Oct. 2023

# Educational Technology & Society

Published by International Forum of Educational Technology & Society
Hosted by National Taiwan Normal University, Taiwan

Volume 26 Issue 4
October 2023
ISSN: 1436-4522 (online)
ISSN: 1176-3647 (print)
DOI: 10.30191/ETS
http://www.j-ets.net/

Educational Technology & Society

vol.26
no.4

Educational Technology & Society has Impact Factor 4.020 and
5-Year impact factor 4.865 according to Thomson Scientific 2022 Journal Citations Report.

http://www.j-ets.net/

# Educational Technology & Society
An International Journal

## Aims and Scope

*Educational Technology & Society (ET&S)* is an open-access academic journal published quarterly (January, April, July, and October) since October 1998. *ET&S* has achieved its purposes of providing an international forum for open access scientific dialogue for developers, educators and researchers to foster the development of research in educational technology. Thanks to all the Authors, Reviewers and Readers, the journal has enjoyed tremendous success.

*ET&S* has established a solid and stable editorial office where the Editors-in-Chief aims to promote innovative educational technology research based on empirical inquiries to echo the pedagogical essentials of learning in the real world—lifelong learning, competency-orientation, and multimodal literacy in the 21st century.

*ET&S* publishes research that well bridges the pedagogy and practice in advanced technology for evidence-based and meaningful educational application. The focus of *ET&S* is not only technology per se, but rather issues related to the process continuum of learning, teaching, and assessment and how they are affected or enhanced using technologies. The empirical research about how technology can be used to overcome the existing problems in the frontline of local education with findings that can be applied to the global spectrum is also welcome. However, papers with only descriptions of the results obtained from self-report surveys without systematic or empirical data or any analysis on learning outcomes or processes are not favorable to be included in *ET&S*.

*ET&S* publishes research that fulfills any of the following description:

- **Evidence-based Research:** Continuous research that is conducted within a sufficient amount of time such that the effectiveness of the research intervention in education can be evaluated and demonstrated convincingly via empirical methods. Preferably, research experience, outcome, and results for a semester long of adopting educational technologies; or an evaluation/experiment that is conducted multiple times and spans over several semesters.
- **Emerging Technology:** New and emerging technology used in education
- **Special Domain Research:** Research with specific participants using technology in education, for example, those with privacy issues in special education or those of younger ages like infants and toddlers, who may not form a large pool of participants





ISSN 1436-4522 (online) and 1176-3647 (print). DOI 10.30191/ETS. 

## Publisher
International Forum of Educational Technology & Society

## Host
National Taiwan Normal University, Taiwan

## Editorial Office
c/o Chair Professor Nian-Shing Chen, Institute for Research Excellence in Learning Sciences, Program of Learning Sciences, National Taiwan Normal University, No.162, Sec. 1, Heping E. Rd., Da-an Dist., Taipei City 10610, Taiwan.

## Supporting Organizations
University of North Texas, USA
University of Piraeus, Greece

## Website Maintenance
Institute for Research Excellence in Learning Sciences, National Taiwan Normal University, Taiwan


## Abstracting and Indexing
*Educational Technology & Society* is abstracted/indexed in Social Science Citation Index, Scopus, ACM Guide to Computing Literature, airiti, Australian DEST Register of Refereed Journals, Computing Reviews, Current Contents/Social & Behavioral Sciences, DBLP, DOAJ, Educational Administration Abstracts, Educational Research Abstracts, Educational Technology Abstracts, Elsevier Bibliographic Databases, ERIC Clearinghouse on Information & Technology, Inspec, ISI Alerting Services, JSTOR, PsycINFO, Social Scisearch, Technical Education & Training Abstracts, and VOCED.

## Guidelines for authors
Submissions are invited in the following categories:

- Peer reviewed publications: Full length articles (up to 8,000 words excluding References and Appendices)
- Special Issue publications

All peer review publications will be refereed in double-blind review process by at least two international reviewers with expertise in the relevant subject area.

For detailed information on how to format your submissions, please see: https://www.j-ets.net/author_guide

For Special Issue Proposal submission, please see: https://www.j-ets.net/journal_info/special-issue-proposals

## Submission procedure
All submissions must be uploaded through our online management system (http://www.j-ets.net). Do note that all manuscripts must comply with requirements stated in the Authors Guidelines.

Authors, submitting articles for a particular special issue, should send their submissions according to the channel specified in the Call for Paper of the special issue.

All submissions should be in electronic form. Authors will receive an email acknowledgement of their submission.

The preferred formats for submission are Word document, and not in any other word-processing or desktop-publishing formats. Please place figures and tables in their respective format in the anonymous manuscript along with all appendices (if any).

Please provide following details with each submission in a separate file (i.e., Title Page):

- Author(s) full name(s) including title(s),
- Name of corresponding author,
- Job title(s),
- Organisation(s),
- Full contact details of ALL authors including email address, postal address, telephone and fax numbers.

In case of difficulties, please contact journal.ets@gmail.com.

**Educational Technology & Society**

**Volume 26 Number 4 2023**

## Table of Contents

### Full Length Articles



### Editorial Note



### Special Issue Articles



ISSN 1436-4522 (online) and 1176-3647 (print). DOI 10.30191/ETS. 

Tseng, Y.-C., Chen, M.-R. A., & Lin, Y.-H. (2023). An Investigation of the Effects of EFL Students' Self-efficacy in an Asynchronous Online Course with Interactive Contents. *Educational Technology & Society*, *26*(4), 1-13. https://doi.org/10.30191/ETS.202310_26(4).0001

# An Investigation of the Effects of EFL Students' Self-efficacy in an Asynchronous Online Course with Interactive Contents

**Yu-Ching Tseng[1], Mei-Rong Alice Chen[2] and Yi-Hsuan Lin[3*]**
[1]Department of English, Tamkang University, Taiwan // [2]Department of English Language and Literature, Soochow University, Taiwan // [3]Department of English Language and Literature, Chinese Culture University, Taiwan // ychtseng@mail.tku.edu.tw // mralice@scu.edu.tw // lyx29@ulive.pccu.edu.tw
[*]Corresponding author

(Submitted July 26, 2022; Revised February 14, 2023; Accepted February 24, 2023)

**ABSTRACT:** This study investigates the role of self-efficacy in an asynchronous online English course enriched with interactive features. Self-efficacy is a strong predictor of academic achievement in conventional classrooms. However, when learning happens in an online environment, the students' learning achievement is also affected by their psychological perceptions of online learning. In this study, the relationship between self-efficacy and affective factors (i.e., learner autonomy, learner–content interaction, and perceptions toward transactional distance) was investigated. The aims of this study were to identify the influence of different levels of self-efficacy on these factors and to explore their relationships in an online EFL course. In total, 286 students were administered the questionnaires before and after the curriculum to probe their self-perception of these affective variables. When asynchronous interactive learning materials came into play, learners with different levels of self-efficacy make statistically different learning achievements. The statistically significant differences were also found between the student's self-efficacy level, their learner autonomy, and their perception toward the interactive contents. However, the difference was not significant between self-efficacy and transactional distance. The cost of asynchronous learning is an increasing transactional distance due to the lack of instructor-learner interaction. This study suggests that interactive content triggered an opposite effect by making the instructor's role invisible rather than absent. A good online course must balance the student's self-determined learning and flexibility with the course structure. Interactive learning content can keep the balance between developing learner autonomy and fostering engagement by dissolving the teacher's role into interactive course material.

**Keywords:** Self-efficacy, Asynchronous learning, Learner perceptions, Transactional distance, Interactive contents

## 1. Introduction

Taiwanese learners who take remedial English courses to graduate from university tend to have lower English proficiency and different levels of self-efficacy. However, the cause of their low English proficiency is not their failure to learn but rather the failure of their teachers and the traditional classroom to teach them. As online education is becoming a global phenomenon, the emergence of technology-mediated learning in the English as a Foreign Language (EFL) classroom provides learners with an asynchronous environment that allows them diverse opportunities to learn English. Today, asynchronous classrooms can replace or enhance language learning. Thanks to the support of the Internet and technology, asynchronous learning does not require the learner to interact directly with the instructor or other learners. It offers temporal and spatial flexibility, individualized learning pace, and repeatability of course material without supervision from the instructor. The typical design of an asynchronous curriculum includes online activities that learners can access anywhere and anytime (Lai & Morrison, 2013; Sulha et al., 2021). Interactive contents enhance interactivity through the use of technologies and network communication. The traditional learning materials are transformed into digitalized formats that give an impetus to reciprocal activities between the learner's input and the immediate feedback from the interactive materials (Domagk et al., 2010; Kaplar et al., 2022). Thus, this study focuses on asynchronous English courses that incorporate recorded interactive instructional videos, HTML 5 materials, interactive readings with vocabulary explanations, pop-up quizzes, and gamification into the content of the interactive learning materials.

Self-efficacy refers to a student's belief in their ability to achieve successful learning. When self-efficacy is combined with a more accessible, flexible, and resourceful environment, the affective factors from the learners' aspect might interact, which can impact their perceptions toward learning and change learning outcomes. Previous research has extensively studied self-efficacy and made important contributions by emphasizing its positive association with academic achievement (Caprara et al., 2011; Honicke & Broadbent, 2016; Talsma et al., 2018; Wang & Bai, 2017). Although self-efficacy is a predictor of academic achievement in classrooms featuring face-to-face interactions, many important questions remain unresolved in the research to date.

First, previous studies have successfully established a positive association between self-efficacy and learning achievement. However, few empirical studies have addressed the relationships between self-efficacy, self-perceived affective factors, and the emotional and attitudinal influence, which may impact one's self-efficacy in a technology-enabled learning setting. The way that a learner perceives their learning process and learning environment influences their confidence in gaining a successful learning experience, and the interference of technology and interactive features embedded in the course materials add further complexity to the issue. Therefore, the present study investigated the relationship between self-efficacy and the affective factors that can influence a learner's decisions regarding the management of their learning process, including learner autonomy, which involves the judgments that the student makes about their ability to take responsibility for their own learning process; learner–content interaction, which involves the students' perception regarding whether their interaction with the course content has been a successful experience; and transactional distance, which involves the perceived distance between the students and the course.

