

Journal Article

Text Mining as an Educational Evaluation Methodology: Analyzing Textual Data Extracted from Online Learning Environments

 **Manuel B. Garcia** ^{a,b,c} *

^a College of Education, University of the Philippines Diliman, Quezon City, Philippines

^b Educational Innovation and Technology Hub, FEU Institute of Technology, Manila, Philippines

^c College of Education, Korea University, Seoul, South Korea

* Correspondence:

Manuel B. Garcia, University of the Philippines Diliman and FEU Institute of Technology.
mbgarcia@feutech.edu.ph

How to cite this article:

Garcia, M. B. (2025). Text Mining as an Educational Evaluation Methodology: Analyzing Textual Data Extracted from Online Learning Environments. *Learning Environments Research*.
<https://doi.org/10.1007/s10984-025-09547-y>.

Article History:

Received: 5 Mar 2024
Revised: 8 Jan 2025
Accepted: 3 Jul 2025
Published: 20 Jul 2025

Abstract:

The rise of digital platforms has led to a massive influx of textual data. While traditional textual analysis techniques have been effective, analyzing large datasets is becoming impractical due to the required time and resources. To demonstrate the usefulness of text mining as an alternative, this study analyzed data extracted from an emergency remote learning (ERL) environment. Free-form responses from a series of cross-sectional surveys (2020-2022) were analyzed using word frequency, collocation, concordance, topic modeling, and sentiment analyses. According to the findings, the most frequently occurring unigram and bigram in the text corpus were "hard" and "mental health", respectively. Three primary themes based on lived experiences were identified, namely individual, academic, and technological challenges, and another three themes emerged from coping strategies, including entertainment, relationship, and health-related mechanisms. Negative sentiment toward the ERL setup was also evident in the text corpus. Overall, the combination of text mining techniques allowed for a comprehensive exploration of the linguistic features of the corpus and provided a multifaceted understanding of the selected phenomenon. Consequently, this study endorses the use of text mining as a methodology for analyzing large volumes of textual data.

Keywords:

Online Learning Environment, Emergency Remote Learning, Text Mining, Student Experience, COVID-19



This is a pre-copyedit version of an article copied from <https://manuelgarcia.info/publication/text-mining-education> and published in the *Learning Environments Research* journal. The final authenticated version is available online at <https://doi.org/10.1007/s10984-025-09547-y>. Any other type of reproduction or distribution of the article is not authorized without written permission from the author and publisher.

INTRODUCTION

Social scientists and humanities scholars have a long history of employing textual analysis in their research projects. This methodology has been recognized as a powerful tool for investigating the social world and understanding human behavior (e.g., Wanniarachchi et al., 2022; Yang & Sheu, 2019). In the context of qualitative research, Kuckartz (2014) defined textual analysis as the process of deriving meaningful information expressed within a document using data analysis techniques commonly applied in social sciences disciplines. Examples of such techniques include classical content analysis, grounded theory, and thematic analysis. Despite the abundance of empirical evidence supporting the effectiveness of these techniques, it has become increasingly impractical to analyze data using a manual text analysis approach (Rahman, 2017). For example, vast amounts of unstructured textual data became available for analysis due to the rise of the Internet, social media networks, and other digital communication channels (Alcober et al., 2020; R. P. Bringula et al., 2022; Macanovic, 2022). Utilizing traditional manual methods of text analysis in such large datasets is nevertheless time-consuming and resource-intensive. Researchers are consequently recognizing the viability of computational text analysis (Wiedemann, 2016), especially when compared to manual methods (Syyrilä et al., 2021).

Computational text analysis, or the use of computer-assisted techniques in analyzing textual data, is often referred to as text mining. Ignatow and Mihalcea (2017) described text mining as a process of extracting consequential patterns from unstructured texts using techniques commonly applied in natural language processing, machine learning, information retrieval, sentiment analysis, named-entity recognition, concordance analysis, and computational linguistics. Examples of such techniques include keyword extraction, topic modeling, and concordance identification. Unlike textual analysis, text mining leverages the power of computational methods to automatically analyze large volumes of text data. For instance, a study conducted by Garcia (2020) analyzed 65,396 tweets related to the COVID-19 pandemic using sentiment and emotion analyses. The study highlights how text mining techniques can be used to quickly and accurately analyze large volumes of textual data that would be difficult, if not impossible, to scrutinize manually. Unsurprisingly, many researchers are starting to support various manual approaches to text analysis with text mining techniques (Macanovic, 2022; Wiedemann, 2016). Other researchers even explored the compatibility of using these text analysis approaches concomitantly (Inaba & Kakai, 2019; Muller et al., 2016; Yu et al., 2011).

Like other websites that are rich in texts (e.g., social media platforms), online learning environments also generate large volumes of textual data that can be analyzed using text mining techniques (Wahyono et al., 2021). These environments provide a variety of text-based communication channels (e.g., messaging systems and discussion forums) and allow students to produce textual data in the form of their course submissions (e.g., essays and assignments). Despite the potential of these platforms, researchers often tend to rely on social media networks as the source of data for text mining when analyzing phenomena related to online education (e.g., Aydin, 2021; Bozkurt, 2021; Zhou & Mou, 2022). By solely relying on social media data,

researchers may be missing out on valuable insights that can only be gleaned from data generated within online learning platforms. For instance, the shift to online learning modality during the COVID-19 pandemic, also known as emergency remote learning (ERL), resulted in a significant increase in the volume of textual data generated within these environments (Fung et al., 2022). This scenario exemplifies how online learning environments can provide a fertile ground for applying text mining techniques to extract valuable educational insights and enhance learning experiences. Unfortunately, Bond (2020) reported that online surveys were the primary data collection tools, and quantitative nonexperimental designs were the main research methods used during the implementation of ERL. The application of text mining techniques in analyzing large qualitative data within these learning environments could have potentially extracted valuable insights for improving the overall learning experience of students in times of crisis.

This study explores the application of text mining techniques to analyze large-scale textual data generated in online learning environments. Given the substantial volume of data produced on these platforms during the COVID-19 pandemic, the period of ERL implementation presents an optimal source for such an analysis. From 2020 to 2022, our students participated in a series of cross-sectional surveys containing quantitative and free-form questions regarding their ERL experience. The findings from the quantitative data analysis have already been presented in a separate study (Garcia & Revano, 2022), allowing this research to focus specifically on the qualitative aspects of the data. With this pseudo-longitudinal dataset, not only is there an opportunity to demonstrate the capabilities of text mining techniques, but the study also addresses the limitation in online learning literature of not using data from different time points (e.g., Mirahmadizadeh et al., 2020; Syyrilä et al., 2021). Additionally, it expands the relatively scant body of research that has been conducted on the analysis of free-form comments in comparison to quantitative responses in student surveys (Alhija & Fresko, 2009; Brockx et al., 2012). This expansion is vital since this deficiency leaves a significant gap in our understanding of online learning environments and experiences, particularly as qualitative data are a rich but underutilized source of feedback crucial for the quality assessment of educational processes and outcomes (Gakhal & Wilson, 2019; Hoon et al., 2015). Given the importance of learning environments in shaping educational outcomes (Cayubit, 2022), understanding these settings is essential. Insights derived from their large-scale qualitative data, which is often underanalyzed, can inform more effective educational strategies and interventions. Therefore, this study will utilize text mining techniques to answer the following research questions (RQ):

- RQ1. Using word frequency and collocation analyses, what are the most frequent unigrams and bigrams in student feedback on their ERL experiences?
- RQ2. Using Key-Word-In-Context (KWIC) concordance analysis, what are the underlying meanings behind the most frequently occurring words in the corpus?
- RQ3. Using latent Dirichlet allocation (LDA) topic modeling, what are the emergent topics from students' lived experiences and coping strategies?
- RQ4. Using lexicon-based sentiment analysis, what is the general polarity of free-form responses generated in and extracted from an ERL environment?

BACKGROUND OF THE STUDY

Online Learning Environments as a Data Goldmine

A learning environment refers to the setting or context in which education occurs, encompassing physical classrooms, digital platforms, and hybrid systems that blend both. In educational research, this domain is a critical area of focus because it significantly influences learner engagement, cognitive development, and educational outcomes (Cayubit, 2022; Rusticus et al., 2023). Theories of educational psychology, such as constructivism, emphasize the importance of the environment in shaping how students interact with content and each other, as well as how it affects their learning processes. With the increasing popularity of digital platforms during the COVID-19 pandemic, online learning environments have become a primary area of interest for educational researchers (Garcia, 2022; Ong & Quek, 2023; Salta et al., 2022). Driven by the necessity to continue educational operations remotely (Lin & Yeh, 2022), schools at various levels have adopted these platforms rapidly. This widespread adoption has made online learning environments a rich source of diverse data types, including, but not limited to, real-time interaction data, learning analytics, and feedback mechanisms. These data are collected through various activities such as gamification (Valderama et al., 2022), video lectures (Garcia & Yousef, 2023), interactive simulations (Ari et al., 2022), virtual classrooms (Islam et al., 2023), and more (Hadad et al., 2024). Each activity provides a detailed picture of student learning behaviors and engagement patterns. Therefore, these environments are goldmines for conducting empirical studies and testing educational theories in situ. The ability to gather and analyze such expansive data sets allows for a more comprehensive understanding of the effectiveness of various online teaching methods as well as student experience within virtual learning environments.

