively utilize their prior knowledge in coding, sampling, inquiring, and designing iterations. In The SAGE handbook of the grounded theory, further developments were compiled and edited by Charmaz and Bryant (2010), featuring the various versions and applications of grounded theory.

Learning visualizations can benefit by adopting the well-established social science research method. The application of the grounded theory generally features two main processes - the collection and the data analysis.

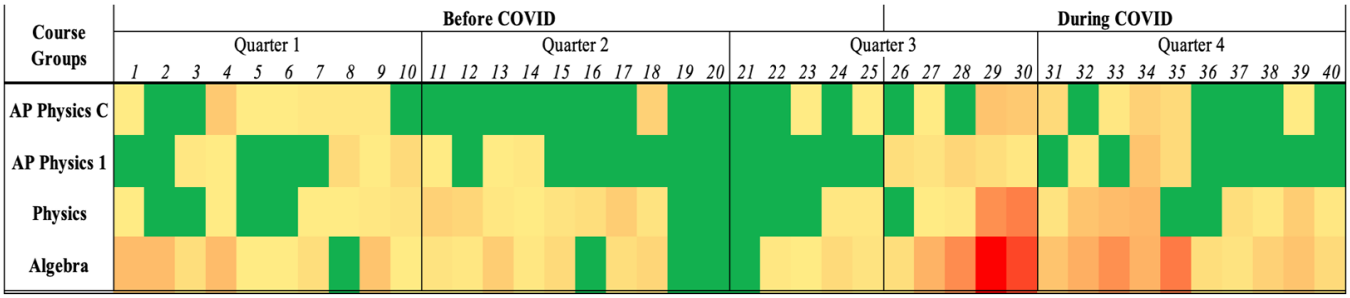

Figure 1. Heat map[1] of the emails received per week across the school year 2019-2020

Data collection may include observations, surveys, and interviews, while data analysis may consist of categorization, coding, comparisons, case analysis, and memos. Diehl et al. (2022) characterized grounded theory in visualizations by comparing grounded theory workflows with the common visualization and analytics workflows confirming that the grounded theory is beneficial for studying discourse data. Thus, our study that uses visualizations applied grounded theory when categorizing the nodes in the network graphs to present a general understanding of the nodes' main concepts.

Class Instructional Modes and Class Activities

Class instructions were done physically and face-to-face from August 2019 to February 2020. This period is considered before COVID-19. However, due to COVID-19 restrictions and guidelines for holding school, class instructions transitioned to online and virtual from the last week of February 2020 to June 2020. This period is considered *during COVID-19*. Hence, class instructional modes in 2019-2020 were face-to-face and virtual *before COVID-19* and during *COVID-19,* respectively.

The primary class activities during the entire school year of 2019-2020 were *lesson/HW*, *quiz*, *test*, *lab*, *practice*, *review,* and *project*. *Lesson/HW* is a regular activity wherein the teacher presents the lesson topic and gives work due to the following meeting. A *quiz* is a formative assessment after covering half of a unit, while the *test* is a summative assessment after completing a chapter of the unit lesson. *Lab* pertains to the hands-on and inquiry-based activities that culminate the whole chapter topic lessons. *Practice* and *review* are similar activities to prepare the students for the upcoming assessment. Finally, the *project* is an activity that requires students to develop or build up practical applications of the topic studied in the class.

The teacher only changed the modes of instruction. The activities given to four classes, physics, AP physics 1, AP physics C, and algebra 2, were not varied even though there was a change in the instructions. For example, the teacher taught lessons, whether face-to-face or virtual. Tests were given every after a unit of chapters. The teacher consistently gave *lab* during both periods.

Data Collection

The primary data in this study are emails sent by international high school students who took classes in physics (17), AP physics 1 (6), AP physics C (14), and algebra 2 (33) in an international school year 2019-2020. The students sent all emails to the teacher and the principal researcher of this study via schoology, the school's primary virtual environment and social networking service. The researcher did not collect any personal information for this study. The number of quarterly emails is 96, 77, 242, and 208 for Q1, Q2, Q3, and Q4. It is essential to note that the data collection was initially done only for exploratory purposes and took place in the summer of 2020 after the school year. Since the student-senders have just graduated or have left for vacation, there was no way for the teacher-researcher to send out consent forms. Instead, the school administration provided permission to use the data.

Web Scraper, a free and easy-to-use scraper for its point-and-click interface, scraped the emails from Schoology. It is a Google Chrome extension that quickly extracts desired contents from a pool of records on a selected website. The emails were saved and downloaded as comma-separated values (CSV). The online scraping tool disregarded and excluded the senders' names from the analysis. The desired contents for this study are only the messages. The Web Scraper automatically scraped the letters without including the names or any identification of the sender. In this manner, the teacher-researcher did not have any clue and did not in any way identify the sender. Furthermore, a third-party data analyst worked on the data processing for an unbiased and credible analysis to ensure the data's quality and trustworthiness. Hence, ethical considerations were observed, and the senders remained anonymous to the teacher-researcher.

Description of the Email Corpus: Email Sending Activity Before and During COVID-19

Figure 1 shows the heat map of the number of emails per week. The earliest email was on *week 1* (5th-9th of August 2019), and the latest was on *week 40* (1st-5th of June 2020).

[1] Legend: green–minimum *f*, yellow–medium *f*, red–high *f*

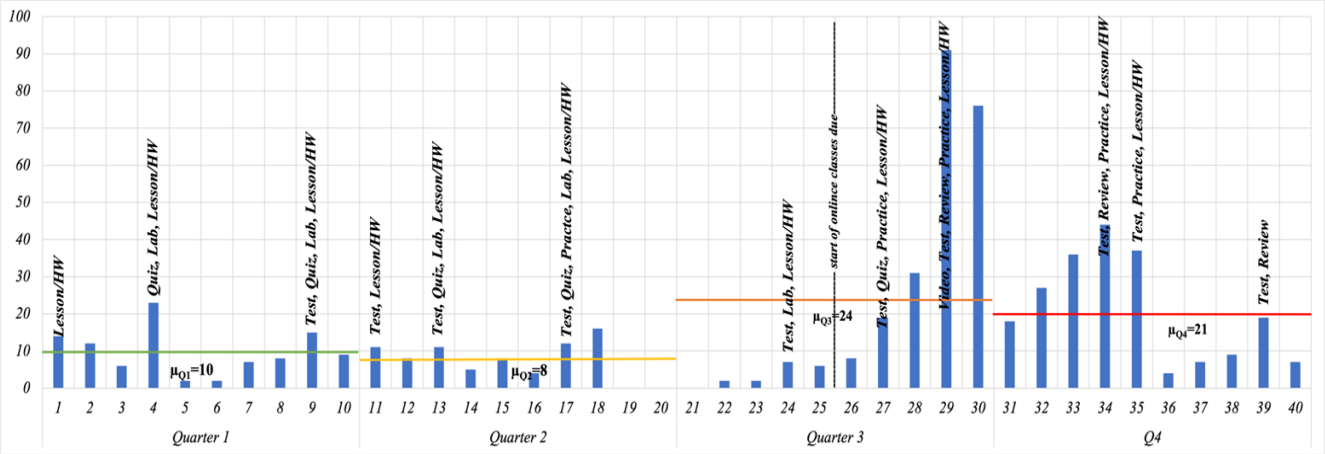

Figure 2. Class activities during the active weeks across the school year 2019-2020

The two highest email sending activities happened on *week 29* (23rd-27th of March 2020) during a week of *video, test, review, practice,* and *lesson/HW* in the fourth quarter. The highest emails received are from the lower-level course groups *physics* and *algebra 2*. In general, the flooding of emails happened during the second semester of the school when the school transitioned to online classes due to COVID-19. The pronounced increase in the number of emails sent was interesting to explore. The term "*active week*" is when the number of emails sent by students per course group is the highest.