Second, the impact of self-efficacy on online learning has not yet been fully understood. An investigation of whether a student's self-efficacy is related to their self-concept in learning gains particular significance for an EFL course that is completely online without mandatory class meetings. In an asynchronous online course, students perceive a psychological distance from the instructor because they are physically separated. The transactional distance triggers two-sided effects. In one aspect, instructors strive to minimize this distance by developing interactive materials that can successfully engage the students' attention and time in learning. An increase in the learner's interaction with the course content reduces the distance between the students and the instructor, which could enhance the students' academic performance (Ekwunife-Orakwue & Teng, 2014; Shea et al., 2016; Swart & Wuensch, 2016). In another aspect, a highly structured and predetermined classroom often allows or requires less interaction between the students and the instructor, and therefore the transactional distance may increase. An increase in the transactional distance enhances the student's autonomy because they need to take more responsibility for their own learning (Moore, 1997). This study explained the two-sided effects on transactional distance by suggesting that a highly-structured asynchronous classroom abundant in interactive contents can enforce and constrain autonomous learning.

This study investigated whether an asynchronous EFL course enriched with interactive features can affect a student's perception of their online learning experiences through the lens of transactional distance theory (Moore, 1997). In particular, this study emphasized how high- and low-efficacious students perceive their learning experiences when they are provided with abundant interactive learning materials in a highly-structured online learning environment. The research efforts were guided by the following questions:

RQ1: Does providing interactive learning content to learners with different levels of self-efficacy make a significant difference in their learning achievements?
RQ2: Does providing interactive learning content to learners with different levels of self-efficacy make a significant difference in their perceptions regarding learner autonomy?
RQ3: Does providing interactive learning content to learners with different levels of self-efficacy make a significant difference in their perceptions regarding learner–content interactions?
RQ4: Does providing interactive learning content to learners with different levels of self-efficacy make a significant difference in their perceptions regarding the transactional distance?
RQ5: What is the relationship between the student's learning achievements, learner autonomy, learner–content interactions, transactional distance, and their self-efficacy?

## 2. Literature review

Social and cognitive theories have postulated the importance of self-concepts in motivating students' learning to achieve academic success (Deci & Ryan, 1987). This study investigated the interaction of self-efficacy with other learning variables (i.e., autonomy, learner–content interactions, and transactional distance) to discuss the effect of interactive learning content in an asynchronous classroom.

### 2.1. Self-efficacy and learning achievement

Self-efficacy is defined as a person's belief in their capacity to successfully accomplish a particular task (Bandura, 1986; Schunk, 1991). It exercises an influence over many different aspects of human behavior and plays a determinant role in predicting the effort that one puts into a designated task. More specifically, self-efficacy refers to an individual's perception of their ability "to exercise control over their own level of functioning and over events that affect their lives" (Bandura, 1991, p. 257).

Self-efficacy, as one of the most important components of self-concepts, has been studied extensively. It has been positively associated with subsequent academic proficiency in cross-sectional and longitudinal studies (Caprara et al., 2011; Wang & Bai, 2017), as well as meta-analyses (Honicke & Broadbent, 2016; Richardson et al., 2012; Talsma et al., 2018). In the educational setting, self-efficacy is a prediction of an individual's learning achievement (Bandura, 1977) because a strong sense of self-efficacy is key for "successful adaptation and change" (Bandura, 1997, p. 32).

In online learning, it is skeptical to assume that a person's sense of self-efficacy directly points to a successful learning experience. Previous studies have shown contradictory results. Some studies suggested that self-efficacy is correlated with academic learning achievement (Kitsantas & Chow, 2007; Lim, 2001; Yukselturk & Bulut, 2007), whereas others found no significant correlation between the two (Cho & Shen, 2013; Crippen et al., 2009). The direct relationship between self-efficacy and learning outcomes can be mediated by other factors. For example, the learner's attitude toward online instruction and familiarity with online learning devices (Cussó-Calabuig et al., 2018) can affect their learning outcomes. Meanwhile, task-oriented students can be motivated by task-based instruction (Yukselturk & Bulut, 2007). Finally, the competitive nature of gamified online learning activities may negatively influence the performance of students who prefer traditional classroom instructions (Charles et al., 2011).

### 2.2. Transactional distance, learner–content interaction, and learner autonomy

This study investigated the relationship between self-efficacy and the key factors under the framework of Transaction Distance Theory (TDT). According to Moore (1997), transaction distance (TD) is the psychological and communicative distance between the students and the instructor in the context where they are physically separate. The perceived distance influences different aspects of teaching and learning, which has been widely studied in the TD literature. For example, Benson and Samarawickrema (2009) designed course activities in distance education and applied components of TD to understand the students' learning experiences. Chen (2001a) conducted a study in an online setting to figure out the extent to which the students perceived psychological or communicational distance. Similar to the work by Chen (2001a; 2001b), other studies on TDT were carried out to measure the students' subjective perception regarding their actual performances in online learning environments (Zhang, 2003).

Teachers tend to reduce TD to facilitate the effect of the student's learning (Benton et al., 2013; Moore, 1984). Moore (1997) identified three factors that influence the extent of distance as follows: structure, dialogue (interaction), and learner autonomy. Structure refers to the level of rigidity of the structure of an online course, including facets such as the course objectives, the pedagogical model, the design of assessment, and the flexibility or adjustability of the course to accommodate individual student needs (Zhang, 2003). Dialogue consists of three forms of interaction in online learning (Moore, 1997), i.e., learner–content interaction, learner–learner interaction, and learner–instructor interaction. Learner autonomy refers to a student's ability to take charge of their own learning (Holec, 1981). It has been argued that learner autonomy is affected by self-efficacy, which predicts the efforts that the students are willing to put into the process of learning (Deci & Ryan, 1987). Further, learner autonomy is also highly related to TD. An increase in distance between the student and the instructor enhances the student's autonomy because they need to take more responsibility for their own learning (Moore, 1997).

Moore (1991) has suggested that there is an inverse relationship between structure and dialogue. A highly-structured course leaves little room for instructor–learner interaction, and so TD increases. This inverse relation is based on the studies that focused on the dialogue between instructor-learner interaction (Huang et al., 2015; Moore, 1991) or learner–learner interaction (Benson & Samarawickrema, 2009). When interactive learning content brings an innovative method of interaction for learners with course content, it becomes important to re-examine the relationship between structure and dialogue, as an online course that is full of interactive features requires more interactions between the students and the content. A rigidly structured online course controls part of the student's learning progress, and therefore learner autonomy may decrease because the student does not need to take full responsibility for the management of their learning. Structural inflexibility influences the different forms of interaction. From the learner's perspective, when the perceived distance changes, the requirement of self-responsibility reduces, which affects their learning achievement.

### 2.3. Transactional distance in distance education

Online distance education is defined by four characteristics: institutionalization, geographical separation, interaction, and formation of a learning community (Schlosser & Simonson, 2009). An online course is institution–based. Although the instructor and the students are inherently separate temporally and geographically, a learning community is formed, where interactions are established between the instructor and the students, the students and the contents, and among the students through the course design and the learning activities.

The technology used in an online classroom is another focus of studies using the framework of TDT. Park (2011) has proposed that mobile learning added a new dimension to the original TDT frameworks. Swart and Wuensch (2016) indicated that the adaptation of digital devices has helped in shortening TD. Along similar lines, some studies have found that decreased TD could increase the learner's academic performance (Ekwunife-Orakwue & Teng, 2014; Shea et al., 2016; Swart & Wuensch, 2016). Huang et al. (2016) found that a good online course facilitates the interaction between the participants and the instructor. However, according to Moore (1997), the way that TD is influenced depends on the type of media used in the classroom. Interactive media can over-structure course material and allow for fewer interactions for the students to communicate with the instructor and their classmates, resulting in an increase in TD. Interactive materials bring more learner–content interaction because the materials respond to the learners or give the learners feedback while they are learning. Interaction with the course content has been identified as a vital factor in the success of online learning (Pham, 2018). Ekwunife-Orakwue and Teng (2014) investigated 342 students who were taking either online or blended courses and found that among the three factors, learner–content interaction exercised a greater effect on the student's learning outcomes than other forms of interaction.

## 3. Method

This study puts forward an example of asynchronous learning for a college-level online EFL course with the purpose to investigate the influence of a student's self-efficacy on their perception regarding the interactive materials and the overall course. This study employed TDT as its theoretical framework. The study utilized a quantitative method and used questionnaires to collect the data. Because the students' perception is important for the planning and revision of our online curriculum, self-reported questionnaires were used along with their academic achievement scores to investigate the effects of self-efficacy on their performance and other affective factors.

### 3.1. Participants

This study recruited 286 participants (Female = 163, Male = 123) from two online courses at the undergraduate level in a northern private university in Taiwan during the spring semester of 2021. The participants were junior and senior students from non-English major departments of the university. These students had completed two years of English courses, mainly through physical, face-to-face teaching. In addition, the high-low difference in the level of self-efficacy related to English learning is obvious among non-English majors. These online courses delivered remedial English lessons to students who had not reached the required level of English proficiency before graduation. The university set the exit requirement that the students must graduate with B1 (intermediate level) in the Common European Framework of Reference (CEFR). All of the students consented to participate in the study by viewing an online explanatory video that was recorded by the instructor/researcher and clicking into the questionnaire links.

### 3.2. Interactive learning content

The 18-week online course contained one synchronous week and 17 asynchronous weeks. The first week was the course orientation, where the instructor and the students interacted in a virtual classroom simultaneously using Microsoft Teams as the live web-based video conferencing. The other 17 weeks were conducted asynchronously, where real-time interaction was almost unattested between the instructor and the students. The instructor used

Moodle as the learning management system (LMS) to provide the students with learning resources and academic support during the process while the students took responsibility for their own learning (Majeski et al., 2016).

The asynchronous classrooms in the present study incorporated the use of interactive multimedia materials that were designed and generated by the course instructor. The highly interactive materials were designed to make learning more engaging and effective for the students. The courses were highly structured, requiring the students to follow a certain order and complete all of the required tasks rendered in different kinds of interactive formats (e.g., videos, readings, HTML5 materials, exercises, quizzes, and gamification).