Student Experience in Emergency Remote Learning

The abrupt transition to online learning has had a considerable impact on the educational landscape, establishing ERL as a critical subject for scholarly investigation. This shift has unveiled unique challenges and opportunities, drawing research interest toward the impact on student interaction outside the traditional classroom environment. As educational institutions navigated this transition, they continually gathered feedback from their students. According to Aldridge and Bianchet (2022), leveraging student feedback about their learning environments is crucial for enabling students to participate actively in shaping and enhancing their educational settings. With these initiatives, there is a wealth of unstructured textual data available that can be effectively analyzed using text mining techniques. Unfortunately, analyses of student experiences during ERL have predominantly employed quantitative research methods, as evidenced by a systematic review (Bond, 2020). Wong (2015) asserted that although quantitative data provides a measurement of reality, qualitative data in the form of students' own words provides valuable insight into the "why" of their lived educational experiences (Tremblay et al., 2021; Vindrola-Padros et al., 2020). Other researchers have thereby proposed a more comprehensive approach

to investigating student experience. For example, Klemenčič and Chirikov (2015) emphasized qualitative research methods as a viable tool for disentangling the complexities of the student experience in higher education. These methodologies provide authentic and trustworthy accounts of real-life experiences, the circumstances in which they occur, and the significance they hold for students that cannot be easily comprehended from numerical data (Gakhal & Wilson, 2019; Peñarrubia-Lozano et al., 2021; Rahman, 2017). Even in the causal explanation of events where a dismissal of qualitative research has been particularly virulent, Maxwell (2012) argued that it could make an essential contribution to causal inquiry in education by capturing the richness of individual experiences, interpretations, and meanings. Nevertheless, Klemenčič and Chirikov (2015) noted that qualitative research can be labor-intensive as it requires significant time and effort for both data collection and analysis. Exploration of digital approaches for qualitative data analysis is warranted to accommodate large datasets (Syyrilä et al., 2021; Wiedemann, 2016).

Text Mining Techniques for Qualitative Data Analysis

In qualitative research, key challenges such as efficiency and accuracy arise when analyzing large amounts of unstructured textual data. Traditional manual analysis methods can be time-consuming and often involve subjective judgments by the researchers, leading to potential bias in the analysis (Galdas, 2017; Klemenčič & Chirikov, 2015). Fortunately, with the rapid increase in digital content and the development of advanced computational techniques, text mining has emerged as a promising approach to tackle these challenges (Garcia & Cunanan-Yabut, 2022; Macanovic, 2022). Text mining techniques provide a promising alternative by allowing researchers to process vast amounts of qualitative data efficiently and systematically. This approach also offers several advantages over manual analysis methods, including greater efficiency, reduced researcher bias, and the ability to identify patterns, themes, and relationships that may not be immediately apparent in traditional methods (Ignatow & Mihalcea, 2017). Many researchers (e.g., Syyrilä et al., 2021; Wiedemann, 2016) have therefore underscored computational text analysis as a valuable tool for identifying patterns, themes, and relationships that may not be immediately apparent through manual analysis.

In the analysis of textual data, scholarly literature recommends various techniques. First, frequency-sorted word lists (i.e., word frequency) have been an essential process in computational text analysis due to their valuable representation of meaning for various purposes, such as text categorization and information retrieval. Word frequency refers to the number of times a word appears within a corpus of texts. Nonetheless, relying solely on this approach does not provide insight into the context in which a given term is utilized within the text (Ignatow & Mihalcea, 2017). Other studies have suggested collocation analysis as a complementary technique to word frequency as it can identify significant associations between words and reveal language use patterns (Cordeiro, 2019). Collocation analysis is a process of identifying the co-occurrence of specific words based on the idea that those that commonly co-occur have a certain level of

semantic relatedness. Meanwhile, O'Donnell (2008) recommends KWIC concordance analysis to facilitate the detection of lexicogrammatical patterns. This text mining technique presents a list of all occurrences of a specific word or phrase, accompanied by the words and phrases that appear before and after it. The computational method of topic modeling is also recommended for the identification of word usage patterns and underlying themes by analyzing the semantic structure of the text. This unsupervised machine learning technique has been used in the analysis of different corpora, such as online reviews (Kwon et al., 2021) and political speeches (Miranda & Bringula, 2021). In the comparative analysis of topic modeling methods, Albalawi et al. (2020) discovered that LDA is one of the two methods that generate the most valuable outputs with diverse ranges and meanings. Sentiment analysis may also be leveraged to extract more human elements from a text corpus (Garcia, 2020). Although each text mining technique has its unique strengths and benefits, their combination has the potential to enhance the quality and depth of the analysis, leading to a more comprehensive understanding of the results.

METHODS

Research Design

Within the field of social science text mining research, Ignatow and Mihalcea (2017) provided a framework that outlines five key design decisions for researchers to consider: research type, level of analysis, mode of analysis, data selection, and inferential logic. First, this study combined the philosophy of idiographic and nomothetic approaches for the research type. The idiographic approach was utilized to acquire an in-depth understanding of the unique experiences of students in the ERL environment. Concurrently, the nomothetic approach was also used to identify common patterns and themes within the dataset. The combination of idiographic and nomothetic approaches allowed for comprehensive data analysis (Beltz et al., 2016). Second, the sociological level of analysis was selected in recognition of the critical role played by the social spaces in which the texts were produced. This decision was made to illuminate the connections between the textual data and the societal context in which they were situated (Nguyen et al., 2020). As noted by Ignatow and Mihalcea (2017), analyzing texts as social information provides valid and relevant insights into social reality. Third, a mixed-methods research design was adopted as the mode of analysis since the research questions necessitated both computational linguistics and qualitative data analysis techniques. According to Creswell (2014), most text mining research projects are best comprehended as mixed methods. Fourth, pre-existing survey data on student experience during ERL was strategically selected to avoid incurring additional time and cost. Using existing data is a commonly observed practice within the domain of text mining research (e.g., Alsayat & Ahmadi, 2022). Finally, abduction was selected as the inferential logic as it allows to generate and test explanations or hypotheses using both qualitative and quantitative data, which aligns with the research questions and the chosen mixed-methods research design.

Data Collection and Procedure

The data were collected from cross-sectional surveys conducted at the end of every trimester in a Philippine higher education institution. These surveys focused on student experiences during ERL and have been conducted from the academic year 2020-2021 until 2022. The survey questionnaire was distributed via the institution's online learning platform as a course assignment. It included both quantitative and qualitative items revolving around their challenges, lived experiences, and coping strategies. The results of the quantitative data analysis were previously reported in a separate study (Garcia & Revano, 2022). For the qualitative section, the open-ended questions asked were: (1) "What challenges did you encounter during online learning in the pandemic?" (2) "How would you describe your overall experience with online learning?" and (3) "What strategies did you use to cope with these challenges?" Only free-form responses were extracted, anonymized, and exported from a database to a text file (original text file). Contents of the original text file were then copied, pre-processed, and saved as another text file (cleaned text file). Data pre-processing methods (e.g., text cleaning, lemmatization, and tokenization) were adopted from the studies of Garcia (2020) and R. Bringula et al. (2022), depending on the specific text mining technique employed. Finally, the original text file was used to report complete feedback on the paper, while the cleaned text file was used for text mining.

Data Analysis and Techniques

The data analysis techniques employed in this study are tailored to the research questions. For RQ1, the most relevant terms were extracted using word frequency analysis. In linguistics, the frequencies of words are used to examine how people communicate (Lijffijt et al., 2011). The statistical measure term frequency-inverse document frequency was used to evaluate how relevant a particular word is in the text corpus. Distinctive keywords from each question were also identified and visually presented using word clouds. Meanwhile, words that were used in different contexts (i.e., words sense disambiguation) were noted for further analysis (e.g., "*I need **support** to learn the lesson*" and "*My parents need **support** for me to continue college*"). Collocation, or words that commonly co-occur, was also performed to identify hidden semantic structures by counting bigrams as one word. As a linguistic concept, collocation is considered a crucial component of semantic analysis (Barnbrook et al., 2013). For RQ2, the word context of the most frequent unigrams was examined using 15-word concordances (left and right) of the keyword of interest. By looking at the surrounding words, it becomes easier to interpret the theme of a search term in a corpus. KWIC concordance analysis can also assist in detecting lexicogrammatical patterns (O'Donnell, 2008). For RQ3, LDA was used to explore the inferred topics and themes further. LDA is a topic modeling technique that uncovers broader underlying themes and patterns within the entire corpus of qualitative data. Evidence suggests that LDA is superior to other statistical topic models in text mining (Liu et al., 2011). This analysis was conducted twice to cover both lived experiences and coping strategies. For RQ4, a sentiment

analysis using a lexicon-based approach was performed to measure the polarity behind the textual data. According to Garcia and Cunanan-Yabut (2022), sentiment analysis can capture an overview of public opinion on various topics. The entire data analysis was carried out using R programming.