Figure 2 shows a time-series graph for the entire academic year 2019-2020, divided into weeks. Each quarter has ten weeks, so there are 40 weeks in one academic year. The average quarterly emails are 10, 8, 24, and 21 for Q1, Q2, Q3, and Q4. The graph identifies the similarities and the differences between the email sending activity before and the during COVID-19.

Figure 2 also shows the typical class activities with higher sending activity for the weeks. The evident similarities identified are when the email sending activity is highest and what class activity occurs. Firstly, the email sending activity increases around the 4th (middle) week and the quarter's 9th (last) week. Secondly, the occurring class activity during those specific weeks is assessment. This assessment can be formative or summative such as *tests, quizzes, lab, projects,* practice, and *homework*. The *test* appears to be the everyday class activity around the 4th and the ninth week of each quarter.

On the other hand, the difference found is the relatively high number of emails sent during COVID-19 compared to those forwarded before COVID-19. Students started communicating through emails since physical meeting with the teacher is prohibited due to the COVID-19 restrictions. The earlier descriptive statistics support this result, implying that the lower-level course group was responsible for this increased email sending activity during COVID-19.

The author investigated the email sending activity and class activity per course group. Although the aggregated graphs are not presented in this report, the author identified the following significant results. Firstly, the lower-level course groups *algebra 2* and *physics* had more email sending activity before COVID-19. Specifically, the *algebra 2* students sent emails during a *quiz* or *test* day, while the higher-level course groups *AP physics C* and *AP physics 1* sent emails during a non-test day (lab or a project). Secondly, the email sending activity of both course groups boosted when the classes switched to online due to COVID-19 restrictions. With a higher number of emails sent (μ=26), a similar observation was found such that the lower-level course groups *algebra 2* and *physics* had the most emails sent during the *test* days, whereas the higher-level course groups *AP physics C* and *AP physics 1* had the most emails during non-*test* days (*video* or *practice*). Finally, the email activities continued in small amounts (μ=21) for the lower-level course groups while stopped for the higher-level course groups as the school year ended. The inference based on the observations is that the email sending activity varies depending on the class activity during a particularly active day. In addition, many busy days during COVID-19, specifically on *test* days, may account for students sending inquiries about the new procedures of administering assessments online.

During COVID-19, the teacher administered assessments through lockdown browsers to limit the students from viewing different browser tabs while taking the examination. In addition, the students were encouraged to turn on their laptop's cameras and microphones and ensure that their faces and half of the upper body appeared on the screen. The administration's time, the starting and ending times, and the exact time of answer submission and upload was also strictly observed. The school implemented these complicated procedures to prevent cheating during online assessments.

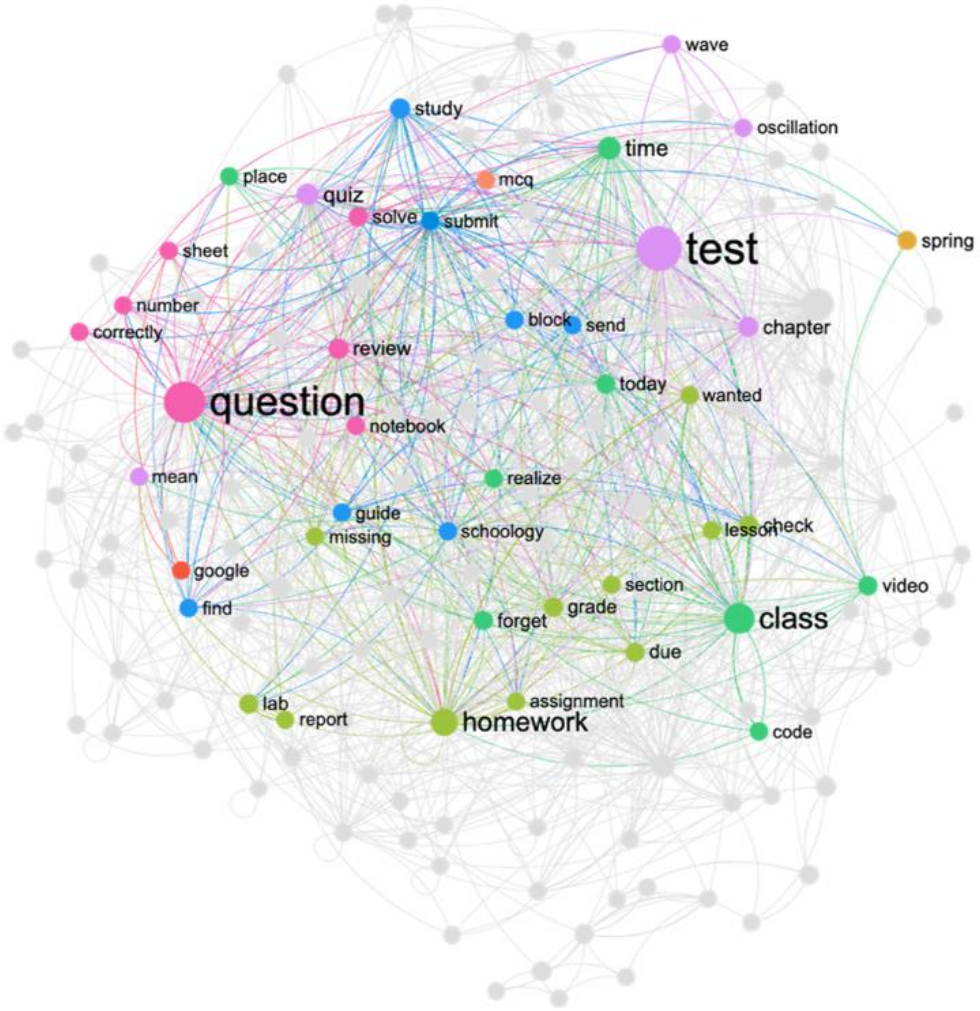

Figure 3. Email network structure revealing the topic clusters and keywords

Furthermore, students used to inquire about the items in the test, the solutions to the test items, and their scores. Since they could no longer interact with the teacher in the classroom after taking a test like they used to do, students tended to email the teacher to clarify or inquire about the test. That was why the email activity became more significant during COVID-19. Another remarkable result is that the higher course level groups stopped communicating with the teacher from the middle of the last quarter 4 when they finished the *advanced placement*® (*AP*) examinations. Hence, their email sending activity was great during *practice* and *review* days in preparation for the *AP* exams.

The descriptive section of this section presented the general backgrounded information of the actual data used for this study. The most significant email sending activity occurred during assessment days and was done mainly by lower-level course groups. Could the lower-level groups be more concerned about assessments and need teacher support?