*Table 1.* Descriptions of interactive contents

| Interactive content format | Description |
|---|---|
| Video | The interactive videos used in the courses allowed the students to interact with the video content by clicking, dragging, hovering, and exercising in the middle of video playing. |
| Reading | The interactive reading used in the courses contained audio and clickable vocabulary explanations. |
| HTML5 materials | The HTML5 interactive materials used in the course focused on grammar and some common usages of phrasal verbs. For each section, the students had to correctly answer all the questions before they could proceed with the material. |
| Exercise | The interactive exercises used in the courses included different kinds of interactive activities for students to complete repetitively, including multiple choices, drag-and-drop, memory games, fill-in-the-blanks, flashcards, word puzzles, and so on. |
| Quiz | The interactive quizzes used in the courses automatically graded the student's scores, and some provided feedback after the student has answered the questions. |
| Gamification | The gamification material used in the courses included web-based escape rooms created by the instructor. Students needed to solve puzzles and riddles to complete the missions. The tasks that were designed to challenge students included interactive exercises, animations, videos, and games. |

### 3.3. Measurement tools

#### *3.3.1. Learning achievement test*

The 18-week class involved 30 quizzes. All of the quizzes were based on the textbooks and developed as well as revised by the instructor/researchers according to the content designed for each week. Each quiz consisted of 10–20 items. This study used the 10th week, which is arranged to be the midterm week at most universities in Taiwan, as the divide to split the quizzes into half. A total number of 14 quizzes were completed before the 10th week, and the average score of these 14 tests was adopted as the pre-test results for each student participating in the study. To demonstrate the construct validity of the pre-tests, three experts reviewed and revised the reading quizzes. Kuder–Richardson 20 was used to measure internal consistency. The researchers selected the article from Unit 2, Can Facebook Change Your Life, from the textbook, *Read to Succeed*, which was published by Live ABC, as it is the standard textbook for intermediate readers. This pre-test had an acceptable reliability of 0.69.

For the post-test results, the average score of the 16 quizzes completed after the 10th week was adopted. It was assumed that the students would be more familiar with learning with technology as the semester progressed. The more time that the students spent on online learning elevated their familiarity with technology and enhanced their ability to learn autonomously. To demonstrate the construct validity of the post-tests, three experts reviewed and revised the reading quizzes. Kuder-Richardson 20 was used to measure internal consistency and reliability. The researchers selected the article from Chapter 2, Travelers' Tales from *Essential Reading 4*, which was published by Macmillan as one of the post-tests because this is also the standard textbook for students of intermediate proficiency. This post-test had an acceptable reliability of 0.70.

#### *3.3.2. Learner autonomy, learner–content interaction, and transactional distance in distance learning*

A large part of the quantitative data was acquired by distributing and collecting the questionnaires to and from the participants. After reviewing the existing literature, we considered distance learning to be a multidimensional construct that requires the inspection of several aspects of the cognitive construct, in which self-perceived autonomy, self-perceived interaction with the learning content, and self-perceived transactional distance are integral elements. To gauge these three elements, the researchers adopted the constructs developed and validated by Huang et al. (2015) to further investigate the interaction of these elements and their relationship with self-efficacy.

To measure learner autonomy, Moore's (1991; 1997) original definition was adopted. Learner autonomy is related to self-directedness and a student's ability to control their learning process to achieve their goals. Therefore, two factors—the items related to the independence of learning (e.g., "Working on my own, I feel happy"), and the items related to study habits (e.g., "I enjoy less-structured courses which require me to take more control of my own learning")—were incorporated into the questionnaire. Expert judgments were incorporated to demonstrate content validity, and a total of 10 items were adopted to measure learner autonomy. The reliability of these questions was acceptable, with Cronbach's Alpha value being 0.78 for the pre-questionnaire and 0.83 for the post-questionnaire.

As regards learner–content interaction, we adopted the idea of "psychological or communication space" (Healey et al., 2008), which suggests how we communicate with the world affects our internal thoughts and self-perception. Therefore, the items related to the course design—such as course variety and individualization (e.g., "I receive individualized feedback on my assignments"; "The course is structured in a way that enables me to work at my own pace to meet the course goals and objectives") and formality (e.g., "Clear guidelines/rubrics on assignments, projects or other course-related tasks are provided for this online course")—were included. Expert judgments were used for content validity to gather constructive feedback (i.e., revisions and suggestions), and a total of 15 questions were adopted. The Cronbach's Alpha value for the pre-questionnaire was 0.89, while that for the post-questionnaire was 0.93.

To measure the learners' perception of transactional distance, the original questionnaire consisted of a set of items that were generated under two factors: learner–instructor interaction and learner–learner interactions. Although the questionnaire that was adopted in this study kept all of the items of the original, three experts revised them by adding the effect of interactive course materials into the formation of questions (e.g., "Interactive contents make me feel a strong sense of belonging to this online course"; "Through interactive contents, I feel closely connected to my instructor in this online course"). As regards the transactional distance questionnaire, the Cronbach's Alpha value for the pre-questionnaire was 0.92, while that for the post-questionnaire was 0.94—both values were at a good level.

#### *3.3.3. Self-efficacy questions*

A student's self-regulated and self-confident attitude during their participation encompasses the concept of self-efficacy, which is particularly important in technology-integrated classrooms. A student's self-efficacy in online learning courses is related to whether the online course meets their basic cognitive needs, such as their achievement in the course, their autonomy, and their interaction with the course materials. This study employed the questionnaire that was developed by Ngo and Eichelberger (2021) to reveal the student's perceived individual beliefs regarding their learning. A median split was used to categorize the sample into high and low self-efficacy (Marashi & Dakhili, 2015). The Cronbach's Alpha value for the questionnaire was 0.93.

### 3.4. Data collection procedures

All the students who volunteered to participate in the study were asked to complete two questionnaires. The self-efficacy questionnaire was collected at the end of the first week. The questionnaires, which probed their perceptions of autonomy, course content, and transactional distance, were conducted at two separate points in the semester. The first point was the end of the first week before any formal class activities occurred, while the second point was the end of the 17th week after the participating students had finished all the materials and tasks distributed on the learning platform. The 286 participants were classified as having either high or low self-efficacy based on the median split of their self-efficacy scores (Marashi & Dakhili, 2015; Ngo & Eihelberger, 2021). A significant between-group difference ($F$ = 478.09, $p$ < .001) was found in the high self-efficacy group ($N$ = 141, 49.3 %, Female = 82, Male = 59, Mean = 4.08 $SD$ = 0.38) and the low self-efficacy group ($N$ = 145, 50.7%, Female = 81, Male = 64, Mean = 2.91, $SD$ = 0.51). As the semester progressed, the scores of their achievement tests were also collected.

## 4. Results

### 4.1. Learning achievement

The students taking this online course were required to complete one or two topics of learning materials per week. Each topic contained one or two graded quizzes that evaluated the learner's overall understanding of the online materials. A total of 30 scores were collected during the semester. The quizzes were split into pre-test and post-test to compare their learning outcome after the extensive implementation of asynchronous instructions.

A one-way ANCOVA was performed to analyze the participants' learning achievements by using their pre-test as a covariate, self-efficacy as an independent variable, and the post-test as a dependent variable. The homogeneity of regression was not violated ($F$ = 2.72, $p$ = .10 > .05), showing a common regression coefficient for one-way ANCOVA. The examination of the self-efficacy in terms of learning achievement through the ANCOVA method showed that there were significant differences between the high self-efficacy group (Adjusted mean = 70.69) and the low self-efficacy group (Adjusted mean = 67.88) with $F$= 4.37 ($p$ = .04 < .05), as presented in Table 2. The results indicate that the learning achievement of high self-efficacy learners was significantly higher than that of low self-efficacy learners.

*Table 2.* ANCOVA of the post-test for achievement

| Group | *N* | Mean | *SD* | Adjusted mean | *SE* | *F* | $\eta^2$ |
|---|---|---|---|---|---|---|---|
| High self-efficacy | 141 | 71.31 | 11.58 | 70.69 | 0.95 | $4.37^*$ | 0.02 |
| Low self-efficacy | 145 | 67.28 | 13.73 | 67.88 | 0.94 | | |

*Note.* $^*p < .05$.

### 4.2. Learner autonomy

Being autonomous reflects the psychological characteristics of individuals. This specific personal trait indicated whether these learners were able to independently direct their learning in the asynchronous English learning environment. The psychometrical scale from Huang et al. (2015) was adopted to measure the participants' perceptions regarding learner autonomy. A one-way ANCOVA was performed to analyze the participants' perceptions regarding learner autonomy by using their pre-questionnaire scores as a covariate, self-efficacy as an independent variable, and the post-questionnaire score as a dependent variable. The homogeneity of regression was not violated ($F$ = 3.62, $p$ = .06 > .05), showing a common regression coefficient for one-way ANCOVA. The examination of the self-efficacy in terms of the learner autonomy through the ANCOVA method showed that there were significant differences between the high self-efficacy group (Adjusted mean = 4.46) and the low self-efficacy group (Adjusted mean = 3.77) with $F$ = 6.08 ($p$ = .05), as presented in Table 3. The results reveal that the learners with high self-efficacy had significantly higher autonomy than the learners with low self-efficacy.

*Table 3.* The ANCOVA results for the learner autonomy of the two groups

| Group | *N* | Mean | *SD* | Adjusted mean | *SE* | *F* | $\eta^2$ |
|---|---|---|---|---|---|---|---|
| High self-efficacy | 141 | 4.13 | 0.51 | 4.05 | 0.05 | $4.23^*$ | 0.01 |
| Low self-efficacy | 145 | 3.84 | 0.69 | 3.18 | 0.05 | | |

*Note.* $^*p < .05$.

### 4.3. Learner–content interaction

Because learner–content interaction can be an indicator of self-paced learning, its effect was statistically tested in this study design. In contrast to real-time, off-campus synchronous learning, asynchronous learning in higher education offers flexibility for students to manage their own learning. Although asynchronous learning does not require interactions that come with scheduled class meetings, successful asynchronous learners demonstrate the ability to interact with the contents at their own pace. This self-paced orientation is important for meaningful learning to take place. Therefore, a one-way ANCOVA was performed to measure the participants' perceptions regarding learner–content interaction by using their pre-questionnaire scores as a covariate, the two groups of high and low self-efficacy students as an independent variable, and the post-questionnaire score as a dependent variable. The homogeneity of regression was not violated ($F$ = 0.90, $p$ = .34 > .05), showing a common regression coefficient for one-way ANCOVA. The examination of the effectiveness of the self-efficacy in terms of the learner–content interaction through the ANCOVA method showed that there were significant differences between the high self-efficacy group (Adjusted mean = 4.46) and the low self-efficacy group (Adjusted mean = 3.91) with $F$ = 3.92 ($p$ = .04 < .05), as presented in Table 4. The findings demonstrate that the learners with high self-efficacy interacted significantly more with the content than the learners with low self-efficacy.

*Table 4.* The ANCOVA result for the learner–content interaction of the two groups

| Group | $N$ | Mean | $SD$ | Adjusted mean | $SE$ | $F$ | $\eta^2$ |
|---|---|---|---|---|---|---|---|
| High self-efficacy | 141 | 4.13 | 0.63 | 4.06 | 0.05 | $3.92^{*}$ | 0.01 |
| Low self-efficacy | 145 | 3.85 | 0.70 | 3.91 | 0.05 | | |

*Note.* $^{*}p < .05$.