RESULTS

Of the 1,246 students who participated in the surveys, 978 textual responses (78.49%) were extracted and analyzed. These textual responses on online experience and coping strategies comprised 125,636 words (128.46 words per student and 26.2 words per sentence) and 95,267 words (97.41 words per student and 21.7 words per sentence), respectively. In terms of language, most students responded using English except for the 94 instances of 24 Filipino words (e.g., “*mabagal*” or slow, “*trabaho*” or employment, “*pera*” or money).

RQ1: Using word frequency and collocation analyses, what are the most frequent unigrams and bigrams in student feedback on their ERL experiences?

Table 1 presents the ten most frequent words in the corpus. Interestingly, “hard” was the most frequent word, indicating the challenging situation students faced during ERL. As has been recently reported by Treceñe (2022), students faced difficulties (e.g., adapting to the new learning mode) when attending online classes. Overall, these results suggest that students' feedback on their ERL experiences was characterized by a focus on a variety of topics, including the challenges and difficulties they encountered, the importance of technology and online resources, concerns about health and well-being, and the impact of ERL on personal relationships.

Table 1. The ten most frequent unigrams in the text corpus

Rank	Word	Frequency	Sample Feedback
1	hard	689	Sometimes I have a hard time catching up to lessons especially with many workloads stacking up.
2	internet	654	Very worrisome and stressful – there's a lot of inconsistencies and you'd always worry about your Internet , deadlines and such.
3	health	591	I've found myself at the end of the line a few times, my mental health got worse and I've experienced breakdowns more often.
4	work	457	Family business got affected and my mother lost her work because of cost cutting it was hugely an adjustment for us because tuition fees were still expensive as usual.
5	sleep	431	While I get to sleep often and do my usual routine which is play games, attend class during school weeks and all that- I feel very pressured and anxious 24/7 because of online classes itself.

circumstances than specifically with their educational experiences (e.g., “*My life during this pandemic has not been the easiest and it is **hard** because I feel anxious*” [Student 25]). This observation may provide a possible explanation for the observed decline in academic performance among students during the pandemic (Engzell et al., 2021). Students exhibited greater concern for daily survival than their academic studies (e.g., “*Our status in **life** is now different because we need to earn more money plus it is **hard** catching up to lessons especially with many workloads stacking up*” [Student 225]). Excessive workload intensified this dilemma since students’ ability to manage time was already a recurring problem when learning online even before the pandemic, as confirmed by a qualitative research synthesis (Blackmon & Major, 2012).

Further insights into the reasons for this disparate impact on students can be gleaned from the subsequent unigrams in the list. Notably, the term “slow” was mentioned in 89% of occurrences of the term “internet,” suggesting that bandwidth and connectivity issues may be significant challenges for students (e.g., “*Our **internet** is too **slow** so I cannot participate during online class*” [Student 66]). This finding is critical because the internet connectivity experience directly influences students’ behavioral intention to use online learning. Poor internet access could lead to reduced opportunities to participate actively in online classes and less time to complete online activities. Consequently, researchers have suggested that educational institutions should adjust their online delivery methods in response to these technological constraints that many students face (Cullinan et al., 2021). Another possible explanation for the challenges faced by students during ERL relates to their well-being, with the term “mental” being mentioned in 95% of occurrences of the word “health” (e.g., “***Mental health** is deteriorating day by day but still trying to survive, it’s hard but this won’t last forever so I try to just go with the flow.*” [Student 531]). Furthermore, “mental health” emerged as the most frequent bigram in the corpus. In contrast, the word “physical” was only mentioned in 25% of occurrences of the term “health.” These findings suggest that mental health may be a more salient issue for students in ERL than physical health (e.g., “*I hope everyone values everyone’s mental **health** because it’s as important as our **physical.***” [Student 243]). Researchers such as Lai et al. (2020), Yang et al. (2021), and Lee et al. (2021) noted the adverse impacts of ERL on students due to a heavy academic workload and a sense of disconnection from the school community, among others.

Regarding the term “work,” it was observed that it was used in two distinct contexts, namely school deliverables and employment. Among the top unigrams, the words “group” (e.g., “*still hard doing **group work** online when you do not know what the groupmates are dealing with*” [Student 229]), “school” (e.g., “*All of my **work** consists of doing it on the computer and not on the **school** itself*” [Student 428]), and “activity” (e.g., “*Being in a physical classroom makes me only focus on my **work** which is not case with **activity** via online*” [Student 85]) were mentioned in 52%, 41%, and 32% of the occurrences of “work,” respectively. Meanwhile, the unigrams “freelance” (e.g., “*I realize that even though I’m still a student, I should think or plan for another sidelines like*

part time jobs or freelance work [Student 711]) and "job" (e.g., "*Family business got affected and my mother lost her **job** so she looked for extra work*" [Student 391]) were identified in 42% and 67% of occurrences of the term "work," respectively. In addition to the previously established issue of academic workload, these findings highlight the adverse economic effects of the pandemic on students' and their parents' labor market participation. As a result, many students have experienced the loss of internships, jobs, or job offers, and those students from lower-income backgrounds have been more likely to postpone graduation than their higher-income peers (Aucejo et al., 2020). The subsequent unigram in the list, "sleep," may have been influenced by the term "work," as students mentioned the term "quality" in 76% of occurrences (e.g., "*My sleeping schedule is off, and I can say that I do not get a **quality sleep** anymore*" [Student 236]). This finding highlights the prevalence of poor sleep quality among students during the pandemic, which is a critical issue as sleep quality has been shown to mediate the relationship between perceived stress and dietary behaviors (Du et al., 2021). Overall, these insights have important implications for educators and institutions seeking to support students in the transition to ERL.

RQ3: Using latent Dirichlet allocation (LDA) topic modeling, what are the emergent topics from students' lived experiences and coping strategies?

Upon investigation, it became apparent that certain words were utilized in varying contexts within the corpus. For instance, the unigram "support" was used to describe both financial assistance (e.g., "*I'm torn between thinking of wanting to graduate as soon as I can to **support** my mom who is our main financial support and thinking of just stopping wasting too much money*" [Student 623]) and educational guidance (e.g., "*I would say that the **support** from my teachers helped me tremendously when it comes to learning while being in the comfort of my home and time*" [Student 125]). LDA topic modeling was employed to identify emergent topics. Additionally, due to the distinctiveness of the lived experiences keywords identified through KWIC analysis, another analysis was conducted to cover the coping strategies as well.

Table 2. Inferred topics from lived experiences

Topic	Words	Sample Sentences	Label
1	internet, wi-fi, computer, poor, connection, access, data, mobile, stable, issue	<ul style="list-style-type: none"> • "It is a little bit harder than face-to-face classes because of internet providers in the Philippines. Almost every day I experience internet problems even though we already pay for the best internet deals in the market." • "In my personal experience, it is actually stressful because I sit and stare at my computer screen every day for more than 16 hours. If we only have a stable connection, I can finish my tasks as early as possible, and I do not need to spend extra money for mobile data." 	Technological Factors

		<ul style="list-style-type: none"> • “My experience during the pandemic is really hard especially at the beginning. I do not have the necessary software and hardware specifications for my computer to perform my activities. But I am still thankful because I know people who do not even have Internet at home or own device.” 	
2	class, teacher, school, attend, project, time, instructions, hard, subject, good	<ul style="list-style-type: none"> • “very tiring and draining especially if you have class from 7-am to 9pm its easy to attend classes and to learn but and its still very different to learn when it was face to face in school.” • “At first, it was okay due to no specialty classes so it was very clear-cut, attend class, do formatives, wait for summatives, repeat. But now it really just makes me lazier during the comfort of my own home and it's still hard to approach people online for group works and even our teacher, still hard doing group work online when you do not know what the groupmates are dealing with especially with the capstone project coming.” • “Everyday feels the same. Boring but at the same time stressful because I kept thinking about the online class, whether or not I can pass the semester.” 	Academic Factors
3	hard, stress, anxiety, mental, lack, health, fear, survive, cope, physical	<ul style="list-style-type: none"> • “Between money issues, mental and physical health, life during COVID-19 is uniquely difficult for everyone to survive. And I'm still one of the ‘lucky ones’. It is incredibly draining. I am tired and my anxiety is always ready to attack me.” • “Hard to cope up especially as an IT student we need to do hands-on activities like coding or other computer tasks since most of us learn a lot during face to face classes.” • “It is very challenging and stressful, the environment is also a big factor on contributing with the stress level, everyday just feels like I have to work on submissions like assignments and assessments nonstop without even learning.” • “I have now grown to be familiar with the fear of mortality and its lasting effect on some people, as well as its lack thereof on others. I feel like I'm somehow comfortable already.” 	Individual Factors