RESULTS

This section reports the analysis of the student's email communication during the school year 2019-2020 using topic modeling through InfraNodus. The researcher interpreted and explained the analysis results to answer the questions of this research study.

What Topic Do Clusters Represent in the Email Corpus Before and During COVID-19?

The structure of the email corpus given by InfraNodus is in **Figure 3**. The email structure is *focused*[2], which means there is a focus on one topical cluster while there is some level of diversity. It shows the main topic clusters. Moreover, it presents the keywords that make up the topic clusters.

Table 1 summarizes the eight topic clusters representing email communication across all quarters, their node percentages, and the degree of the influential keywords in the email corpus.

[2] This kind of network structure indicates its level of diversity based on the modularity measure (Blondel et al., 2008) combined with the measure of influence distribution, as well as the percentage of nodes in the top community.

Table 1. Topic cluster node %, degrees of keywords, and topic clusters quarterly %

Topic cluster	Node (%)	Keyword	Degree	Before COVID-19			During COVID-19		
				Q1	Q2	Ave	Q3	Q4	Ave
Question-answer-review	27	question	45	18	22	20	24	28	26
Class-time-homework	23	class	21	25	23	24	22	18	20
Test-quiz-wonder	21	test	53	32	31	32	28	18	23
Conference-Schoology-find	16	conference	8	16	9	12	13	16	14
Pdf-adobe-document	5	pdf	3	8	7	8	6	5	5
Account-google-college	3	account	3	7	6	7	5	5	5
Unit-mcq-frq	3	unit	2	0	0	0	5	5	5
Study-block-guide	3	study	2	4	3	3	4	4	4

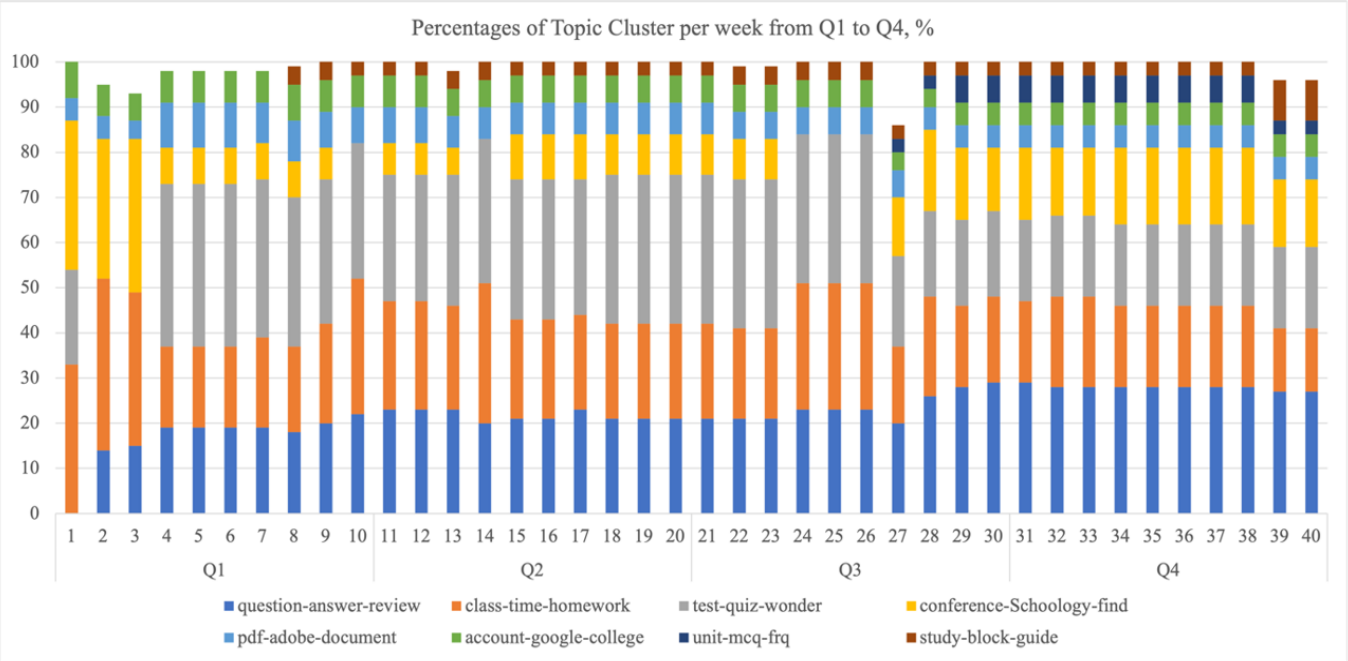

Figure 4. Topic cluster percentages per week across the school year 2019-2020

InfraNodus uses the *Jenks elbow cutoff*[3] algorithm in selecting the top prominent keywords of significantly more substantial influence. These most influential keywords are the topics with the highest *betweenness centrality*[4] or the highest degree. The complete statistics of the most influential keywords are in the appendices. Based on the topic modeling analysis results run by InfraNodus, students generally email the teacher concerning class assessments associated with the topics clusters *question-answer-review*, class-*time-homework*, and *test-quiz-wonder,* the most influential keywords *test*, *question*, and *class*. In this regard, it will be explored in the following sections if the email sending activities happened during the assessments.

Using the InfraNodus, the average percentages of the topic clusters before COVID-19 (Q1-Q2) and during COVID-19 (Q3-Q4) shown in **Table 1** were obtained. A similarity between the rates of topic clusters before and during COVID-19 is that the topic clusters related to assessment garnered at least 20%. These topic clusters are the *test-answer-review, class-time-homework,* and *test-quiz-wonder*. These main topic clusters include the most influential words such as a *question, homework,* and *test* related to assessment.

Conversely, the clear difference between the percentages of the topic clusters before and during COVID-19 is the distinct increase or decrease in their values. First, the increased percentages of the topic clusters *question-answer-review*, *conference-schoology-find*, and *study-block-guide* correspond to 6%, 2%, and 1%, respectively. Second, the percentage decrease in the topic clusters *class-time-homework*, *test-quiz-wonder, pdf-adobe-document,* and *account-google-college* decreased by 4%, 9%, 3%, and 2%, respectively. Last, the topic cluster that includes *units* only appeared during COVID-19 with a 5%.

Figure 4 plots the percentages against the weeks of the entire academic year 2019-2020 to show these differences.

[3] a 1-dimensional implementation of the K Means algorithm.
[4] the shortest path between two randomly chosen nodes.