### 4.4. Transactional distance

Given that online asynchronous English learning relies on technologically mediated interactions, the perception of psychological distance has long been considered an important construct. This aspect of technological mediation was investigated through statistical measures to obtain the full picture of the student's online learning perception. A one-way ANCOVA was performed to examine the participants' transactional distance by using their pre-questionnaire scores as a covariate, the self-efficacy as an independent variable, and the post-questionnaire score as a dependent variable. The homogeneity of regression was not violated ($F$ = 2.28, $p$ = .13 > .05), showing a common regression coefficient for one-way ANCOVA. The examination of the effectiveness of the self-efficacy in terms of the transactional distance through the ANCOVA method showed that there were no significant differences between the high self-efficacy group (Adjusted mean = 3.90) and the low self-efficacy group (Adjusted mean = 3.78) with F = 2.28 ($p$ = .13 > .05), as presented in Table 5. The results indicate no significant difference in the perception of transactional distance between the learners with high and the learners with low self-efficacy.

*Table 5.* ANCOVA Analysis of perception toward transactional distance

| Group | $N$ | Mean | $SD$ | Adjusted mean | $SE$ | $F$ | $\eta^2$ |
|---|---|---|---|---|---|---|---|
| High self-efficacy | 141 | 3.98 | 0.68 | 3.90 | .056 | 2.28 | .01 |
| Low self-efficacy | 145 | 3.71 | 0.75 | 3.78 | .055 | | |

### 4.5. The relationship between learner autonomy, learner–content interaction, transactional distance, self-efficacy, and learning achievement

With the collected quantitative data, the researchers examined the relationships among the variables in the context of technology-mediated asynchronous English learning. The relationships between the indicators of transactional distance and the students' learning achievements are important for future curriculum design. As Table 6 demonstrates, learner autonomy correlated with learner–content interaction, transactional distance, and self-efficacy; learner–content interaction correlated with transactional distance and self-efficacy; transactional distance correlated with self-efficacy; and self-efficacy correlated with learning achievement. Most importantly, there was a statistically significant correlation between self-efficacy and learning performance, learner autonomy, learner–content interaction, and transactional distance. Therefore, in this study, self-efficacy can be interpreted as an essential factor of asynchronous online courses.

*Table 6.* Correlations of the variables

| | Learning achievement | Learner autonomy | Learner–content interaction | Transactional distance |
|---|---|---|---|---|
| Learner autonomy | .055 | 1 | | |
| Learner–content interaction | .065 | $.822^{**}$ | 1 | |
| Transactional distance | .011 | $.767^{**}$ | $.790^{**}$ | 1 |
| Self-efficacy | $.175^{**}$ | $.607^{**}$ | $.515^{**}$ | $.554^{**}$ |

*Note*. $N$ = 286; $^{**}p < .01$.

## 5. Discussion

The present study demonstrated that in an asynchronous online course that is highly structured and incorporates highly interactive learning contents, self-efficacy could be a predicator of a student's learning achievement. We found significant differences between the students with high and low self-efficacy in their learning achievements, learner autonomy, and their perceptions of learner-content interaction. However, no statistically significant relationship was found between the learners with different levels of self-efficacy and their perceived transactional distance. The findings also indicate that all of the variables correlated with each other, whereas learning achievement only correlated with self-efficacy.

### 5.1. Self-efficacy predicts students' achievements, autonomy, and interactions with content

The findings indicate that when learners were provided with interactive learning content, there was a statistically significant relationship between these variables—learners with different levels of self-efficacy performed differently in terms of learning achievement, developed different perception regarding learner autonomy, and demonstrated different learner–content interactions. These findings echo those of the previous studies that associated self-efficacy with academic achievement (Caprara et al., 2011; Honicke & Broadbent, 2016; Talsma et al., 2018; Wang & Bai, 2017). Further, the present study revealed that self-efficacy also correlates with the affective factors that are important to students' self-perception when taking online courses, including their self-perception regarding whether they can take responsibility of their own learning or not and whether their interaction with the course content constitutes a successful experience. Asynchronous classes incorporate technologies in students' learning to promote their interaction with course materials so they can better comprehend the content. It is important for online learners to develop a positive self-perception regarding their online learning experiences. The findings of our study indicate that self-efficacy is an important factor that affects not only a student's learning achievement but also how they react to online activities, manage their time to organize learning procedure, and complete the learning process to fulfill the requirements set by the instructor. Given that high self-efficacious students are more autonomous and interact significantly more with the content than low self-efficacious students, all of these positive self-perceptions assist the students in achieving more successful learning outcomes.

### 5.2. A well-structured asynchronous course triggers the two-sided effects of transactional distance

The findings also indicate that in a highly-structured asynchronous course constructed primarily by interactive contents, the perceived distance to the online course does not show differences between the high- and low-efficacious students. This result explains the two-sided effects of transactional distance. In one respect, TD should increase if the teacher's role disappears in the classroom, whereas in another respect, the interactive content reduces TD because a course incorporating rich interactive learning content keeps the students engaged, dissolving the teacher's role in the curriculum.

When interactive elements are blended in an asynchronous online course, access to interactive content makes individualized learning possible. The interactive content provides instant feedback to the students during the process of learning and allow them to learn through playing. An asynchronous teacher is part of the course content. Through the materials, the teacher interacts with the students, supervises their progress, and invisibly evaluates their performance. Interactive learning content strikes a balance between developing learner autonomy and fostering course engagement by making the teacher's role invisible. Swan (2002) argued that three factors—a clear and consistent course structure, frequent instructor-learner interactions, and a valued and dynamic discussion—are key to the success of an online course. However, a well-structured online classroom requires the instructor to interfere with the student's learning less because highly interactive content encourages the students to personalize and control their own pace of learning in a restrictive way. Moreover, interactive content nullifies the need for students to interact with the instructor because the teacher's role dissolves into the material. In a highly organized online course, the teacher may seem to be invisible, but, in fact, the teacher is everywhere. The students do not need to be physically close to the teacher if the instructions are clearly conveyed through the interactive materials. This can be verified through the student posts in every week's Q&A section. The two classes collected a total number of 61 posts, and none of the feedback was about the content-related clarifications. The students mostly posted comments to inquire about course management or to report LMS problems.

### 5.3. Transaction distance, learner autonomy, and interactive content are interwoven

The findings of this study revealed that the three affective variables (i.e., TD, learner autonomy, and learner–content interaction) are correlated. Moreover, these variables also correlate with self-efficacy, which, in turn, correlates with learning achievement. A major goal of education is to develop a student's autonomy in learning (Bembenutty, 2011), but self-autonomy is not necessarily connected to their increased or decreased TD when learning occurs in an online setting characteristic of interactive contents. Asynchronous courses force the rules and regularities of student learning, as a well-organized online course requires the instructor to set clear goals, plan the detailed learning procedures, prepare learning resources, and determine appropriate evaluative methods for the learners. The interactive features that are embedded in the learning content forces the students to obey the regulations and follow the same path to complete the tasks. However, autonomous learners are innovative and creative. When it comes to learning, innovation and creativity pull the students out of regularity, often leading to a weaker bond with the course. In another aspect, interactive course content requires the students to become more autonomous in terms of learning, as they have to spare some time every week to complete all of the designated material and exercises. As the semester progresses, the more time that they have spent on interactive materials familiarizes them with online learning and improves their overall ability of learning with technology. Bandura (1986; 1997) identified mastery experiences as one of the four key sources of self-efficacy. In the educational setting, mastery experiences refer to a student's successful experiences in performing tasks. In contrast to traditional classrooms, students in asynchronous classrooms cannot simply sit without receiving any information from the lecture. Asynchronous classrooms reduce the student's mental and psychological stress by allowing them to repeat the material as many times as they need until they feel confident in their learning. Besides, they can also access the material at any time to refresh their memory and comprehension of the material. Students also have to actively participate in the learning activities to complete their weekly assignments. When interactive materials lead to more successful experiences for the students, they increase their learning achievements through their increased self-efficacy with online learning.

## 6. Conclusion

As English teachers are increasingly moving toward online teaching, it has become imperative to understand how the different variables related to a learner's perception interact with digitalized interactive learning contents to produce better learning outcomes with online instructions. Our statistical findings confirmed that a student's learning achievement in an asynchronous classroom can be enhanced by higher self-efficacy. Therefore, online instructors need to stress the importance of improving self-efficacy among students. For this, the instructors are encouraged to use interactive content to help their students establish more successful learning experiences. The flexibility and repeatability of online resources allow students to establish their own learning routines and determine their own pace of learning. Further, the instant feedback and interactive features provided by online learning materials can keep students engaged. Successful learning experiences can enhance a student's autonomous learning, while a stronger learner autonomy facilitates improved self-efficacy, which predicts the opportunity for better academic success.

In this study, we sought to examine how TDT (Moore, 1997) works in the context of an asynchronous online English course. Our findings have successfully established a relationship among the three learning variables that are considered important elements of describing a student's learning perception. For a long time, TDT has been used to explain the mechanisms of online education. The present study found that the difference in the overall transactional distance between the two groups of students and the course itself was not statistically significant because the relationship between TD and other affective factors is complex. This suggests that each factor is closely interwoven with the other factors, and there is no straightforward way to discuss any of them individually.

By combing all these factors, the results of this study bring the importance of interactive learning contents to the forefront. In an asynchronous environment, interactive learning content supports self-determined learning for both high and low self-efficacious learners. These learners accumulated successful learning experiences through the repetitive use of interactive materials. Positive learning experiences trigger higher course engagement, while higher engagement, in turn, brings in more positive experiences, and the whole process leads to a successful learning. Moreover, interactive learning materials encourage autonomous learning. When students become independent learners, the asynchronous classroom turns student-centric, and online learning no longer remains teacher-centric. In this context, students receive education not only from the instructor but also through the use of interactive learning content to discover their ability as independent learners. When interactive content is underpinned by a well-organized structure, it constrains autonomous learning by forcing students to obey course regulations and complete all of the tasks on time. Pedagogically speaking, the opposing but not conflicting features of interactive materials have successfully pushed the teacher's role behind the scene. Today, students can benefit from autonomous learning, even if the teacher is absent. However, the preset interactive activities must guarantee that the flexibility and repeatability of learning come with limitations.