Results from the LDA analysis indicate that the challenges experienced by students during the ERL implementation can be summarized into three distinct themes. As shown in Table 2, the topics revolved around technological factors (e.g., unreliable internet access and lack of devices), academic factors (e.g., heavy workload and request for academic calendar cancellation), and individual factors (e.g., isolation, mental health difficulties, financial distress within the household, and need for a part-time job). These outcomes align with the contextual interpretations obtained from the KWIC analysis, which supports their validity. These findings demonstrate that some of the challenges experienced by students were not necessarily a direct result of ERL but instead

emerged due to the pandemic situation. Therefore, educational leaders can still leverage ERL as a platform to offer academic support programs that help students overcome these challenges.

Table 3. Inferred topics from coping strategies

Topic	Words	Sample Sentences	Label
1	games, facebook, netflix, boring, tutorial, interaction, playstation, movies, youtube, fun	<ul style="list-style-type: none"> • “My life became quite boring if I’m being honest. I have less interaction with people that I used to talk to and I cope by playing computer games with friends and starting finding ways to help me grow and learn new things.” • “Mixed bag, while I do love that there is something to do besides play video games and watch movies all day some people are making this type of learning a huge burden.” • “All I do is facebook 24/7, sometimes watch educational videos, and even documentaries, netflix tv series and movies so that I will not feel isolated and lonely.” 	Entertainment Strategies
2	family, help, friends, support, relationship, work, busy, fun, home, parents	<ul style="list-style-type: none"> • “I was able to do things I haven’t done before for my artistic side as I have more time at home and I am lucky to have my supportive family and friends despite of many lockdowns.” • “I make myself busy with school and startup work with friends or sometimes with just myself to at least earn extra money while trying to escape from this reality. • “Our whole family got COVID last year and even though it was hard academically and financially, we got through it together that’s why I’m still thankful.” 	Relationship Strategies
3	exercise, busy, mental, videos, sleep, work, safe, health, physical, active	<ul style="list-style-type: none"> • “I feel like trap in a cage and being isolated from the outside world and somehow drains my energy, so I try to be active, do exercise at home, make myself pretty busy and forget the covid crisis. This is how I escape every day.” • “As my coping strategy, I try to enjoy and appreciate the value of sleep because I was not able to sleep consistently 8 hours a day before when the school is still face-to-face. • “Just like what I always tell my friends, do not settle on doing nothing or being less active than usual just because you’re always staying at home. Do something, exercise, walk, jog, make your mental, not only physical, health.” 	Health Strategies

In the context of coping strategies (see Table 3), the LDA analysis revealed three inferred topics: entertainment strategies, relationship strategies, and health strategies. Entertainment strategies include playing video games and watching movies; relationship strategies include connecting to friends via online platforms and getting more involved with the family; and health strategies include physical exercise and getting quality sleep. According to recent studies, these

coping mechanisms are valid and effective strategies. For instance, Barr and Copeland-Stewart (2021) reported that playing video games during the pandemic has been associated with a positive impact on the players' perceived well-being. This effect is attributed to video games providing an enjoyable means of stress relief, maintaining social contact, and offering a mentally stimulating escape from the impact of lockdown. Meanwhile, Boursier et al. (2021) observed that watching television series can act as a recuperation tactic for managing emotional turmoil by enabling individuals to seek momentary refuge in imaginary worlds. Consequently, students were able to devise their coping strategies to endure the pandemic, independent of any educational establishment's influence. Nevertheless, despite the effectiveness of these coping mechanisms (Garcia & Revano, 2022), it is recommended that educational institutions design interventions and academic support initiatives that can more effectively assist students, particularly with respect to the psychoeducational ramifications of the pandemic. Such interventions can facilitate students' adaptation to the "new normal" and help them build resilience and coping skills that can serve them beyond the pandemic (Almeida, 2023; Lamsal, 2022; Tomé & Coelho, 2023).

RQ4: Using lexicon-based sentiment analysis, what is the general polarity of free-form responses generated in and extracted from an ERL environment?

In contrast to earlier research that reported positive sentiments (Lau & Sim, 2020) and perceptions (Khan et al., 2021), the results of the sentiment analysis in the present study revealed a predominantly negative experience among students concerning the ERL setup. The findings suggest that students are not satisfied with their experience of ERL despite the efforts made by educators and institutions to enhance their online learning environment (e.g., Ofofu-Ampong et al., 2024). The negative sentiment may be attributed to various factors, such as the lack of face-to-face interaction, limited socialization opportunities, and technical difficulties encountered during the learning process (Guillaume et al., 2022; Leahy et al., 2021). In addition, a chi-square test revealed that there is an unequal number of positive and negative sentiments in the text corpus ($\chi^2(1) = 37.71$), which is unlikely to have occurred from a sampling error ($p < .05$). This finding suggests that the prevalence of negative sentiments in the text corpus is not merely a chance occurrence, but rather reflects a genuine trend in the students' experiences of ERL. Moreover, the significance level of $p < .05$ suggests that this trend is statistically meaningful and not likely to be a spurious result. Table 4 shows some of the most frequently occurring words and sample sentences from the text corpus with positive and negative sentiments.

Table 4. Sample sentences that show positive and negative sentiments

Word	Positive	Negative
family	“The good thing about pandemic for me is I get to spend more time with my family and do fun stuff together.”	“It is depressingly sad and feel lonely even though I have my family with me more frequently than before.”

exercise	“I really hate the lockdown part of the pandemic since all gyms are closed and I cannot concentrate to exercise at home even before”	“What I hate about this pandemic is the lockdown because I cannot go outside to exercise , socialize, and go to school.”
sleep	“I sleep longer now because our online class is flexible and many of our activities are asynchronous which is a good thing for me.”	“My life became unstable because of this pandemic, I cannot sleep anymore properly because of many assignments.”
help	“I have more time now to play computer games with friends and even find ways to help me grow and learn new things.”	“Even as a privileged Filipino, I cannot help but worry about me and my family’s well-being and the fear of having the virus.”
facebook	“Good thing facebook was invented before pandemic because even though there is a lockdown, I can still contact my friends and talk to my family from other places.”	“I am demotivated to study these days and I see myself scrolling the news feed of my facebook account for hours every day without doing any assignments.”
safe	“Online class is the only solution I believe to continue education while staying safe as well in the comfort of our homes.”	“I fear going out because it is not safe just to even breathe the same air with other people so I isolate myself in my room”

DISCUSSION

The advent of the Internet, social media networks, and other digital communication channels has resulted in the proliferation of substantial volumes of textual data. Despite the established effectiveness of traditional textual analysis techniques, it has become increasingly impractical to analyze such large datasets due to the necessitated time and resources (Galdas, 2017; Klemenčič & Chirikov, 2015). As a viable alternative to traditional methods, text mining is increasingly being acknowledged as an effective methodology for analyzing large volumes of textual data (Macanovic, 2022; Wiedemann, 2016). Like other textually rich websites, online learning platforms also contain significant amounts of textual content. The sudden transition of instructional delivery to ERL during the COVID-19 pandemic has intensified this trend of increased textual data production. Analyzing these data is an opportunity to prove the viability of text mining techniques in the education discipline. The present study thus examined data extracted from an ERL environment using text mining techniques, such as word frequency, collocation, KWIC concordance, topic modeling, and sentiment analysis. Assessing online learning environments provides a valuable opportunity to explore text mining as an educational evaluation methodology because these platforms are repositories of extensive and varied textual data.

According to the findings, the most frequent word in the text corpus was “hard”, with a total of 689 instances constituting 70.45% of all occurrences. Interestingly, the word "life" was observed more frequently than "school," which implies that individuals experience personal difficulties more acutely than academic challenges. The analysis of text data indicated that mental health is a significant issue among students in ERL, as evidenced by the prevalence of this topic

in top bigrams. Other researchers such as Lai et al. (2020), Yang et al. (2021), and Lee et al. (2021) have similarly raised pandemic-induced mental health concerns, such as stress, depression, and anxiety. More importantly, Tee et al. (2020) reported that students experience more severe psychological impacts and exhibit greater symptoms of mental health problems compared to those who are employed. Numerous factors have been proposed to account for poor mental health among students, including difficulties with maintaining focus during online classes, a rigorous academic curriculum, the demand for academic success, the physical distance from educational institutions, and financial constraints related to educational costs. Given the complex and varied nature of these challenges, educational institutions should undertake measures to address the pressing mental health crisis that is affecting students.