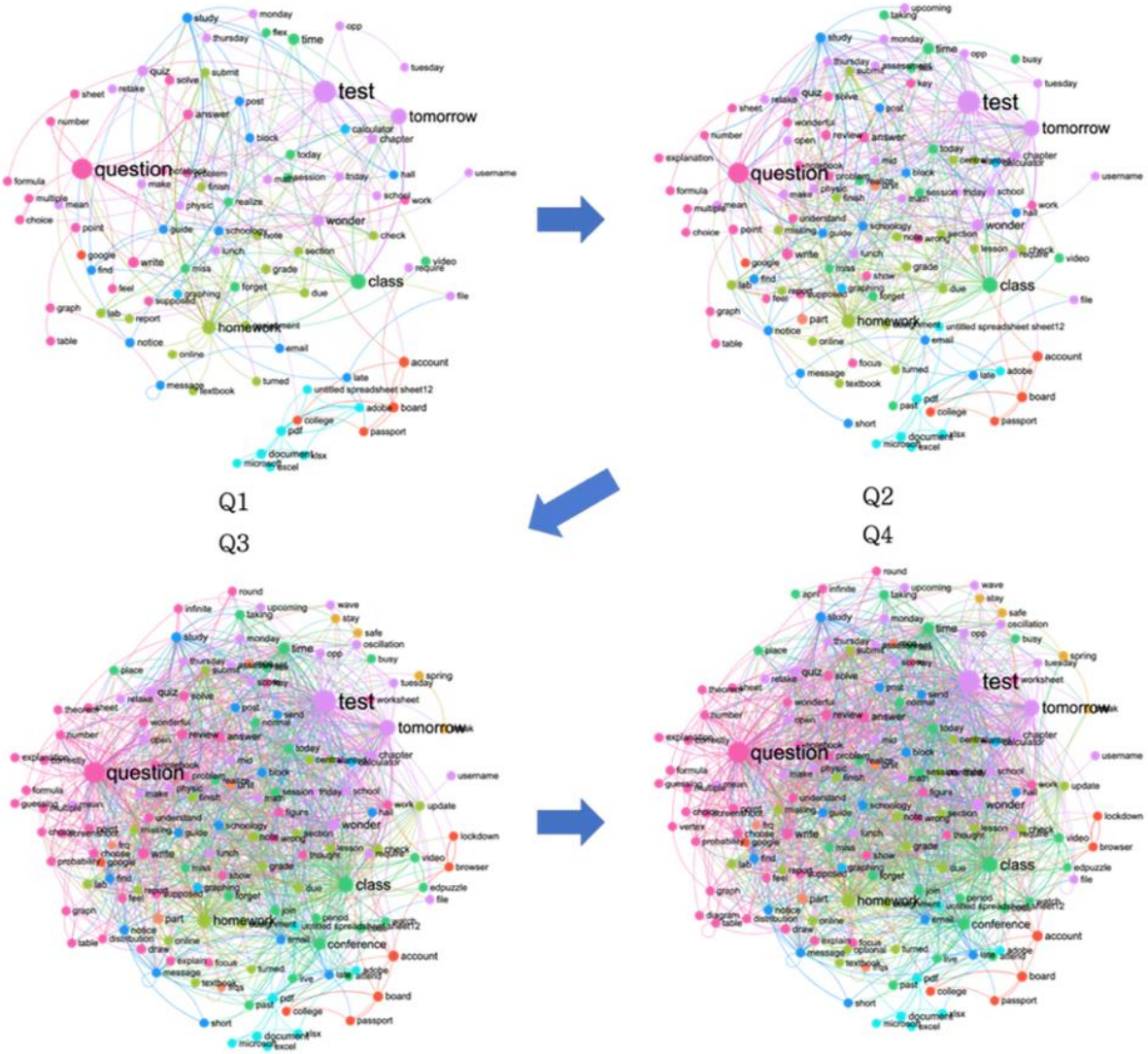

Figure 5. Topic cluster evolution from quarter 1 to quarter 4

The observed increase in the percentages of the topic clusters *question-answer-review*, *conference-schoology-find*, and *study-block-guide* may attribute to the fact that students become more active in email communication to inquire about things they used to do personally with the teacher. These are not limited to asking about test schedules, requesting a review or clarification on a lesson, feedback on tests, planning or recording the conferences, and asking for extra help and materials for test preparations.

Teachers were encouraged to unload the burden of the online class by lessening the number of hours of teaching virtual, decreasing the number of items given in any assessments. Conversely, the observed decrease in the percentages of the topic clusters *class-time-homework*, *test-quiz-wonder*, *pdf-adobe-document*, and *account-google-college* may explain the deloading imposed during COVID-19. Hence, the student emails would not contain many topics during COVID-19. Finally, the topic cluster *unit-mcq-frq* only appeared in the student's emails during COVID-19 because these terms are for assessment forms of the *advanced placement*® exams. Although the teacher started preparing the students for that test before the COVID-19 period, it was only during COVID-19 that there were specific emails on this matter.

Based on the results, test preparation became a big deal to both students and teachers, validating the current results of the Sintema (2020) study on the effect of COVID-19 on the performance of grade 12 students. He predicted that Zambian students would be affected in exam preparations due to the decreased contact hour-learning with the teacher. The study participants were not evaluated regarding their performances, but they initiated and sent emails to the teacher to inquire about the tests. In this manner, email communication became a solution to the lack of contact hours with the teachers.

The network structures run by the text-network tool InfraNodus further show the similarities and differences of the topic clusters of the whole email corpus. This tool uses a combination of clustering and graph community detection algorithms to identify the groups of more densely connected nodes than the rest of the network, as shown in **Figure 5**.

The nodal words that co-occur together in the same context are aligned closer on the graph, grouped by colors and network topology. The topical groups are pushed apart in the chart to show the network structure clearly. The similarity in the network graphs is that nodes of a *question, test, homework, and class* are present in all network graphs Q1, Q2, Q3, and Q4. Three nodes are related to assessment that has been a consistent part of the results of this study.

The structure of the network graphs can be *diversified* or *focused*. During Q1, before COVID-19, the network was *diversified* with a modularity of 0.54, and the email communication structure had an optimal level of plurality and diverse topic representation. Conversely, the graphs are focused on Q2 before COVID-19, Q3, and Q4 during COVID-19. They have modularity of 0.46, 0.34, and 0.39, respectively, which means there is a focus on one topical cluster while there is some level of diversity. The summary of the descriptions of the network graphs is.

In the case of a discourse network, high mind-viral immunity means that the text proposes multiple points of view and propagates its influence using highly influential concepts and more minor, secondary topics. The mind-viral immunity before COVID-19 Q1 and Q2 is high, while during COVID-19 Q3 and Q4 are medium. The higher the network's structure diversity, the network will be more resilient, adaptive, and less diverse.

The *diversified* network structure of the emails before COVID-19 may suggest many reasons why students emailed the teacher. Possibly, they had many concerns during that period. However, the *focused* network structure of the emails during COVID-19 may suggest that students had problems about why they sent emails during that period.

So far, in general, the most significant email sending activity occurred during assessment days and mainly was sent by lower-level course groups. Lower-level course groups seemed more concerned about the assessments revealed by the main topic clusters related to *testing, question,* and *homework.* The results also show that the email sending activity of the students varies depending on the class activity during a particularly active day. There were more active days during COVID-19, specifically on *test* days. This finding accounts for the students inquiring about the new administering assessments online. It supports the result that the network graph of the emails during COVID-19 is *focused, indicating* the specificity of their concerns regarding why they sent emails regarding assessments.

How Can the Topic Nodes Be Categorized Before and During COVID-19?