Finally, a student's attitude toward English learning might not be persistent when they are learning with different types of interactive contents. This could trigger a positive learning effect for students with lower levels of English proficiency. Interactive learning resources can facilitate students who fail in traditional ways of teaching by giving them control over their own learning. The learning process is multifaceted and complicated, as students with different learning styles prefer different kinds of material and do not respond to all kinds of materials similarly. Therefore, future research could go one step further by investigating how academic self-efficacy might fluctuate when different interactive learning materials are used in an asynchronous setting, along with how they affect a student's academic engagement. Additionally, future studies might also be able to explain the mechanism behind the relationship between interactive materials and learning achievement for EFL learners.

## Statements on open data, ethics, and conflict of interest

The participants were protected by hiding their personal information during the research process. They knew that the participation was voluntary and they could withdraw from the study at any time. There is no potential conflict of interest in this study. The data can be obtained by sending a request email to the corresponding author.

## Acknowledgement

The research is supported by the Grants and awards for distance courses, Center for Distance Education Development, Tamkang University, and partially supported by the National Science and Technology Council under contract numbers NSTC-111-2410-H-031-092-MY2 and Applied AI research, Soochow University under contract numbers 111160605-0014.

## References

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84,* 191-215.

Bandura, A. (1986). *Social foundations of thought and action: A Social cognitive theory*. Prentice-Hall.

Bandura, A. (1991). Social cognitive theory of self- regulation. *Organizational Behavior and Human Decision Processes, 50,* 248-287.

Bandura, A. (1997). *Self-efficacy: The Exercise of control*. Freeman.

Bembenutty, H. (2011). Introduction: self-regulation of learning in postsecondary education. In H. Bembenuty (Ed.), *Self-Regulated Learning: New Directions for Teaching and Learning* (pp. 3–8). Jossey-Bass, Wiley. https://doi.org/10.1002/tl.439

Benson, R., & Samarawickrema, G. (2009). Addressing the context of e-learning: Using transactional distance theory to inform design. *Distance Education, 30,* 5–21.

Benton, S. L., Li, D., Gross, A., Pallett, W. H., & Webster, R. J. (2013). Transactional distance and student ratings in online college courses. *American Journal of Distance Education, 27*(4), 207–217.

Caprara, G. V., Vecchione, M., Alessandri, G., Gerbino, M., & Barbaranelli, C. (2011). The Contribution of personality traits and self-efficacy beliefs to academic achievement: A Longitudinal Study. *British Journal of Educational Psychology, 81*(1), 78–96.

Charles, D., Charles, T., McNeill, M., Bustard, D., & Black, M. (2011). Game-based feedback for educational multi-user virtual environments. *British Journal of Educational Technology, 42*(4), 638-654. https://doi.org/10.1111/j.1467-8535.2010.01068.x

Chen, Y. J. (2001a). Transactional distance in World Wide Web learning environments. *Innovations in Education and Teaching International, 38*(4), 327–338.

Chen, Y. J. (2001b). Dimensions of transactional distance in the world wide web learning environment: A Factor analysis. *British Journal of Educational Technology, 32*(4), 459–470.

Cho, M. H., & Shen, D. (2013). Self-regulation in online learning. *Distance Education, 34,* 290–301. https://doi.org/10.1080/01587919.2013.835770

Crippen, K. J., Biesinger, K. D., Muis, K. R., & Orgill, M. (2009). The Role of goal orientation and self-efficacy in learning from web-based worked examples. *Journal of Interactive Learning Research, 20,* 385–403.

Cussó-Calabuig, R., Farran, X. C., & Bosch-Capblanch, X. (2018). Effects of intensive use of computers in secondary school on gender differences in attitudes towards ICT: A Systematic review. *Education and Information Technologies, 23,* 2111–2139.

Deci, E. L., & Ryan, R. M. (1987). The Support of autonomy and the control of behavior. *Journal of Personality and Social Psychology, 53*(6), 1024–1037.

Domagk, S., Schwartz, R. N., & Plass, J. L. (2010). Interactivity in multimedia learning: An Integrated model. *Computers in Human Behavior, 26,* 1024–1033. https://doi.org/10.1016/j.chb.2010.03.003.

Ekwunife-Orakwue, K. C., & Teng, T. L. (2014). The Impact of transactional distance dialogic interactions on student learning outcomes in online and blended environments. *Computers & Education, 78*, 414–427.

Feldman, D. B., & Kubota, M. (2015). Hope, self-efficacy, optimism, and academic achievement: Distinguishing constructs and levels of specificity in predicting college grade-point average. *Learning and Individual Differences, 37,* 210-216.

Healey, P. G. T., White, G., Eshghi, A., Reeves, A. J., & Light, A. (2008). Communication spaces. C*omputer Supported Cooperative Work (CSCW), 17*, 169–193.

Holec, H. (1981). *Autonomy in foreign language learning*. Oxford.

Honicke, T., & Broadbent, J. (2016). The Influence of academic self-efficacy on academic performance: A Systematic review. *Educational Research Review, 17*, 63–84.

Huang, X., Chandra, A., DePaolo, C. A., & Simmons, L. L. (2016). Understanding transactional distance in web-based learning environments: An empirical study. *British Journal of Educational Technology, 47*(4), 734-747.

Huang, X., Chandra, A., DePaolo, C., Cribbs, J., & Simmons, L. (2015). Measuring transactional distance in web-based learning environments: An Initial instrument development. *Open Learning: The Journal of Open, Distance and e-Learning, 30*(2), 106–126. https://doi.org/10.1080/02680513.2015.1065720

Kara, M. (2021). Transactional distance and learner outcomes in an online EFL context. *Open Learning: The Journal of Open, Distance and e-Learning, 36*(1), 45-60.

Kaplar, M., Radović, S., Veljković, K., Simić-Muller, K., & Marić, M. (2022). The Influence of interactive learning materials on solving tasks that require different types of mathematical reasoning. *International Journal of Science and Mathematics Education, 20,* 411–433. https://doi.org/10.1080/02680513.2020.1717454

Kitsantas, A., & Chow, A. (2007). College students' perceived threat and preference for seeking help in traditional, distributed and distance learning environments. *Computers & Education, 48*, 383–395. https://doi.org/10.1016/j.compedu.2005.01.008

Lai, C., & Morrison, B. (2013). Towards an agenda for learner preparation in technology-enhanced language learning environments. *CALICO Journal, 30*(2), 154–162.

Lim, C. K. (2001). Computer self-efficacy, aca- demic self-concept, and other predictors of satisfaction and future participation of adult distance learners. *American Journal of Distance Education, 15*(2), 41-51.

Majeski, R. A., Stover, M., & Ronch, J. (2016). Making asynchronous online learning more learner-oriented: An Integrated conceptual model with applications for course design and instruction. *Educational Gerontology, 42*, 109–119.

Marashi, M. & Dakhili, M. (2015). Self-efficacy and anxiety among EFL learners with different kinds of multiple intelligences. *Theory and Practice in Language Studies, 5*(12), 2636-2645.

Moore, M. G. (1984). On a theory of independent study. In D. Stewart, D. Keegan, & B. Holmberg (Eds.), *Distance education: International perspectives* (pp. 68–94). Routledge.

Moore, M. G. (1991). Editorial: Distance education theory. *American Journal of Distance Education, 5*(3), 1–6. https://doi.org/10.1080/08923649109526758

Moore, M. G. (1997). Theory of transactional distance. In D. Keegan (Ed.), *Theoretical Principles of Distance Education* (pp. 22-38). Routledge.

Ngo, H., & Eichelberger, A. (2021). College students' perceived self-efficacy and use of information and communication technologies in EFL learning. *International Journal of Education and Development using ICT, 17*(1), 34-44. http://ijedict.dec.uwi.edu/viewarticle.php?id=2875

Park, Y. (2011). A Pedagogical framework for mobile learning: Categorizing educational applications of mobile technologies into four types. *International Review of Research in Open and Distributed Learning, 12*(2), 78–102.

Pham, N. T. (2018). Learner-content interaction in an online English learning course at a Vietnamese university. *VNU Journal of Foreign Studies, 34*(5). https://doi.org/10.25073/2525-2445/vnufs.4307

Schlosser, L. A., & Simonson, M. (2009). *Distance education: Definition and glossary of terms* (3rd ed.). Information Age Publishing.

Schunk. D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist, 26*, 207-132.

Shea, J., Joaquin, M. E., & Wang, J. Q. (2016). Pedagogical design factors that enhance learning in hybrid courses: A Contribution to design-based instructional theory. *Journal of Public Affairs Education, 22*(3), 381–397.

Sulha, A. H., Famela, A. K., & Harahap A. T. A. (2021). The Implementation of synchronous and asynchronous learning in English as foreign language setting. *Indonesian Journal of Research and Educational Review, 1*(1), 17-27.

Swan, K. (2002). Building learning communities in online courses: The Importance of interaction. *Education, Communication & Information*, *2*(1), 23-49. https://doi.org/10.1080/1463631022000005016

Swart, W., & Wuensch, K. L. (2016). Flipping quantitative classes: A Triple win. *Decision Sciences Journal of Innovative Education, 14,* 67-89.

Talsma, K., Schüz, B., Schwarzer, R., & Norris, K. (2018). I believe, therefore I achieve (and vice versa): A Meta-analytic cross-lagged panel analysis of self-efficacy and academic performance. *Learning and Individual Differences, 61*, 136–150.

Wang, C., Kim, D. H., Bai, R., & Hu, J. (2014). Psychometric properties of a self-efficacy scale for English language learners in China. *System, 44,* 24–33.

Wang, C., & Bai, B. (2017). Validating the instruments to measure ESL/EFL learners' self-efficacy beliefs and self-regulated learning strategies. *TESOL Quarterly, 51*(4), 931-947.

Wang, C., Schwab, G., Fenn, P., & Chang, M. (2013a). Self-efficacy and self-regulated learning strategies for English Language learners: Comparison between Chinese and German College students. *Journal of Educational and Developmental Psychology, 3*(1). https://doi.org/10.5539/jedp.v3n1p173

Wang, C., Shannon, D. M., & Ross, M. E. (2013b). Students' characteristics, self-regulated learning, technology self-efficacy, and course outcomes in online learning. *Distance Education, 34*(3), 302-323.

Yukselturk, E., & Bulut, S. (2007). Predictors for student success in an online course. *Educational Technology & Society, 10*(2), 71–83.

Zhang, A. M. (2003). Transactional distance in web-based college learning environments: *Toward measurement and theory construction* (Unpublished doctoral dissertation). Commonwealth University, Virginia.