When it comes to academic challenges, the text mining analysis unveiled the heavy academic workload as a major issue. This conclusion was substantiated by the frequency of the word "work" in 46.73% of the text corpus, making it the fourth most frequent unigram. Conversely, Khalil et al. (2020) and Barrot et al. (2021) identified the greatest academic-related challenges for students to overcome as content understanding and learning environment, respectively. A plausible justification for the prevalence of excessive workload during the pandemic could be the insufficient pedagogical skills and training of educators for conducting purely online instruction (Jain et al., 2020). Despite their digital literacy skills, teachers may lack the necessary competence to effectively teach online courses. As such, teachers tend to provide more supplementary materials to support learning goals at home and, in some cases, additional assessments (e.g., assignments, reports, case studies) to measure the attainment of these goals. Although teachers express a willingness to minimize the workload of students (Lepp et al., 2021), online learning necessitates increased effort, energy, and responsibility from students. For example, the intensive use of synchronous videoconferencing platforms has led to the emergence of "Zoom fatigue" – a phenomenon that is more draining than face-to-face meetings due to the heightened requirement for sustained attention (Spataro, 2020). The management implication of this result suggests that educational institutions must strategize on how to alleviate the burden on students and ensure a conducive academic environment in times of crisis. It is also imperative that students are provided with a manageable workload and are instructed by teachers who are more understanding towards their students' current educational and mental conditions.

Consistent with recent discoveries (e.g., Cullinan et al., 2021), the present study similarly exhibited technological barriers impeding the effective implementation of ERL despite the digital revolution of the past few decades. However, contrary to the results of a systematic review conducted by Rasheed et al. (2020) that identified technology use as the primary challenge, the findings of this study highlight inadequate internet access as the primary impediment. This finding is supported by the KWIC analysis showing that the word occurrence of "internet" in 66.86% (second most frequent word) of the text corpus also mentioned the word "slow" in its 89% occurrences ($N = 582$). The issue of inadequate access to quality internet service that is observed

in this study is analogous to that experienced by other developing nations (Khan et al., 2021; Mpungose, 2020; Zalat et al., 2021). This finding suggests that the first-level digital divide (e.g., access to the Internet) is a more pressing problem than the second-level digital divide (e.g., technology use). For developed countries, there is evidence demonstrating comparable difficulties, particularly for students from low socioeconomic status backgrounds and rural regions (Peñarrubia-Lozano et al., 2021). Given that a student's intention to use learning management systems is directly impacted by their experience with internet connectivity, the importance of establishing a dependable network infrastructure when relying exclusively on ERL as an instructional delivery platform is consequently emphasized. In the absence of such reliable infrastructure, alternative modalities such as radio and television must be considered.

Overall, the application of text mining techniques has profoundly enhanced our understanding of the impact of the sudden shift of instructional delivery from traditional to online modes of teaching. This transition placed many students at an extreme disadvantage, a conclusion reflected in the language they used in their free-form responses. Through the employment of various text mining methods, significant linguistic patterns were uncovered that highlighted the challenges faced by students in this new educational environment. Specifically, the most frequent unigrams and bigrams identified by word frequency and collocation analyses were strongly indicative of these challenges. The examination of contextual snippets of the most frequently used words using a KWIC concordance analysis presented a more nuanced understanding of language use. This analysis also facilitated the identification of patterns and relationships between words that may not be discernible through word frequency and collocation analyses. It also demonstrated the existence of certain words that were used in different contexts. The process of topic modeling helped identify the underlying thematic structure of the corpus and provided insight into the key topics that emerged from the data. From a macro perspective, sentiment analysis uncovered the overall polarity of the corpus. Collectively, the combination of these text mining techniques allowed for a comprehensive exploration of the linguistic features of the corpus and provided a multifaceted understanding of the lived experience and coping strategies of students. Consequently, this study endorses the use of text mining as an effective methodology for analyzing large volumes of textual data and understanding complex educational phenomena.

While the COVID-19 pandemic and the transition to ERL provided a context for this research, the focus of the study is not solely on these events. Instead, the primary objective is to demonstrate the potential of text mining as a method for analyzing large volumes of textual data—a task that traditional qualitative techniques may find challenging. As repeatedly pointed out in many research articles (Galdas, 2017; Klemenčič & Chirikov, 2015), a traditional qualitative analysis is often manual and time-consuming and requires iterative processes that can be impractical for handling large datasets. Conversely, text mining automates much of the data processing, enabling researchers to efficiently analyze vast quantities of text to identify trends, patterns, and sentiments that would be difficult to discern manually. This study contributes to the

growing body of evidence suggesting that text mining approaches can support and, in some aspects, enhance the capabilities of a traditional qualitative data analysis (Janasik et al., 2008). By proving that text mining can address complex social science research questions effectively, this study reinforces the method's validity and applicability in academic and practical settings.

Despite the successful analyses using text mining techniques, there are limitations to the study that should be acknowledged. Primarily, while text mining provided significant insights into the language used by students, it did not explore the reasons behind the use of specific language. This limitation restricts our understanding of the surface level of textual expressions and omits deeper explorations of the underlying motivations, attitudes, and beliefs that shape these expressions. From the educational research perspective, relying solely on written text as the source of data can overlook critical aspects of the learning experience, such as verbal communication, nonverbal cues, and social interactions, which are often vital in fully understanding student engagement and learning dynamics. Future research could benefit from exploring the use of text mining techniques in combination with other data sources (e.g., video or audio recordings). Finally, the study was limited to text mining techniques because it was the scope of the study. Future researchers are encouraged to consider using text mining techniques in combination with other textual analysis methods. Comparing the strengths and limitations of different approaches may identify the most appropriate methods of data analysis.

CONCLUSION

This study aimed to test the efficacy of text mining as an educational evaluation methodology by analyzing student experience within online learning environments. It examined free-form comments from a series of cross-sectional surveys to understand how effectively text mining techniques can uncover insights from large volumes of textual data. Through the application of these techniques, the analyses revealed "hard" and "mental health" as the most frequent unigram and bigram, respectively. KWIC analysis was utilized to uncover the contexts and meanings behind these prevalent words. Additionally, topic modeling identified themes such as individual, academic, and technological challenges related to the learning experiences, alongside themes of entertainment, relationships, and health in coping strategies. The analysis also detected a predominately negative sentiment toward the experiences within these online environments. Unlike the limited insights typically derived from quantitative and traditional qualitative methods, the findings of this study underscore the depth and uniqueness of information that can be extracted through computational text analysis. Furthermore, this research provides empirical evidence of the effectiveness of text mining techniques in leveraging computational methods to achieve a comprehensive understanding of complex phenomena within large-scale datasets. This approach not only enhances our comprehension of student experiences but also enriches the methodologies available for educational research.