The network structure indicates its diversity based on *modularity* combined with *influence distribution* and the *percentage of nodes* in the top community. The network graph of the central topic cluster in the email corpus shown in **Figure 3** has a modularity of 0.32 and influence distribution of 50%, making the graph *focused*. This kind of network structure indicates a focus on one topical cluster while there is some level of diversity. The main influential words in the email corpus are the *test*, *question*, and *class* with the degrees of 122, 118, and 89, respectively. Moreover, these words have 170, 147, and 107, respectively. They are considered the main *hubs* of the entire network graph. A hub is a node that connects most or all other nodes in the network. The *bigrams* are the most prominent relationships between the nodes in the graph shown in **Figure 6**. The weights of the *bigrams* for both directions are 39 (*test-question*), 23 (*test-class*), 9 (*question-class*), 9 (*test-test*), and 3 (*question-question*).

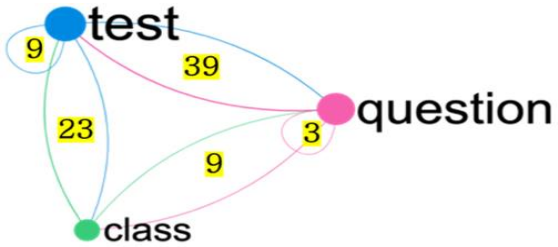

Figure 6. Network graph showing only the most influential topics

The same core network graph strongly connects the *test* and *question* nodes. Hence, this section examines the network graphs of the topic clusters containing the nodes *test* and *question*, determines what categorical concerns relate to the assessment, and how these found categories vary before COVID-19 and during COVID-19. These topics pertain to assessment concerns with the lower-level course groups in the preceding section.

InfraNodus provide the network graph shown in **Figure 7** for the *question-answer-review* and *test-quiz-wonder* topic clusters presenting how each node is close to each other and the central hubs *question* and *test.*

The researcher categorized the qualitative description, the context of the nodes, and the concern of the emails sent by the students regarding the assessment using grounded theory. In **Figure 7**, the categories related to assessment are, as follows:

(1) *Content,*

(2) *goal* of the assessment,

(3) *elements* in the assessment and the student's reflection, or

(4) self-assessment.

In addition, the other categories in **Figure 7** are, as follows:

(5) *subject,*

(6) *type,*

(7) *schedule,* and

(8) *grading* of the assessment.

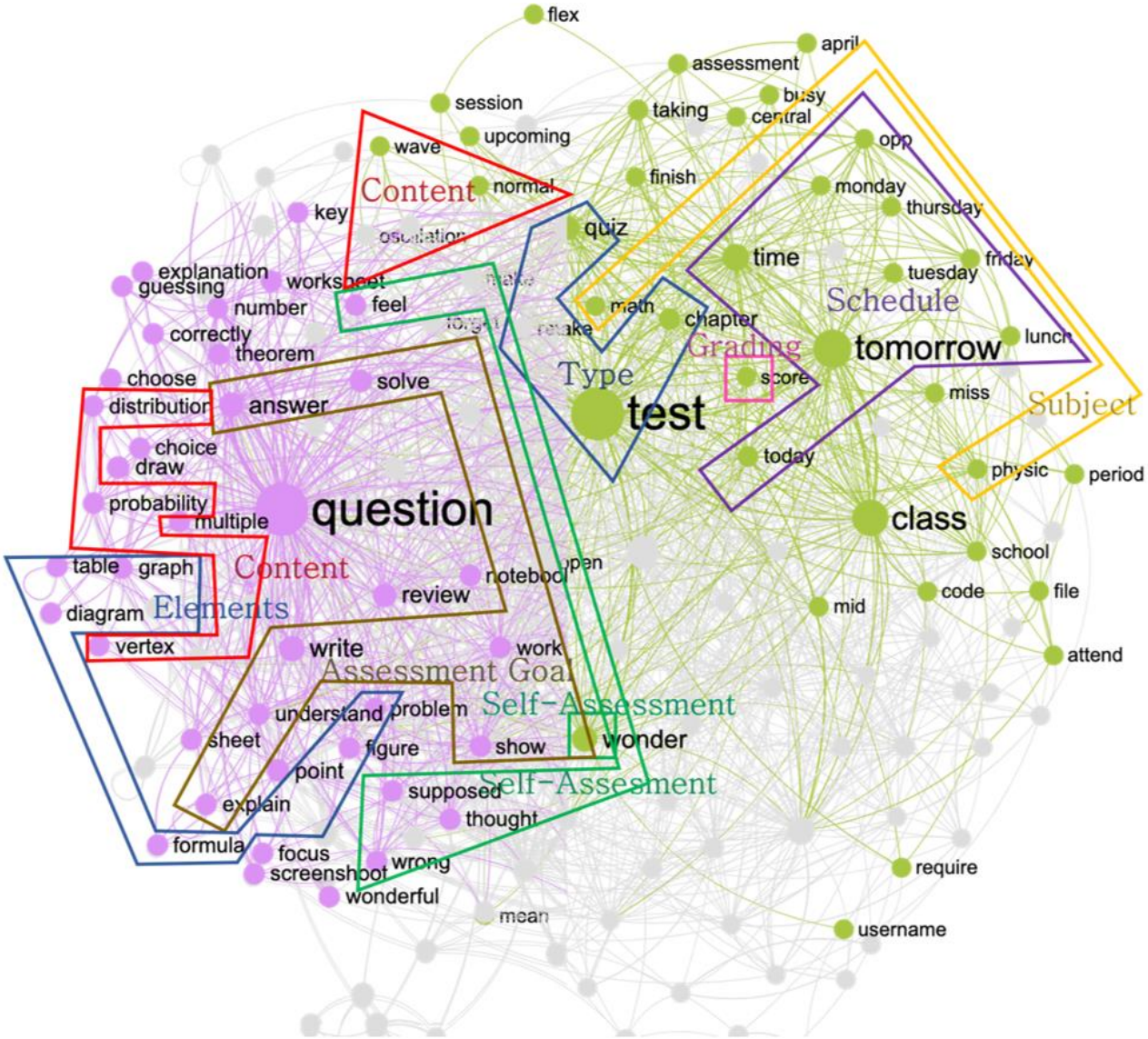

Figure 7. Contexts of question-answer-review and test-quiz-wonder topic clusters presented in the email network structure

The nodes included in the polygon of the category are the basis of categorization. The similar types for both topic clusters *question* and *tests* are *content* and *self-assessment* (self-reflection). It implies that students emailed the teacher due to concerns about the test coverage and how they think they would or have performed in an assessment.

The network graphs of the assessment-related topic clusters *question-answer-review* and *test-quiz-wonder* from Q1 to Q4 were run, analyzed, and coded to determine whether the categories presented in the previous section differ before COVID-19 and during COVID-19. Although the graphs are not shown in this report, **Table 2** summarizes the result of that analysis.

The method used to identify the assessment-related categories in the preceding graph was the same method applied to produce the results. Nodal words included in each category and the context of these words in the emails were the basis of categorization. Categories deem either *q* or *t*, or both, indicating either in the *question-answer-review, test-quiz-wonder*, or both.