Shadiev, R., & Zhou, J. (2023). A Review of Research on the Use of 360-degree Video Technology in Language Learning. *Educational Technology & Society*, *26*(4), 14-37. https://doi.org/10.30191/ETS.202310_26(4).0002

# A Review of Research on the Use of 360-degree Video Technology in Language Learning

**Rustam Shadiev[1*] and Junpei Zhou[2]**
[1]College of Education, Zhejiang University, China // [2]School of Education Science, Nanjing Normal University, China // rustamsh@gmail.com // zhoujp_99@163.com
[*]Corresponding author

(Submitted July 8, 2022; Revised February 6, 2023; Accepted February 27, 2023)

**ABSTRACT:** The present study is set to systematically review articles on the use of 360-degree video technology in language learning. The study selected and reviewed twenty-four articles in the following aspects: (1) tools related to 360-degree video technology; (2) languages and skills involved; (3) theories and pedagogical approaches in reviewed articles; (4) methodology of reviewed studies; (5) applications of 360-degree video technology to language learning; (6) reported findings; and (7) reported problems in reviewed studies. The results demonstrated that the tools related to 360-degree video technology can be grouped according to the following three ways of using them: (1) creating or editing videos/images, (2) obtaining videos/images, and (3) viewing videos/images. The participants in most studies recorded or edited 360-degree videos to develop their own learning content, rather than using existing ones. In most studies, the participants used head-mounted displays (HMDs) to view 360-degree videos and low-cost HMDs were used more frequently. Scholars often focused on English and Chinese, and they targeted speaking and writing skills in their research. Various theories were used to frame research and the embodied cognition theory was the most popular. The most commonly used pedagogical approach was task-based learning. Fewer studies focused on students from primary or junior school. Many studies lasted for more than one month. Different language skills were mainly measured using scales or tests. Findings related to learning outcomes, learners' perceptions of using 360-degree video technology and motivation were most frequently reported in the reviewed studies. Finally, problems related to methodology, technology implementation and learning process were identified in the reviewed studies and they are reported in the present research. Based on the results, several suggestions were made and implications derived.

**Keywords:** 360-degree video technology, Language learning, Review

## 1. Introduction

360-degree video technology, also called spherical video-based virtual reality (Ye et al., 2021), refers to technology that helps create and watch 360-degree videos and images. Such videos or images can be created by using 360-degree cameras (Rupp et al., 2019). 360-degree videos or images can be viewed by using a smartphone, tablet, computer or head-mounted display (Snelson & Hsu, 2020). Viewers can pan and tilt the phone or use the mouse or the keyboard arrows when using a computer to choose what they want to see. When using HMD, a viewer can turn the head to control the viewing direction (Repetto et al., 2021; Rupp et al., 2019).

To date, 360-degree video technology has been widely applied in various domains of knowledge, e.g., medical education and healthcare (Fukuta et al., 2021; Zulkiewicz et al., 2020), science (Wu et al., 2021), language learning (Huang et al., 2020; Repetto et al., 2021), and sports (Kittel et al., 2019). Researchers explained that 360-degree video technology creates such learning environments in which ethical principles can be maintained and problems related to time and space can be overcome (Concannon et al., 2019). Furthermore, such environments can help virtually experience dangerous situations and increase physical accessibility (Geng et al., 2021). For example, in the study by Herault et al. (2018), medical students learned how to treat patient trauma and to communicate with patients and their relatives effectively in authentic situations created by 360-degree video technology. In the study by Li et al. (2012), students learned about tower crane dismantling (i.e., one of the most dangerous activities in the construction industry) and practiced their skills using 360-degree video technology. Wu et al. (2021) explored the effects of applications of 360-degree video technology to scientific inquiry instruction on learners' problem-solving abilities. Liu et al. (2020) employed 360-degree video technology and cyclists watched them to study sidewalks and paved shoulders. The results reported by scholars such as knowledge gain or acquisition of certain skills were positive in most cases.

Scholars attempted to review existing studies on 360-degree video technology and their applications to assist learning and instruction. Shadiev et al. (2021a) systematically reviewed fifty-two research articles on 360-degree video technology and its applications in the field of education published between 2015 and 2020. They focused on exploring tools used in the reviewed articles, theories used in reviewed articles, methodologies that were applied by researchers and reported findings. Pirker and Dengel (2021) carried out a systematic review of sixty-four research articles to explore the potential of 360-degree video technology for education. Pirker and Dengel (2021) discussed the use cases, advantages, and limitations of 360-degree video technology in their study. Snelson and Hsu (2020) reviewed twelve research articles on 360-degree video technology applications in education published between 2017 and 2019. The scholars focused on the extent and nature of research on the educational use of 360-degree video technology and the benefits or drawbacks of applications in reviewed studies.

The overview of related review studies showed that they all focus on applications of 360-degree video technology in education in general. That is, there are no studies that explore the use of 360-degree video technology in specific domain knowledge. Therefore, educators and researchers need to address this gap in the literature. So we make the first step in this direction, and as experts in technology-assisted language learning field we carry out the present research to explore the use of 360-degree video technology in language learning by systematically reviewing related research articles.

To begin with, 360-degree video technology has been recognized by scholars as a potential tool to assist language learning. In real life, language learners often lack the authentic language learning environment to practice target languages (Shadiev & Yu, 2023). Fortunately, 360-degree video technology, as one of the types of virtual reality, can provide them with a realistic learning environment because 360-degree video is generated with real-world footage (Snelson & Hsu, 2020). That is, an authentic environment created by 360-degree video technology enables learners to feel immersed and sense of presence in a real language learning environment (Huang et al., 2020). Unlike traditional 3D animation-based virtual reality with complex techniques to develop and expensive cost, 360-degree video technology is affordable and easy to use by instructors and language learners in school settings (Huang et al., 2020). For example, there are many studies in which teachers and students created their own 360-degree video and image content and then used it in the language learning process (Chen & Hwang, 2020; Nobrega & Rozenfeld, 2019).

The literature on 360-degree video technology-assisted language learning is growing, however, as we mentioned earlier, there are no studies that systematically reviewed them. For example, Peixoto et al. (2021) and Parmaxi (2023) focused more on language learning assisted by immersive VR created by using computer-based 3D technology. Dhimolea et al. (2022) suggested that VR can be low- (LiVR) and high-immersive (HiVR). Users experience LiVR on a flat screen and interact with VR content using a mouse or keyboard. On the other hand, users experience HiVR using HMD and interact with content by using the buttons on HMD, a controller, or haptic systems. HMD presents an artificial environment that replaces or replicates users' real-world surrounding contexts so convincingly that users perceive the created environment as being spatially realistic and fully engage with it (Shadiev & Li, 2023). Compared to LiVR, HiVR provides higher levels of immersion. Different technologies can be used to create HiVR; for example, computer-based 3D technology or 360-degree video technology (Kim et al., 2022). These two technologies are different from each other in terms of cost, authenticity, presence and flexibility (see Shadiev et al., 2021a for more details). For example, 360-degree video technology presents content recorded by a camera, so objects, people, and scenes in virtual reality look the same as they are in the real world. For this reason, the degree of authenticity is higher and the cost is lower for virtual reality created by 360-degree video technology.

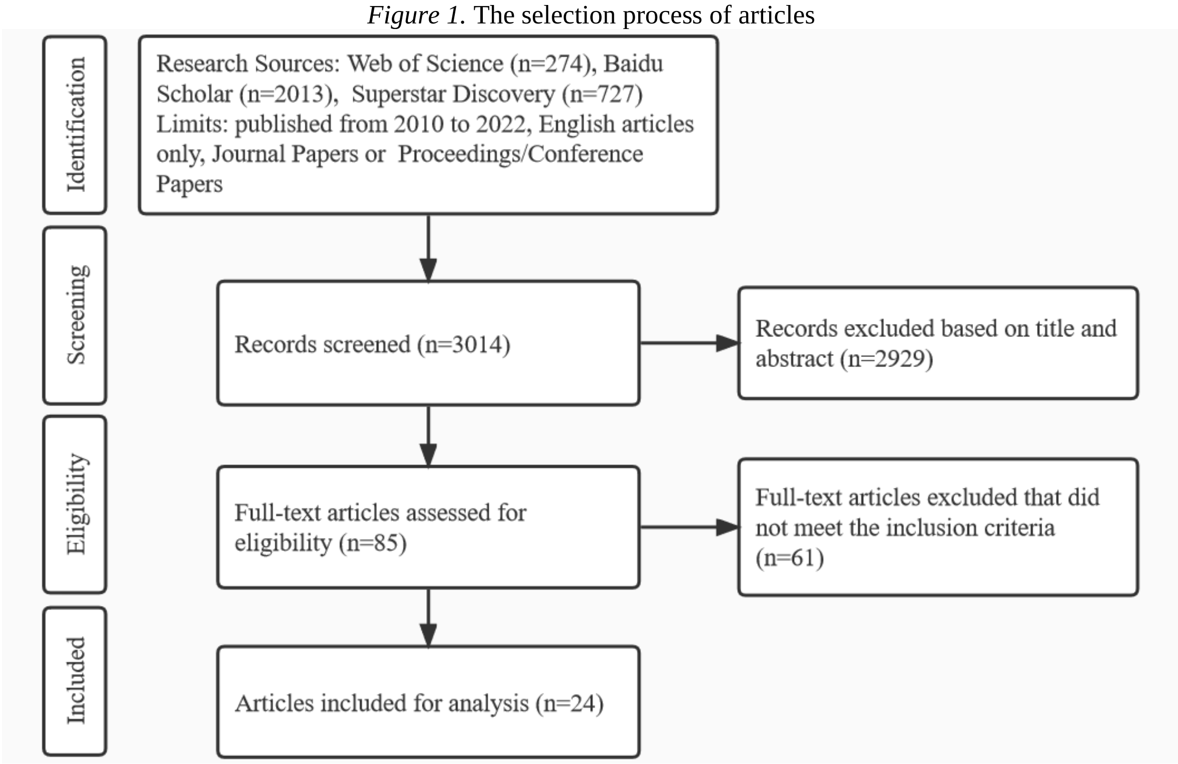
Therefore, the present review study goes beyond existing studies. First, the scopes of the review studies are different, i.e., previous studies explored PC-based immersive VR or they focused on education in general. In contrast, our study specifically focuses on the usage of 360-degree video technology in language learning context. Since 360-degree video technology has become popular and is being used in language learning in recent years, it is necessary to conduct this review study to fill the existing gap in. Second, dimensions of review studies are different. We consider some important dimensions that were rarely or not explored at all in other studies, such as pedagogical approaches or applications of 360-degree video technology in language learning. Related studies mentioned that 360-degree video technology is developing fast. Many kinds of this technology exist nowadays, and they can be used in different ways (e.g., not only to watch videos or pictures but also to create learning content as a part of the learning process) (Shadiev et al., 2021a). As we focus on the use of 360-degree video technology in the field of language learning, it is very important to know what languages were involved and what skills were targeted in the reviewed articles. For example, what are the most popular languages and skills in this field, and what languages and skills received little attention. Theories were reported in earlier review studies, but because they focused on education in general, such information may have little relevance to the field of language learning. Similarly, methodologies, findings and reported problems can be different between two contexts, i.e., education in general and language learning. Therefore, the present study is set to provide much needed information that can inform and guide future research. To this end, in the present study, the following research questions were addressed: (1) What types of 360-degree video technology were used for language learning in reviewed articles? (2) What languages and skills researchers targeted in the reviewed articles? (3) What theories and pedagogical approaches did researchers use in reviewed studies? (4) What methodologies were applied to the reviewed studies? (5) How was 360-degree video technology applied in language learning? (6) What findings did the researchers report? (7) What were reported problems?