REFERENCES

- Albalawi, R., Yeap, T. H., & Benyoucef, M. (2020). Using Topic Modeling Methods for Short-Text Data: A Comparative Analysis. *Frontiers in Artificial Intelligence*, 3(42), 1-14. <https://doi.org/10.3389/frai.2020.00042>
- Alcober, G. I., Revano Jr, T. F., & Garcia, M. B. (2020). E-Safety in the Use of Social Networking Application. *International Journal of Advanced Trends in Computer Science and Engineering*, 9(1), 71-75. <https://doi.org/10.30534/ijatcse/2020/1291.22020>
- Aldridge, J. M., & Bianchet, S. (2022). Using Student Feedback About the Learning Environment as a Starting Point for Co-Construction. *Learning Environments Research*, 25(3), 939-955. <https://doi.org/10.1007/s10984-021-09403-9>
- Alhija, F. N.-A., & Fresko, B. (2009). Student Evaluation of Instruction: What Can Be Learned From Students' Written Comments? *Studies in Educational Evaluation*, 35(1), 37-44. <https://doi.org/10.1016/j.stueduc.2009.01.002>
- Almeida, R. S. d. (2023). Redefining Health Education in the Post-Pandemic World: How to Integrate Digital Technologies into the Curricula? In *Instructional Technologies in Health Education and Allied Disciplines*. IGI Global. <https://doi.org/10.4018/978-1-6684-7164-7.ch001>
- Alsayat, A., & Ahmadi, H. (2022). A Hybrid Method Using Ensembles of Neural Network and Text Mining for Learner Satisfaction Analysis from Big Datasets in Online Learning Platform. *Neural Processing Letters*. <https://doi.org/10.1007/s11063-022-11009-y>
- Ari, F., Arslan-Ari, I., Abaci, S., & Inan, F. A. (2022). Online Simulation for Information Technology Skills Training in Higher Education. *Journal of Computing in Higher Education*, 34(2), 371-395. <https://doi.org/10.1007/s12528-021-09303-0>
- Aucejo, E. M., French, J., Ugalde Araya, M. P., & Zafar, B. (2020). The Impact of COVID-19 on Student Experiences and Expectations: Evidence from a Survey. *Journal of Public Economics*, 191, 1-15. <https://doi.org/10.1016/j.jpubeco.2020.104271>
- Aydin, C. (2021). Examining Public Opinion Regarding Online Learning during COVID-19 Outbreak: Sentiment Analysis. *European Journal of Science and Technology*(29), 425-431. <https://doi.org/10.31590/ejosat.1035267>
- Barnbrook, G., Mason, O., & Krishnamurthy, R. (2013). The Concept of Collocation. In G. Barnbrook, O. Mason, & R. Krishnamurthy (Eds.), *Collocation: Applications and Implications* (pp. 3-31). Palgrave Macmillan UK. https://doi.org/10.1057/9781137297242_1
- Barr, M., & Copeland-Stewart, A. (2021). Playing Video Games During the COVID-19 Pandemic and Effects on Players' Well-Being. *Games and Culture*, 17(1), 122-139. <https://doi.org/10.1177/15554120211017036>
- Barrot, J. S., Llenares, I. I., & del Rosario, L. S. (2021). Students' Online Learning Challenges during the Pandemic and How they Cope with Whem: The Case of the Philippines. *Education and Information Technologies*, 26(7321-7338). <https://doi.org/10.1007/s10639-021-10589-x>
- Beltz, A. M., Wright, A. G. C., Sprague, B. N., & Molenaar, P. C. M. (2016). Bridging the Nomothetic and Idiographic Approaches to the Analysis of Clinical Data. *Assessment*, 23(4), 447-458. <https://doi.org/10.1177/1073191116648209>
- Blackmon, S. J., & Major, C. (2012). Student Experiences in Online Courses: A Qualitative Research Synthesis. *Quarterly Review of Distance Education*, 13(2), 77-85.
- Bond, M. (2020). Schools and Emergency Remote Education During the COVID-19 Pandemic: A Living Rapid Systematic Review. *Asian Journal of Distance Education*, 15(2), 191-247.
- Boursier, V., Musetti, A., Gioia, F., Flayelle, M., Billieux, J., & Schimmenti, A. (2021). Is Watching TV Series an Adaptive Coping Strategy During the COVID-19 Pandemic? Insights From an Italian Community Sample. *Frontiers in Psychiatry*, 12(554), 1-9. <https://doi.org/10.3389/fpsy.2021.599859>
- Bozkurt, M. (2021). Online Learning Communities in Covid-19 Days: Mining Twitter Data. *International Technology and Education Journal*, 5(2), 67-74. <https://files.eric.ed.gov/fulltext/EJ1338068.pdf>
- Bringula, R., Ulfa, S., Miranda, J. P. P., & Atienza, F. A. L. (2022). Text mining analysis on students' expectations and anxieties towards data analytics course. *Cogent Engineering*, 9(1), 1-17. <https://doi.org/10.1080/23311916.2022.2127469>

- Bringula, R. P., Catacutan-Bangit, A. E., Garcia, M. B., Gonzales, J. P. S., & Valderama, A. M. C. (2022). "Who is gullible to political disinformation?": Predicting Susceptibility of University Students to Fake News. *Journal of Information Technology & Politics*, 19(2), 165-179. <https://doi.org/10.1080/19331681.2021.1945988>
- Brockx, B., Van Roy, K., & Mortelmans, D. (2012). The Student as a Commentator: Students' Comments in Student Evaluations of Teaching. *Procedia - Social and Behavioral Sciences*, 69, 1122-1133. <https://doi.org/10.1016/j.sbspro.2012.12.042>
- Cayubit, R. F. O. (2022). Why Learning Environment Matters? An Analysis on How the Learning Environment Influences the Academic Motivation, Learning Strategies and Engagement of College Students. *Learning Environments Research*, 25(2), 581-599. <https://doi.org/10.1007/s10984-021-09382-x>
- Cordeiro, C. M. (2019). A corpus-based approach to understanding market access in fisheries and aquaculture international business research: A systematic literature review. *Aquaculture and Fisheries*, 4(6), 219-230. <https://doi.org/10.1016/j.aaf.2019.06.001>
- Creswell, J. D. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage.
- Cullinan, J., Flannery, D., Harold, J., Lyons, S., & Palcic, D. (2021). The Disconnected: COVID-19 and Disparities in access to Quality Broadband for Higher Education Students. *International Journal of Educational Technology in Higher Education*, 18(26), 1-21. <https://doi.org/10.1186/s41239-021-00262-1>
- Du, C., Zhan, M. C. H., Cho, M. J., Fenton, J. I., Hsiao, P. Y., Hsiao, R., Keaver, L., Lai, C.-C., Lee, H., Ludy, M.-J., Shen, W., Swee, W. C. S., Thirvikraman, J., Tseng, K.-W., Tseng, W.-C., Doak, S., Folk, S. Y. L., & Tucker, R. M. (2021). The Effects of Sleep Quality and Resilience on Perceived Stress, Dietary Behaviors, and Alcohol Misuse: A Mediation-Moderation Analysis of Higher Education Students from Asia, Europe, and North America during the COVID-19 Pandemic. *Nutrients*, 13(2), 442-462. <https://doi.org/10.3390/nu13020442>
- Engzell, P., Frey, A., & Verhagen, M. D. (2021). Learning loss due to school closures during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences*, 118(17), 1-7. <https://doi.org/10.1073/pnas.2022376118>
- Fung, C. Y., Su, S. I., Perry, E. J., & Garcia, M. B. (2022). Development of a Socioeconomic Inclusive Assessment Framework for Online Learning in Higher Education. In *Socioeconomic Inclusion During an Era of Online Education* (pp. 23-46). IGI Global. <https://doi.org/10.4018/978-1-6684-4364-4.ch002>
- Gakhal, S., & Wilson, C. (2019). Is Students' Qualitative Feedback Changing, Now it is Online? *Assessment & Evaluation in Higher Education*, 44(3), 476-488. <https://doi.org/10.1080/02602938.2018.1521371>
- Galdas, P. (2017). Revisiting Bias in Qualitative Research: Reflections on Its Relationship With Funding and Impact. *International Journal of Qualitative Methods*, 16(1), 1609406917748992. <https://doi.org/10.1177/1609406917748992>
- Garcia, M. B. (2020). Sentiment Analysis of Tweets on Coronavirus Disease 2019 (COVID-19) Pandemic from Metro Manila, Philippines. *Cybernetics and Information Technologies*, 20(4), 141-155. <https://doi.org/10.2478/cait-2020-0052>
- Garcia, M. B. (2022). *Socioeconomic Inclusion During an Era of Online Education*. IGI Global. <https://doi.org/10.4018/978-1-6684-4364-4>
- Garcia, M. B., & Cunanan-Yabut, A. (2022). Public Sentiment and Emotion Analyses of Twitter Data on the 2022 Russian Invasion of Ukraine. *2022 9th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, 242-247. <https://doi.org/10.1109/ICITACEE55701.2022.9924136>
- Garcia, M. B., & Revano, T. F. (2022). Pandemic, Higher Education, and a Developing Country: How Teachers and Students Adapt to Emergency Remote Education. *2022 4th Asia Pacific Information Technology Conference*, 111-115. <https://doi.org/10.1145/3512353.3512369>
- Garcia, M. B., & Yousef, A. M. F. (2023). Cognitive and Affective Effects of Teachers' Annotations and Talking Heads on Asynchronous Video Lectures in a Web Development Course. *Research and Practice in Technology Enhanced Learning*, 18(20), 1-23. <https://doi.org/10.58459/rptel.2023.18020>
- Guillaume, M., Toomarian, E. Y., Van Rinsveld, A., Baskin-Sommers, A., Dick, A. S., Dowling, G. J., Gonzalez, M. R., Hasak, L., Lisdahl, K. M., Marshall, A. T., Nguyen, Q. T. H., Pelham, W. E., Pillai, C. C., Sheth, C., Wang, A. M., Tapert, S. F., & McCandliss, B. D. (2022). Negative Impacts of Pandemic Induced At-Home Remote Learning Can Be Mitigated by Parental Involvement. *Frontiers in Education*, 7, 1-12. <https://doi.org/10.3389/feduc.2022.804191>