The *goal* of the assessment, *subject* of the evaluation, *type* of the assessment, and *students' self-assessment* are the categories that appeared consistently in the network graphs before and during COVID-19. The categories are content, elements, schedule, and grading. The finding is the similarity between before and during COVID-19. Specifically, the *goal* category consistently appeared in the network graphs of *question-answer-review*, while the *subject* and *type* categories consistently appeared in the network graphs of *test-quiz-wonder*. A notable result is that the *self-assessment* category both appeared from Q1 to Q4 in the network graphs of both topic clusters. The *content* and *features* categories appeared in the *question-answer-review* network graph, while the *schedule* and *grading* categories appeared in the *test-quiz-wonder* network graph. Contrarily, some categories only appeared during COVID-19.

Table 2. Summary of the assessment-related categories revealed in the topic cluster network graphs before COVID-19 and during COVID-19

Categories	Before COVID-19		During COVID-19	
	Q1	Q2	Q3	Q4
Goal	*q, t*	*q*	*q*	*q*
Subject	*t*	*t*	*t*	*t*
Content	-	-	*q, t*	*q*
Type	*t*	*q, t*	*t*	*t*
Elements	*t*	-	*q*	*q*
Schedule	-	*q, t*	*t*	*t*
Grading	-	-	*t*	*t*
Self-assessment	*q, t*	*q, t*	*q, t*	*q, t*

Using the network graphs provided by InfraNodus, the structure of the graphs offered more insights into the assessment-related topic clusters *question-answer-review* and *test-quiz-wonder*. These topic clusters contain nodal words, questions, and tests, respectively. The categories that refer to assessment are the content, goal, schedule, subject, elements, type, and grading of the examination and the student's self-assessment or reflection about the evaluation. The content, elements, schedule, and grading categories are prominent during COVID-19.

After transitioning to the online class, students email the teacher regarding the lesson's content currently being discussed or evaluated. Since time constraints were meeting online, students would email regarding this matter to inquire about or request an assessment schedule due to the sudden online transition of classes. Lastly, students tend to ask about the test grading and quickly determine their scores during face-to-face meetings. In the last section, the emails represented by the topic clusters of nodes *question* and *test* determine the specific concerns and the corresponding support provided.

What Concerns About Teacher Support Do Students Have During COVID-19?

Based on the results, student emails increased during COVID-19 after the school transitioned to online classes. The students may have many inquiries, including the need for teacher support especially regarding assessments. The information supports the study of Naji et al. (2020), claiming that students need help with the type of learning activity and student assessments in the case of this study.

Indeed, teacher support is critical in the abrupt change of teaching mode - offline to online due to the pandemic. Students have become more concerned about assessments after the transition showing a need for the teacher to observe these students' needs and provide timely help.

The result in the previous sections on the increase in the email sending activity and the high percentages of topic clusters related to assessment may suggest that students have needed teacher support during COVID-19. This section reviews the emails by considering the topic clusters *question-answer-review* and *test-quiz-wonder*. These topic clusters contain the most influential keywords in *question* and *test*. The student emails were reviewed and explored to identify the concerns in general or specific to the assessment–the problems identified through systematization (Porras-Hernández & Salinas-Amuesca, 2013).

During COVID-19, emails related to the influential keywords *question* and *test* were retrieved using InfraNodus. These emails are the main essential statements linked with the topic clusters containing the clear nodes. Each message was analyzed to identify the corresponding concerns based on the context of the emails. These concerns are then classified into complementary support that the teacher provides the students. Although not shown in this report, the identified problems and the corresponding classification of supports were tabulated if observed in different course groups to determine what kind of support physics specifically and math course groups have. The summary of the result is in **Table 3**.

Specifically, the lower-level course group was more concerned about the test schedule, deadline, and conference. Moreover, their concerns were about an item, a topic, or a skill covered in assessments. The problems of the lower-course group were mostly about immediate feedback, grading, and accommodation. Except for the grading, these concerns are the higher-level course group's primary concern.

Lastly, both groups brought out the problem of giving access to any material posted online. The interpretations made for practical considerations and support were in the subject class–physics and math.

To aid learning in math, students would want clear instructions on the goal and schedule of the test. After any assessment, they would like to review the test items and learn how to answer them. If necessary, the students also look forward to receiving test accommodations and learning about grading.

To aid learning in physics, aside from the students would want to be well-informed on schedules, they would also like to have a suitable amount of time learning the material and have proper guidance in performing a lab.

Moreover, regarding the test, they also want to receive feedback about the answer or solution to any problem given, learn about their scores, and have accommodations. Most importantly, they highly demand technical support because both classes are taught online due to COVID-19. Meeting the technological support may help improve students' motivation, concentration, and skills during this online learning, as Naji et al. (2020) and Nambiar (2020).

Table 3. Summary of support and concerns for all course groups

CG		Concerns	Example emails
Lower-level	Algebra-2	Test goal clarification	*Hello, I'll send the pictures right now. However, we didn't learn 5.7 at school, and I remember you didn't require us to do 5.7 homework. 5.7 also didn't come out in the test.*
		Quiz item review & Answer clarification	*However, I think you missed one question for the quiz during the review, and I want to know why I got the answer wrong. I attached a screenshot of the question. It would be so helpful if you could answer today. (Because the test is tomorrow!) Or are we having a conference tomorrow to ask questions?*
		Grading inquiry	*However, I don't understand why you changed my grade for the midchapter from 12/13 to 11/14. Question number 2 is not answerable because it is not a multiple-choice question or a question with a quantifiable and measurable number of answers. I don't understand why you also took a point away for something you already awarded. I think my transcribed grade was wrong when you reopened question number 8, and instead of a 13/14, I received an 11/14.*
		Test accommodation	*I'm so sorry for just seeing this. I've been busy with AP testing and studying and haven't been feeling well these past few days. The quiz was quite difficult for me, and I wasn't feeling very well when taking it, so I think that played an effect. My mind went blank, and I couldn't focus after running out of time. Thank you so much for this opportunity, and I will send in my answers later today. Sorry for just responding. Thank you so much.*
		Online assignment completion and review access	*It says the test is 1:55 pm. Is this true, or is the test actually in class tomorrow? Hello! I hope this message finds you well. I wonder if you would be willing to open the UNIT 7 FRQ to complete it for submission and use it to review for upcoming testing. Thank you so much, sir! I thought it was due at the end of the day (today, Tuesday), not at 2 pm. I am sincerely sorry for the inconvenience this may cause you.*
	Physics	Inquiry lab questions	*I tried getting ahead of my work, and I looked through the inquiry lab and had a couple of questions.*
		Assignment submission schedule	*I was wondering if this review is due tonight or tomorrow night. It says tonight, but our test is later in the week, so I'm confused. Please let me know. Test will only be one question to question I just solved, right?*
		Conference schedule	*I'm very sorry. Will we have a conference, next class, as well? I promise to be on next time.*
		Test content questions	*I was wondering if you hadn't received my emails. I had some questions about the Oscillation and Waves test, and I could answer them. Here is the picture. If you see, I answered 8, 9, and 10 even though I don't know them well. I asked my parents about the study guide, but I also didn't get questions from them, but the paper where they explained it just memorized it about 30 times.*
		Answer clarification	*Could you send me the answer to the question you asked us to solve? I solved it, but I'm not sure if it's correct. Also, the test will only be one question similar to the question. I just cracked, right?*
		Grading inquiry	*I was looking over my grades for physics, and I realized I got a 98 for my Chapter Test: Gravity and Circular Motion. I asked my friend what he got, and he said he got a 100. However, I remember getting a 2% higher raw score than him (I got 94, and he got 92). Since the test is on Tuesday, will we have class tomorrow.*
		Test retake	*Do I have to retake the test I took today?*
		Online assignment completion & RA	*I saw that you opened it to see the overall score. Would it be possible for you to open the test to check the particular questions I got wrong?*
Higher-level	AP Physics 1	Discussion time limit & problem-solving help	*I'm going through the FRQs and MCQs right now, and I'm finding that I just can't do them. We spent so little time on these topics, and I have no idea how to answer some questions. Can we just work on them together? I really cannot do this.*
		Making up for a missed lab activity	*Because of this, I never really got a chance to work in the lab. I was absent from a class and had to make up for the test the next day. I don't even have a partner. What should I do?*
	AP Physics C	Test procedures	*Is the test tomorrow open notes?*
		Tests reschedule request	*I hope this message finds you well. I noticed that you have posted a test on oscillations for tomorrow and wondered whether we have a test tomorrow. Since we found out about the test today, we were surprised. Regardless of whether it is an open note test, a test still requires a great degree of preparation, and I wondered if you would be able to postpone the test so that the rest of the class could have ample time to prepare. It would be so helpful if you could answer them today. (Because the Test is tomorrow!) Or are we having a conference tomorrow to ask questions? Thank you for your time and see you tomorrow!*
		Missing test element request; feedback on the solution	*I am currently solving the chapter test right now. To receive feedback on our solutions on why things may be wrong or have an answer sheet posted on Schoology to look back, we can see the point distribution and what each question may be looking for. I encountered a question where I needed to refer to a figure to solve, but the question didn't provide the figure. Can you please show what the figure is to solve the question?*