## 2. Method

The present review was carried out using the preferred reporting items for systematic reviews and meta-analyses (PRISMA) methodology. It is a generally accepted approach that helps scholars prepare and report systematic reviews and meta-analyses (Shadiev & Wang, 2022). Articles for the present review were searched on Web of Science, Baidu Scholar and Superstar Discovery databases. Web of Science "is one of the most extensive, popular and relevant research databases for the academic community" (Caseiro & Santos, 2018, p. 8). Web of Science contains about 9,000 kinds of world authoritative journals in many different academic disciplines, and it features rich and powerful search functions (Shadiev & Feng, 2023; Shadiev & Li, 2023; Snelson & Hsu, 2020). We also used such Chinese databases as Superstar Discovery and Baidu Scholar because our search for the articles was carried out in China. These two databases are comprehensive China-based search platforms that provide access to a large amount of local and international literature. Therefore, all these three databases are considered as the most authentic citation databases, offer citation indexing of the social sciences, and have been used in many review studies carried out locally and internationally. Considering the two aspects (i.e., 360-degree video technology and language learning) of the subject of our review, and the definition of 360-degree video technology, we identified search keywords and combined them into the following Boolean search string: "360 video" OR "spherical video" OR "panoramic" OR "virtual tour" OR "360 VR" OR "360 virtual reality" AND "language". The following inclusion criteria were adopted in the selection of research articles: (1) articles published from 2010 to 2022 (March); (2) full-text of articles was available; (3) articles were published in English; (4) in journals or conference proceedings; and (5) articles focused on the applications of 360-degree video technology to assist language learning.

The systematic search process of the study is shown in Figure 1. A total of 3,014 papers were found from the search. After titles and abstracts were screened, 2,929 articles that did not match our selection criteria were removed. After reviewing the main text, 61 articles were excluded, and so, 24 articles were selected for the present review. Two researchers were involved in the search and selection process. They searched for articles, independently examined all of them and selected relevant articles based on the above-mentioned criteria. Once there was a difference during the selection process, the researchers discussed it until their agreement.

*Figure 1.* The selection process of articles



To answer the research questions of the study, the researchers proposed an analytical framework with the following dimensions: (1) tools—the tools and devices used for language instruction and learning; (2) languages and skills—the target languages and skills that were assisted by technology; (3) theories and pedagogical approaches—the theoretical foundations of studies and the approaches that instructors used for language instruction and learning; (4) methodology—included a research method, participants, study duration and data collection; (5) applications of 360-degree video technology to assist language learning—strategies or steps to assist language learning with the technology; (6) findings— reported results in reviewed studies; (7) problems—reported problems in reviewed studies. The framework provided the basis for reviewing articles and coding their content.

After the articles were selected, the researchers analyzed the content using open coding approach. This approach allowed them to segment information and form categories of information related to the phenomena under consideration. Two researchers were involved in the coding process. They read selected articles, and highlighted and coded their content based on the analytical framework. Then, the codes were grouped into different categories and the properties of each category were identified. Finally, the researchers re-examined the reviewed articles and discussed codes and categories under question or if there were any disagreements between researchers until a consensus was reached.

## 3. Results

### 3.1 Tools

The results related to tools are summarized in Appendix 1. The results show that tools can be grouped according to the following three ways of using them: (1) creating or editing 360-degree videos/images, (2) obtaining 360-degree videos/images, and (3) viewing 360-degree videos/images. In terms of creating 360-degree videos/images, 360-degree cameras such as Insta 360 (*n* = 2), Samsung Gear 360 (*n* = 2) and LG 360 CAM (*n* = 1) were used. The results show that researchers also used Google Tour Creator (*n* = 6) and EduVenture® (*n* = 6) platforms to edit 360-degree videos/images. In terms of obtaining 360-degree videos/images, in this case, the participants in the reviewed studies obtained 360-degree videos/images that were created and shared by other people on different online platforms. For example, Google Expeditions (*n* = 4) was the most frequently used platform in the reviewed studies for obtaining 360-degree videos/images. Scholars also used Let's date (*n* = 1) and video hosting sites such as YouTube (*n* = 2) and Youku (*n* = 1) to obtain 360-degree video/image learning content. Tools used for viewing 360-degree videos/images can be divided into head-mounted displays (HMDs), mobile phones, and tablets. With respect to HMD, Google Cardboard (n=9) was the most popular tool in the reviewed studies. Other HMDs such as Samsung Gear VR (*n* = 1), iHarbort ® VR-G (*n* = 1), VR BOSS Z5 headsets (*n* = 1), Oculus VR (*n* = 1), MI VR (*n* = 1), and VIOTEK goggle (*n* = 1) were also used. Scholars did not identify HMDs used in seven studies.

### 3.2 Languages and skills

Target languages are summarized in Appendix 2. According to the results, scholars targeted English (*n* = 15) and Chinese/Mandarin (*n* = 6) more frequently. Learners also learned other languages in the reviewed articles, e.g., French (*n* = 1), German (*n* = 1) and Korean (*n* = 1). The results related to skills are reported in Appendix 3. Speaking (*n* = 9), writing (*n* = 6), vocabulary (*n* = 3), listening (*n* = 2), and reading (*n* = 1) skills were concerned by researchers in the reviewed studies. Some studies focused on one language skill only and some studies on more than one skill. Other skills such as intercultural competence (*n* = 3) and intracultural knowledge (*n* = 1) were also targeted in the reviewed studies. No specific skills were indicated in one study.

### 3.3. Theories and pedagogical approaches

Theories on which the reviewed studies were based are reported in Appendix 4. According to the results, twelve different theories were identified. Embodied cognition theory (*n* = 3), experiential learning theory (*n* = 2), situated learning theory (*n* = 2), cognitive theory of multimedia learning (*n* = 2) and the hypothetical model of immersive cognition (*n* = 2) were used in more than two studies whereas the rest theories were used only once. It should be noted that some studies mentioned more than one theory. Besides, there were ten studies in which scholars did not identify any theories. The results related to pedagogical approaches are summarized in Appendix 5. The results show that task-based learning (*n* = 19) was the most frequently used pedagogical approach. Other pedagogical approaches such as experiential learning (*n* = 2), a progressive question prompt-based peer-tutoring approach (*n* = 1), problem-based learning (*n* = 1), dyadic learning (*n* = 1), and collaborative learning (*n* = 1) were also reported in the reviewed studies. It should be noted that there were some studies in which researchers used more than one pedagogical approach. In addition, it is worth noting that there are three studies in which pedagogical approaches were not identified.

### 3.4. Methodology

The research methods used in the reviewed studies are summarized in Appendix 6. The most frequently used methods were experiment/quasi-experiment (*n* = 10) and mixed methods (*n* = 7). Other methods such as quantitative research (*n* = 3), qualitative research (*n* = 1), and action research (*n* = 1) were rarely used in the reviewed studies. In addition, there were two studies in which research methods were not identified. Appendix 7 presents the data related to participants. The data shows that scholars frequently recruited less than fifty (*n* = 15) participants. There were six studies with participants' number between 51 and 100, and two studies with more than 100 participants. One study did not provide the number of participants. The data in Appendix 8 presents academic level of the participants. The participants were college students (*n* = 17), high school students (*n* = 3), junior school students (*n* = 1), primary school students (*n* = 1), and vocational school students (*n* = 1). Participants from different academic levels in a study can be found in a few reviewed articles. Besides, there were two studies where the academic level of participants was not identified. The study duration of the reviewed studies is summarized in Appendix 9. The studies were grouped based on their duration, i.e., more than one month (*n* = 11), from one week to one month (*n* = 5), and less than one day (*n* = 1). Scholars did not indicate the duration of their studies in seven articles. The results related to data collection are summarized in Appendix 10. According to the results, researchers in the reviewed studies frequently used questionnaires/scales (*n* = 18), interviews (*n* = 14), and tests (*n* = 13). Less frequently used instruments were observation (*n* = 6), recordings (*n* = 5), and student reflection (*n* = 4). Besides, there was one study in which data collection method was not identified.

### 3.5. Applications of 360-degree video technology to assist language learning

The results related to how 360-degree video technology was applied to assist language learning are summarized in Appendix 11. The results are presented with respect to two main aspects: (a) viewing 360-degree videos/images and (b) creating 360-degree videos/images. During viewing 360-degree videos/images, strategies that students used were interacting with video content (*n* = 8), completing relevant learning tasks (*n* = 5), discussion (*n* = 4), oral presentation (*n* = 4), question and answer (*n* = 3), collaboration (*n* = 3), and peer assessment (*n* = 1). In creating 360-degree videos/images, students followed the following steps: drafting scripts (*n* = 3), evaluating and revising scripts (*n* = 2), creating or editing 360-degree video/image (*n* = 8), presenting 360-degree video/image to others (*n* = 2), viewing others' works (*n* = 3), and giving feedback (*n* = 1).

### 3.6. Findings

Findings reported in the reviewed studies are summarized in Appendix 12. Scholars frequently reported findings related to learning outcomes (*n* = 21), perceptions of using technology (*n* = 12), and motivation (*n* = 7). In addition, scholars reported their findings related to cognitive load (*n* = 3), self-efficacy (*n* = 3), language anxiety (*n* = 3), learning behaviors (*n* = 2), and thinking skills (*n* = 2). Findings related to learning outcomes included speaking performance (*n* = 7), writing performance (*n* = 5), vocabulary performance (*n* = 3), intercultural communicative competence (*n* = 2), listening performance (*n* = 1), reading performance (*n* = 1), intercultural knowledge (*n* = 1), and problem-solving abilities (*n* = 1). Learning behavior included engagement (*n* = 1) and patterns (*n* = 1). Thinking skills included creative thinking tendency (n=1) and critical thinking skills (*n* = 1). The results related to the affordances of 360-degree video technology are summarized in Appendix 13. The results can be divided into the following aspects: creating authentic context (*n* = 22), providing immersive experience (*n* = 20), facilitating language learning (*n* = 12), providing learning content/material (*n* = 11), giving a sense of presence (*n* = 10), enabling interaction with learning content (*n* = 10), increasing motivation (*n* = 5), reducing speaking anxiety (*n* = 3), improving self-efficacy (*n* = 1), and promoting creative tendency (*n* = 1).