- Hadad, S., Shamir-Inbal, T., & Blau, I. (2024). Pedagogical Strategies Employed in the Emergency Remote Learning Environment During the COVID-19 Pandemic: the Tale of Teachers and School ICT Coordinators. *Learning Environments Research*. <https://doi.org/10.1007/s10984-023-09487-5>
- Hoon, A., Oliver, E., Szpakowska, K., & Newton, P. (2015). Use of the 'Stop, Start, Continue' Method is Associated with the Production of Constructive Qualitative Feedback by Students in Higher Education. *Assessment & Evaluation in Higher Education*, 40(5), 755-767. <https://doi.org/10.1080/02602938.2014.956282>
- Ignatow, G., & Mihalcea, R. (2017). *An Introduction to Text Mining: Research Design, Data Collection, and Analysis*. SAGE Publications. <https://doi.org/10.4135/9781506336985>
- Inaba, M., & Kakai, H. (2019). Grounded Text Mining Approach: A Synergy between Grounded Theory and Text Mining Approaches. In A. Bryant & K. Charmaz (Eds.), *The SAGE Handbook of Current Developments in Grounded Theory*. SAGE Publications. <https://doi.org/10.4135/9781526485656>
- Islam, M., Mazlan, N. H., Al Murshidi, G., Hoque, M. S., Karthiga, S. V., & Reza, M. (2023). UAE University Students' Experiences of Virtual Classroom Learning During COVID 19. *Smart Learning Environments*, 10(1), 1-16. <https://doi.org/10.1186/s40561-023-00225-1>
- Jain, S., Lall, M., & Singh, A. (2020). Teachers' Voices on the Impact of COVID-19 on School Education: Are Ed-Tech Companies Really the Panacea? *Contemporary Education Dialogue*, 18(1), 58-89. <https://doi.org/10.1177/0973184920976433>
- Janasik, N., Honkela, T., & Bruun, H. (2008). Text Mining in Qualitative Research: Application of an Unsupervised Learning Method. *Organizational Research Methods*, 12(3), 436-460. <https://doi.org/10.1177/1094428108317202>
- Khalil, R., Mansour, A. E., Fadda, W. A., Almisnid, K., Aldamegh, M., Al-Nafeesah, A., Alkhalifah, A., & Al-Wutayd, O. (2020). The Sudden Transition to Synchronized Online Learning during the COVID-19 Pandemic in Saudi Arabia: A Qualitative Study Exploring Medical Students' Perspectives. *BMC Medical Education*, 20(1), 1-10. <https://doi.org/10.1186/s12909-020-02208-z>
- Khan, M. A., Kamal, T., Illiyan, A., & Asif, M. (2021). School Students' Perception and Challenges towards Online Classes during COVID-19 Pandemic in India: An Econometric Analysis. *Sustainability*, 13(9), 4786-4800. <https://doi.org/10.3390/su13094786>
- Klemenčič, M., & Chirikov, I. (2015). How Do We Know How Students Experience Higher Education? On the Use of Student Surveys. In A. Curaj, L. Matei, R. Pricopie, J. Salmi, & P. Scott (Eds.), *The European Higher Education Area: Between Critical Reflections and Future Policies* (pp. 361-379). Springer International Publishing. https://doi.org/10.1007/978-3-319-20877-0_24
- Kuckartz, U. (2014). *Qualitative Text Analysis: A Guide to Methods, Practice & Using Software*. SAGE Publications. <https://doi.org/10.4135/9781446288719>
- Kwon, H.-J., Ban, H.-J., Jun, J.-K., & Kim, H.-S. (2021). Topic Modeling and Sentiment Analysis of Online Review for Airlines. *Information*, 12(2), 78-91. <https://doi.org/10.3390/info12020078>
- Lai, A. Y.-k., Lee, L., Wang, M.-p., Feng, Y., Lai, T. T.-k., Ho, L.-m., Lam, V. S.-f., Ip, M. S.-m., & Lam, T.-h. (2020). Mental Health Impacts of the COVID-19 Pandemic on International University Students, Related Stressors, and Coping Strategies. *Frontiers in Psychiatry*, 11(1082), 1-13. <https://doi.org/10.3389/fpsy.2020.584240>
- Lamsal, B. (2022). Exploring Issues Surrounding a Safe and Conducive Digital Learning Space in Nepal: A Preparation for Online Education in the Post-Pandemic Era. In *Socioeconomic Inclusion During an Era of Online Education* (pp. 246-263). IGI Global. <https://doi.org/10.4018/978-1-6684-4364-4.ch012>
- Lau, S. L., & Sim, T. Y. (2020). Feedback of University Students on Online Delivery Learning During the COVID-19 Pandemic Period. *2020 IEEE Conference on e-Learning, e-Management and e-Services (IC3e)*, 13-18. <https://doi.org/10.1109/IC3e50159.2020.9288409>
- Leahy, F., Newton, P., & Khan, A. (2021). *Learning During the Pandemic: Quantifying Lost Learning* (Education, Universities and Childcare, Issue).
- Lee, J., Jeong, H. J., & Kim, S. (2021). Stress, Anxiety, and Depression Among Undergraduate Students during the COVID-19 Pandemic and their Use of Mental Health Services. *Innovative Higher Education*, 46, 519-538. <https://doi.org/10.1007/s10755-021-09552-y>
- Lepp, L., Aaviku, T., Leijen, Ä., Pedaste, M., & Saks, K. (2021). Teaching during COVID-19: The Decisions Made in Teaching. *Education Sciences*, 11(2), 47-67. <https://doi.org/10.3390/educsci11020047>

- Lijffijt, J., Papapetrou, P., Puolamäki, K., & Mannila, H. (2011). Analyzing Word Frequencies in Large Text Corpora Using Inter-arrival Times and Bootstrapping. *Machine Learning and Knowledge Discovery in Databases*, 341-357. https://doi.org/10.1007/978-3-642-23783-6_22
- Lin, E. C. P., & Yeh, A. J. (2022). Fighting Through COVID-19 for Educational Continuity: Challenges to Teachers. In *Socioeconomic Inclusion During an Era of Online Education* (pp. 177-203). IGI Global. <https://doi.org/10.4018/978-1-6684-4364-4.ch009>
- Liu, Z., Li, M., Liu, Y., & Ponraj, M. (2011). Performance Evaluation of Latent Dirichlet Allocation in Text Mining. *2011 Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, 4, 2695-2698. <https://doi.org/10.1109/FSKD.2011.6020066>
- Macanovic, A. (2022). Text Mining for Social Science – The State and the Future of Computational Text Analysis in Sociology. *Social Science Research*, 108, 1-17. <https://doi.org/10.1016/j.ssresearch.2022.102784>
- Maxwell, J. A. (2012). The Importance of Qualitative Research for Causal Explanation in Education. *Qualitative Inquiry*, 18(8), 655-661. <https://doi.org/10.1177/1077800412452856>
- Mirahmadizadeh, A., Ranjbar, K., Shahriarirad, R., Erfani, A., Ghaem, H., Jafari, K., & Rahimi, T. (2020). Evaluation of Students' Attitude and Emotions Towards the Sudden Closure of Schools During the COVID-19 Pandemic: A Cross-Sectional Study. *BMC Psychology*, 8(1), 1-7. <https://doi.org/10.1186/s40359-020-00500-7>
- Miranda, J. P. P., & Bringula, R. P. (2021). Exploring Philippine Presidents' Speeches: A Sentiment Analysis and Topic Modeling Approach. *Cogent Social Sciences*, 7(1), 1-13. <https://doi.org/10.1080/23311886.2021.1932030>
- Mpungose, C. B. (2020). Emergent Transition from Face-to-Face to Online Learning in a South African University in the Context of the Coronavirus Pandemic. *Humanities and Social Sciences Communications*, 7(1), 113-121. <https://doi.org/10.1057/s41599-020-00603-x>
- Muller, M., Guha, S., Baumer, E. P. S., Mimno, D., & Shami, N. S. (2016). Machine Learning and Grounded Theory Method: Convergence, Divergence, and Combination. *Proceedings of the 2016 ACM International Conference on Supporting Group Work*, 3-8. <https://doi.org/10.1145/2957276.2957280>
- Nguyen, D., Liakata, M., DeDeo, S., Eisenstein, J., Mimno, D., Tromble, R., & Winters, J. (2020). How We Do Things With Words: Analyzing Text as Social and Cultural Data. *Frontiers in Artificial Intelligence*, 3, 1-14. <https://doi.org/10.3389/frai.2020.00062>
- O'Donnell, M. B. (2008). KWICrouper: designing a tool for corpus-driven concordance analysis. *International Journal of English Studies*, 8(1).
- Oforu-Ampong, K., Agyekum, M. W., & Garcia, M. B. (2024). Long-Term Pandemic Management and the Need to Invest in Digital Transformation: A Resilience Theory Perspective. In *Transformative Approaches to Patient Literacy and Healthcare Innovation* (pp. 242-260). IGI Global. <https://doi.org/10.4018/979-8-3693-3661-8.ch012>
- Ong, S. G. T., & Quek, G. C. L. (2023). Enhancing Teacher–Student Interactions and Student Online Engagement in an Online Learning Environment. *Learning Environments Research*, 26, 681-707. <https://doi.org/10.1007/s10984-022-09447-5>
- Peñarrubia-Lozano, C., Segura-Berges, M., Lizalde-Gil, M., & Bustamante, J. C. (2021). A Qualitative Analysis of Implementing E-Learning during the COVID-19 Lockdown. *Sustainability*, 13(6), 3317-3344. <https://doi.org/10.3390/su13063317>
- Rahman, M. S. (2017). The Advantages and Disadvantages of Using Qualitative and Quantitative Approaches and Methods in Language “Testing and Assessment” Research: A Literature Review. *Journal of Education and Learning*, 6(1), 102-112. <https://doi.org/10.5539/jel.v6n1p102>
- Rasheed, R. A., Kamsin, A., & Abdullah, N. A. (2020). Challenges in the Online Component of Blended Learning: A Systematic Review. *Computers & Education*, 144, 1-17. <https://doi.org/10.1016/j.compedu.2019.103701>
- Rusticus, S. A., Pashootan, T., & Mah, A. (2023). What Are the Key Elements of a Positive Learning Environment? Perspectives From Students and Faculty. *Learning Environments Research*, 26(1), 161-175. <https://doi.org/10.1007/s10984-022-09410-4>
- Salta, K., Paschalidou, K., Tsetseri, M., & Koulougliotis, D. (2022). Shift From a Traditional to a Distance Learning Environment during the COVID-19 Pandemic. *Science & Education*, 31(1), 93-122. <https://doi.org/10.1007/s11191-021-00234-x>
- Spataro, J. (2020). *The future of work—the good, the challenging & the unknown*. Microsoft. <https://www.microsoft.com/en-us/microsoft-365/blog/2020/07/08/future-work-good-challenging-unknown/>