Note. CG: Course groups

DISCUSSION

The literature shows that email communication is vital in online learning because it connects students with teachers to implore support. In this study, the students' email communications accurately reflect their words, thoughts, and actions. The researcher aims to report the same accuracy by using the entire schoology messages and interpreting any comments, ideas, and actions from the student's perspective. As such, topic modeling, a technique to analyze verbal data, was performed on

these emails. The researcher investigated the emails by topic modeling to determine the most influential topics or topic clusters. He also utilized the network graphs of the topic clusters based on the suggestion of Fougt et al. (2018).

Hence, for this present study, the researcher explored the emails sent by the students before and during COVID-19 through topic modeling and determined valuable information about the student's concerns. InfraNodus, a tool that can perform topic modeling and visualize the network graphs, was utilized in this study.

The study analyzed the student emails sent in 2019 when classes transitioned online due to the COVID-19 pandemic. Email communications became widely used since students could not communicate personally with the teacher, unlike before the transition. 623 emails sent by 70 students who took physics and math classes were analyzed using InfraNodus. InfraNodus is a tool that can perform topic modeling and visualize network graphs of verbal text data such as emails. Utilizing these two analyses answered the research questions.

The emails were explored and analyzed by topic modeling to identify the overall corpus's topic clusters. The grounded theory categorizes the topic nodes of the main topic clusters' network graphs related to assessment. Moreover, the study explained the differences between these topic clusters before and during COVID-19. Similarly, the study also described the differentiation of these categorizations before and during COVID-19. Furthermore, the graphs of the topic clusters related to assessment determined that the teacher supported the students in physics and math needed during the sudden transition to the online class.

The main topic clusters identified by topic modeling are *question-answer-review*, *class-time-homework*, and *test-quiz-wonder*. These topic clusters are all related to assessment. The question-answer-review among these three topic clusters increased by 6% before and during COVID-19. This result can be attributed to the increased email sending activity of the students to inquire about concerns on test schedule, feedback on the test, and asking for help for test preparations. Conversely, *class-time-homework* and *test-quiz-wonder* decreased by 4% and 9% before and during COVID-19. Mainly due to the shortened hours of virtual learning and the modifications in administering the assessment online after transitioning to an online class.

In addition, there was a strong connection between the *test-quiz-wonder* and *question-answer-review* topic clusters given by the bigram of 39; hence, these two topic clusters become the focus of the analysis. Grounded theory, the nodes in the graph categorized in assessments are similar to both topic cluster network graphs. They were *content* and *self-assessment*. It implies that student emails were generally concerned about the coverage of the evaluation or how they think they would or have performed in an assessment during the transition to an online class. Transitioning to an online class caused the students to email the teacher regarding the lesson's content studied. Email communication has increased since there were time constraints to discuss this matter online. Particularly, students sent emails to inquire about or request the assessment schedule. Most assessment schedules changed due to the sudden transition to an online class. Furthermore, the students emailed to inquire about the test's scoring and grades.

Finally, the increase in the students' email sending activity during COVID-19 was attributed to concerns regarding teacher support on assessments. These were identified by reviewing the student emails representing topic clusters and differed depending on the course group. The lower-level course group seemed concerned about an item, a topic, or a skill in the assessment, feedback, grading, and accommodation. Specifically, students want to be well-informed about the solution to a problem in a test, their test scores, and accommodations for the physics class.

Based on the results of this study, the topic modeling and the use of network graphs can allow the teacher to identify the student's central idea of email communications. Consecutively, the teachers can be aware of the student's concerns and support them accordingly. Thus, this paper showed that another type of verbal data, student emails, can be analyzed using topic modeling. The finding is different from the previous works that used students; more comprehensive written assessments (Pardo et al., 2018), chat data from collaborative learning (Cai et al., 2017), and scientific discourse in interactive multimedia presentations (Shaw et al., 2017), and student discourse in computer-supported collaborative learning (Csanadi et al., 2017).

Furthermore, this study provided detailed information on teacher support on assessment, a part of pedagogy, that advanced the definition of *supportive pedagogy* by Yates et al. (2020) in terms of evaluation. Hence, *supportive pedagogy* can now refer to clarifying the test goals, schedule, content, and procedures, reviewing the test solutions and answers, and providing necessary test accommodations to students affected by the transition to online learning due to the COVID-19 pandemic.

CONCLUSION

In the context of education, topic modeling helps understand the corpora of email communications made by students. Analyzing these email communications may help determine what kind of demand or support students need in the teaching-learning process, either physically or virtually. The topic modeling can identify the meaning of many text corpora. Hence, this will provide greater insight to science educators as to how they can optimize the teaching-learning process and

intentionally address the need of the students, especially in a virtual-learning setup. Since this study involved data network analysis, it may predict the student's need or demand for teacher support. Hence, this can be a startup of future works on automated feedback.

This study only explored and analyzed the student emails before and during COVID-19 through topic modeling. The survey considered a whole corpus, including emails from physics and math classes, and the findings apply to a general context. To understand the results in the context of physics, a more specific analysis using the emails of the physics students is a recommendation that may lead to a further understanding of the students' concerns on teacher support during an online class. Moreover, providing this teacher support to the student's performance may be investigated for future studies.

Author contributions: All authors have sufficiently contributed to the study and agreed with the results and conclusions.

Funding: No funding source is reported for this study.