- Syyrilä, T., Vehviläinen-Julkunen, K., & Härkänen, M. (2021). Healthcare professionals' perceptions on medication communication challenges and solutions - text mining and manual content analysis - cross-sectional study. *BMC Health Serv Res*, 21(1), 1226. <https://doi.org/10.1186/s12913-021-07227-0>
- Tee, M. L., Tee, C. A., Anlacan, J. P., Aligam, K. J. G., Reyes, P. W. C., Kuruchittham, V., & Ho, R. C. (2020). Psychological Impact of COVID-19 Pandemic in the Philippines. *Journal of affective disorders*, 277, 379-391. <https://doi.org/10.1016/j.jad.2020.08.043>
- Tomé, A., & Coelho, J. L. (2023). Physiotherapy Education in the Digital Era: A Roadmap of Educational Technologies for Allied Health Educators. In *Instructional Technologies in Health Education and Allied Disciplines*. IGI Global. <https://doi.org/10.4018/978-1-6684-7164-7.ch002>
- Treceña, J. K. D. (2022). COVID-19 and Remote Learning in the Philippine Basic Education System: Experiences of Teachers, Parents, and Students. In *Socioeconomic Inclusion During an Era of Online Education* (pp. 92-110). IGI Global. <https://doi.org/10.4018/978-1-6684-4364-4.ch005>
- Tremblay, S., Castiglione, S., Audet, L.-A., Desmarais, M., Horace, M., & Peláez, S. (2021). Conducting Qualitative Research to Respond to COVID-19 Challenges: Reflections for the Present and Beyond. *International Journal of Qualitative Methods*, 20, 1-8. <https://doi.org/10.1177/16094069211009679>
- Valderama, A. M., Tuazon, J. B., & Garcia, M. B. (2022). Promoting Student Engagement by Questioning the Faculty and Playing a Gamified Learning Activity. *2022 IEEE 14th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*. <https://doi.org/10.1109/HNICEM57413.2022.10109470>
- Vindrola-Padros, C., Chisnall, G., Cooper, S., Dowrick, A., Djellouli, N., Symmons, S. M., Martin, S., Singleton, G., Vanderslott, S., Vera, N., & Johnson, G. A. (2020). Carrying Out Rapid Qualitative Research During a Pandemic: Emerging Lessons From COVID-19. *Qualitative Health Research*, 30(14), 2192-2204. <https://doi.org/10.1177/1049732320951526>
- Wahyono, I. D., Asfani, K., Mohamad, M. M., Saryono, D., Putranto, H., & Said, M. N. H. B. M. (2021). Text Mining in Chat Room of Online Learning for Detection Emotion using Artificial Intelligence. *2021 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE)*, 63-67. <https://doi.org/10.1109/ICOMITEE53461.2021.9650099>
- Wanniarachchi, V. U., Scogings, C., Susnjak, T., & Mathrani, A. (2022). Fat stigma and body objectification: A text analysis approach using social media content. *DIGITAL HEALTH*, 8, 1-14. <https://doi.org/10.1177/20552076221117404>
- Wiedemann, G. (2016). Summary: Integrating Qualitative and Computational Text Analysis. In G. Wiedemann (Ed.), *Text Mining for Qualitative Data Analysis in the Social Sciences: A Study on Democratic Discourse in Germany* (pp. 251-260). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-15309-0_6
- Wong, A. C. K. (2015). Understanding Students' Experiences in Their Own Words: Moving beyond a Basic Analysis of Student Engagement. *Canadian Journal of Higher Education*, 45(2), 60-80. <https://doi.org/10.47678/cjhe.v45i2.184369>
- Yang, C., Chen, A., & Chen, Y. (2021). College Students' Stress and Health in the COVID-19 Pandemic: The Role of Academic Workload, Separation from School, and Fears of Contagion. *PLOS ONE*, 16(2), 1-16. <https://doi.org/10.1371/journal.pone.0246676>
- Yang, L.-X., & Sheu, C.-F. (2019). A text analysis approach to analyzing gender differences in breakup posts on social media. *Chinese Journal of Psychology*, 61, 209-230. <https://psycnet.apa.org/record/2019-72276-003>
- Yu, C. H., Jannasch-Pennell, A., & DiGangi, S. (2011). Compatibility Between Text Mining and Qualitative Research in the Perspectives of Grounded Theory, Content Analysis, and Reliability. *The Qualitative Report*, 16(3), 730-744. <https://doi.org/10.46743/2160-3715/2011.1085>
- Zalat, M. M., Hamed, M. S., & Bolbol, S. A. (2021). The Experiences, Challenges, and Acceptance of e-Learning as a Tool for Teaching during the COVID-19 Pandemic Among University Medical Staff. *PLOS ONE*, 16(3), 1-12. <https://doi.org/10.1371/journal.pone.0248758>
- Zhou, M., & Mou, H. (2022). Tracking public opinion about online education over COVID-19 in China. *Educ Technol Res Dev*, 70(3), 1083-1104. <https://doi.org/10.1007/s11423-022-10080-5>

RELATED RESEARCH

Book Chapter

Understanding Student Engagement in AI-Powered Online Learning Platforms: A Narrative Review of Key Theories and Models

Garcia, M. B., Goi, C., Shively, K., Maher, D., Rosak-Szyrocka, J., Happonen, A., Bozkurt, A., & Damaševičius, R. (2025). *Cases on Enhancing P-16 Student Engagement with Digital Technologies*. <https://doi.org/10.4018/979-8-3693-5633-3.ch001>.

Book Chapter

Development of a Socioeconomic Inclusive Assessment Framework for Online Learning in Higher Education

Fung, C. Y., Su, S. I., Perry, E. J., & Garcia, M. B. (2022). *Socioeconomic Inclusion During an Era of Online Education*, 23-46. <https://doi.org/10.4018/978-1-6684-4364-4.ch002>.

Conference Paper

Theories Integrated with Technology Acceptance Model (TAM) in Online Learning Acceptance and Continuance Intention: A Systematic Review

Mustafa, A. S. & Garcia, M. B. (2021). *2021 1st Conference on Online Teaching for Mobile Education (OT4ME)*, 1-5. <https://doi.org/10.1109/OT4ME53559.2021.9638934>.

LET'S COLLABORATE!

If you are looking for research collaborators, please do not hesitate to contact me at mbgarcia@feutech.edu.ph.



ABOUT THE CORRESPONDING AUTHOR:

Manuel B. Garcia is a professor of information technology and the founding director of the Educational Innovation and Technology Hub (EdITH) at FEU Institute of Technology, Manila, Philippines. His interdisciplinary research interest includes topics that, individually or collectively, cover the disciplines of education and information technology. He is a licensed professional teacher and a proud member of the National Research Council of the Philippines – an attached agency to the country's Department of Science and Technology (DOST-NRCP).