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

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

REFERENCES

Alawamleh, M., Al-Twait, L. M., & Al-Saht, G. R. (2020). The effect of online learning on communication between instructors and students during the COVID-19 pandemic. *Asian Education and Development Studies, 11*(2). https://doi.org/10.1108/AEDS-06-2020-0131

Alghamdi, R., & Alfalqi, K. (2015). A survey of topic modeling in text mining. *International Journal of Advanced Computer Science and Applications, 6*(1), 147-153. https://doi.org/10.14569/IJACSA.2015.060121

Blei, D. M., & Lafferty, J. D. (2009). Topic models. In D. M. Blei, & J. D. Lafferty (Eds.), *Text mining* (pp. 101-124). Chapman and Hall/CRC. https://doi.org/10.1201/9781420059458-12

Cai, Z., Eagan, B., Dowell, N., Pennebaker, J., Shaffer, D., & Graesser, A. (2017). Epistemic network analysis and topic modeling for chat data from the collaborative learning environment. In X. Hu, T. Barnes, A. Hershkovitz, & L. Paquette (Eds.), *Proceedings of the 10th International Conference on Educational Data Mining* (pp. 104-111). EDM Society.

Charmaz, K., & Bryant, A. (2010). *The SAGE handbook of grounded theory*. SAGE. https://doi.org/10.1016/B978-0-08-044894-7.01581-5

Csanadi, A., Eagan, B., Kollar, I., Shaffer, D. W., & Fischer, F. (2018). When coding-and-counting is not enough: using epistemic network analysis (ENA) to analyze verbal data in CSCL research. *International Journal of Computer-Supported Collaborative Learning, 13*(4), 419-438. https://doi.org/10.1007/s11412-018-9292-z

Dickinson, A. (2017). Communicating with the online student: The impact of email tone on student performance and teacher evaluations. *Journal of Educators Online, 14*(2), 1-10. https://doi.org/10.9743/jeo.2017.14.2.5

Diehl, A., Abdul-Rahman, A., Bach, B., El-Assady, M., Kraus, M., Laramee, R. S., & Chen, M. (2022). Characterizing grounded theory approaches in visualization. *arXiv:2203.01777*. https://doi.org/10.48550/arXiv.2203.01777

Dürscheid, C., Frehner, C., Herring, S. C., Stein, D., & Virtanen, T. (2013). Email communication. *Handbooks of Pragmatics,* (9), 35-54. https://doi.org/10.1515/9783110214468.35

Fougt, S. S., Siebert-Evenstone, A., Eagan, B., Tabatabai, S., & Misfeldt, M. (2018). Epistemic network analysis of students' longer written assignments as formative/summative evaluation. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (pp. 126-130). ACM. https://doi.org/10.1145/3170358.3170414

Gibbs, S. (2016). How did email grow from messages between academics to a global epidemic? *The Guardian*. https://www.theguardian.com/technology/2016/mar/07/email-ray-tomlinson-history

Glaser, B. G., & Strauss, A. L. (2017). *Discovery of grounded theory: Strategies for qualitative research*. Routledge. https://doi.org/10.4324/9780203793206

Glesne, C. (1999). *Glesne, Corrine, becoming qualitative researchers: An introduction*. Longman.

Kayumova, L. R., Gainullina, L. N., Akhmadieva, R. S., Matvienko, V. V., & Kabakhidze, E. L. (2021). Using interactive platform "round" to organize online leisure activities for children during the pandemic. *EURASIA Journal of Mathematics, Science and Technology Education, 17*(10), em2016. https://doi.org/10.29333/ejmste/11182

Maphosa, C., Van Den Berg, G., & Mudau, P. K. (2021). Assessment of the perceived usefulness of mobile phone technology for communication in learning by distance education students in a rural-based university. *African Perspectives of Research in Teaching and Learning, 5*(2), 45-61.

Martin, A. (2020). How to optimize online learning in the age of coronavirus (COVID-19): A 5-point guide for educators. *UNSW Newsroom, 53*(9), 1-30.

Miller, M. A. (2020). The importance of tone and attitude in email and the online classroom. In C. N. Stevenson, & J. C. Bauer (Eds.), *Enriching collaboration and communication in online learning*

communities (pp. 52-67). IGI Global. https://doi.org/10.4018/978-1-5225-9814-5.ch004

Naji, K. K., Du, X., Tarlochan, F., Ebead, U., Hasan, M. A., & Al-Ali, A. K. (2020). Engineering students' readiness to transition to emergency online learning in response to COVID-19: Case of Qatar. *EURASIA Journal of Mathematics, Science and Technology Education*, *16*(10), em1886. https://doi.org/10.29333/ejmste/8474

Nambiar, D. (2020). The impact of online learning during COVID-19: Students' and teachers' perspectives. *The International Journal of Indian Psychology, 8*(2), 783-793.

Paranyushkin, D. (2019). InfraNodus: Generating insight using text network analysis. In *Proceedings of the World Wide Web Conference* (pp. 3584-3589). https://doi.org/10.1145/3308558.3314123

Porras-Hernández, L. H., & Salinas-Amescua, B. (2013). Strengthening TPACK: A broader notion of context and the use of teacher's narratives to reveal knowledge construction. *Journal of Educational Computing Research*, *48*(2), 223-244. https://doi.org/10.2190/EC.48.2.f

Shaffer, D. W., Collier, W. & Ruis, A. R (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics, 3*, 9-45. https://doi.org/10.18608/jla.2016.33.3

Shaw, E., Phillips, R., Kao, A. & Torres, R (2017). Using topic modeling to assess middle school science discourse based on next generation science standards. In *Proceedings of the 2017 International Conference on Advanced Technologies Enhancing Education* (pp. 42-45). Atlantis press. https://doi.org/10.2991/icat2e-17.2016.11

Sintema, E. J. (2020). Effect of COVID-19 on the performance of grade 12 students: Implications for STEM education. *EURASIA Journal of Mathematics, Science and Technology Education*, *16*(7), em1851. https://doi.org/10.29333/ejmste/7893

Strauss, A., & Corbin, J. M. (1997). *Grounded theory in practice*. SAGE.

Yang, Y., & Cornelius, L. F. (2004). Students' perceptions towards the quality of online education: A qualitative approach. In *Proceedings of the 27th Association for Educational Communications and Technology* (pp. 861-877).

Yates, A., Starkey, L., Egerton, B., & Flueggen, F. (2020). High school students' experience of online learning during COVID-19: The influence of technology and pedagogy. *Technology, Pedagogy and Education, 30*(1), 59-73. https://doi.org/10.1080/1475939X.2020.1854337

Yun, E. (2020). Review of trends in physics education research using topic modeling. *Journal of Baltic Science Education, 19*(3), 388-400. https://doi.org/10.33225/jbse/20.19.388

Zhang, W. Y., & Perris, K. (2004). Researching the efficacy of online learning: A collaborative effort amongst scholars in Asian open universities. *Open Learning: The Journal of Open, Distance and e-Learning, 19*(3), 247-264. https://doi.org/10.1080/0268051042000280110

Zhou, S., Gu, H., & Yao, J. (2018). From "strict learning" to "smart learning": An analysis of the key factors affecting students' academic performance: An empirical study based on the educational quality monitoring data of 262245 students in Jiangsu province. *Primary Middle School Management, 32*(11), 39-42.