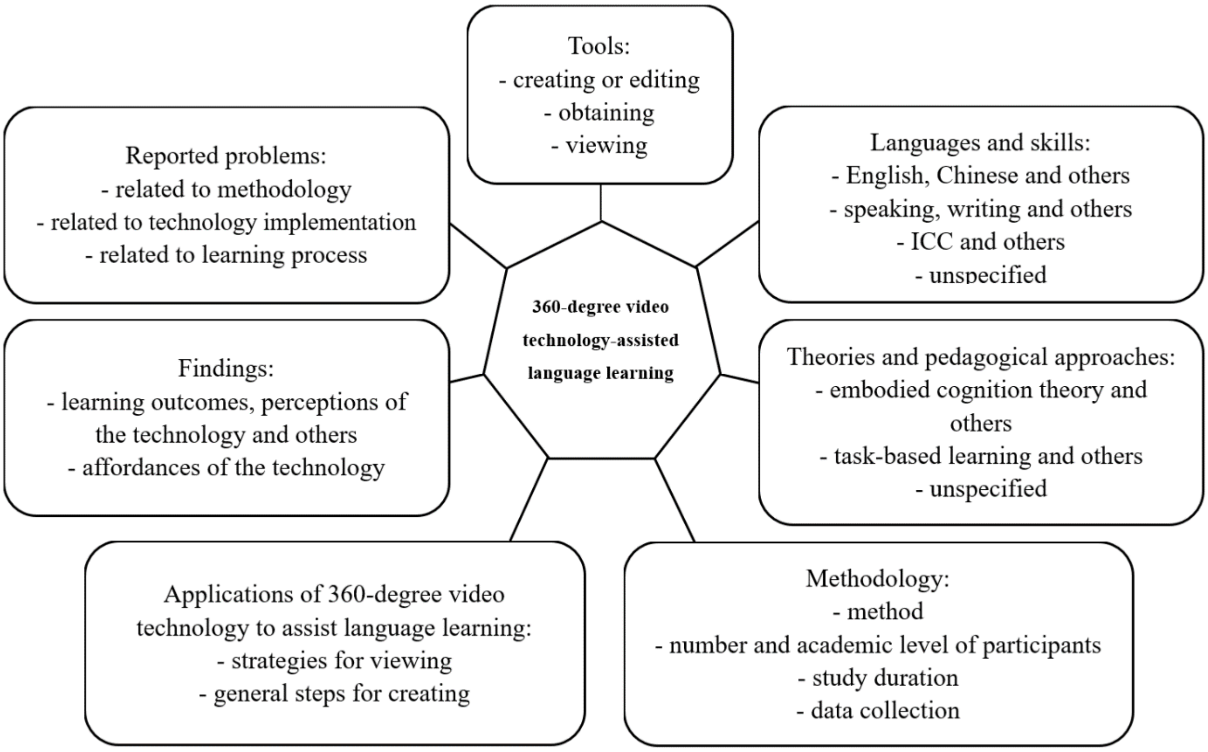
### 3.7. Reported problems

The problems identified in the reviewed studies can be grouped into the following categories: methodology, technology implementation, and learning process. The most frequently reported methodological problems (Appendix 14) were small sample size (*n* = 7), data collection strategy (*n* = 6), short period (*n* = 5) and lack of a control group (*n* = 4). In addition, there were other problems such as participants (*n* = 2), experimental design (*n* = 1) and lack of formal assessment of English proficiency (*n* = 1). In the technology implementation category (Appendix 15), the most frequent problems were physical problems (n=9) and technical difficulties (*n* = 7). Less frequently reported problems included low quality of videos (*n* = 2), length of videos (*n* = 2), small number of devices (*n* = 2), novelty effect (*n* = 2), unfamiliarity with technology (*n* = 1), missing texts (*n* = 1), unsuitable VR goggle size (*n* = 1) and increased cognitive load (*n* = 1). The problems related to the learning process (Appendix 16) included lack of adequate feedback from the instructor (*n* = 2), lack of attention to the learning status of participants (*n* = 1), lack of consideration of participants' technological competency (*n* = 1), participants' insufficient understanding of the project instruction (*n* = 1), participants' inaccurate pronunciation or unfamiliar vocabulary (*n* = 1) and distraction (*n* = 1).

## 4. Discussion

Figure 2 is an overall representation of the results of the present review study in such dimensions as (1) 360-degree video technology, (2) target languages and skills, (3) theories and pedagogical approaches, (4) research methodology, (5) applications of 360-degree video technology to assist language learning, (6) reported findings, and (7) reported problems. Figure 3 includes our suggestions for educators and researchers in the field based on the results.

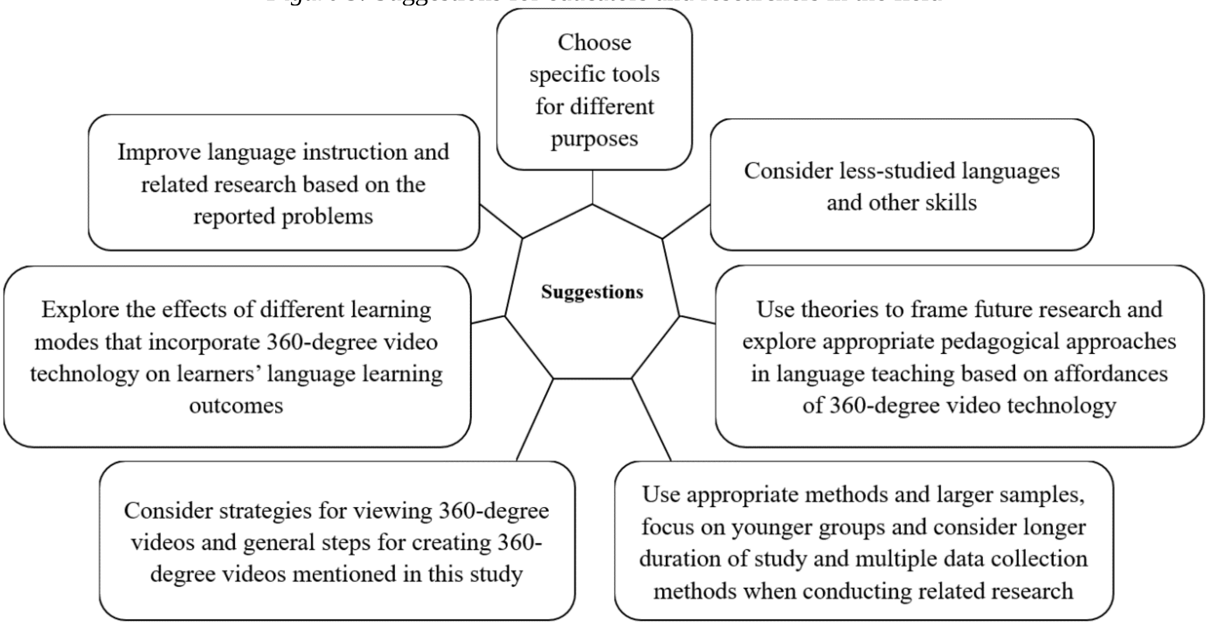
*Figure 2.* An overview of our findings




### 4.1. Tools

The findings revealed that 360-degree video tools can be divided into three categories based on their usage, i.e., creating or editing 360-degree videos/images, obtaining 360-degree videos/images, and viewing 360-degree videos/images. With respect to creating 360-degree videos, researchers or learners often used 360-degree camera because it can capture all perspectives and directions through multiple built-in lenses. Due to its simplicity of operation, not only instructors used it to develop instructional materials, but also language learners to build their own learning content. In terms of obtaining 360-degree videos/images, instructors and language learners could access and view existing 360-degree videos on various platforms (e.g., YouTube and Youku). Google Expeditions was frequently used to view 360-degree virtual tours. This platform is preloaded with thousands of scenes (Ebadi & Ebadijalal, 2020; Xie et al., 2019), so users can explore built-in sites in a 360-degree mode, such as iconic landmarks, architecture, or historical heritage. As for viewing 360-degree videos/images, language learners used HMDs to view 360-degree videos more frequently than mobile phones or tablets. The possible reason is that HMDs can provide an immersive experience for learners and they are affordable nowadays because of their low cost (e.g., Google Cardboard). Learners can rotate their heads to reorient the video when using HMD. For instance, participants in Shadiev et al. (2021a) used HMDs to watch 360-degree introduction videos created by their foreign partners to know them, their cultural background, and their school life better.

*Figure 3.* Suggestions for educators and researchers in the field




Our findings suggest that participants created their own 360-degree videos/images more frequently than just simply obtained 360-degree videos/images elsewhere. This implies that participants in more studies did not use existing content but created their own that could better fit their learning or instructional goals and needs. So, educators and researchers need to notice that such work (i.e., to create videos/images) requires substantial time and effort for planning, shooting, evaluating and revising content, and so this should be considered in the future when they plan to design language learning and teaching activities. To edit videos, the participants used Google Tour Creator or EduVenture® frequently because these platforms enabled adding interactive multimedia elements (e.g., text, images or sound) to 360-degree videos to make them more interactive and useful for language learning. However, it is worth noting that Google Expeditions & Tour Creator are no longer available to users since June 30, 2021. However, many of the 360-degree virtual tours from Expeditions can be found on Google Arts & Culture. Educators and researchers can look for more alternative platforms to help them edit 360-degree videos. One such potential platform is WondaVR and it enables adding multimedia elements such as text, image, quiz, score card, etc. As 360-degree video technologies are very popular nowadays, we believe that more editing platforms will emerge in the market in the nearest future.

### 4.2. Languages and skills

With regard to the target languages, English and Chinese were the most commonly used languages in the reviewed studies. The reason for this may be that they are popular languages and are spoken by a large number of people around the world. Besides, other languages such as French, German and Korean were also used but not so frequently. In future studies, researchers and educators may consider paying attention to lesser-involved languages and explore the potential and effectiveness of applying 360-degree video technology to support learning them.

Our results showed that language output (such as writing and speaking) received more attention. It is possible that the improvement of language output skills is more dependent on the affordances of 360-degree video technology. In traditional writing activities, students may not have a deep perception and experience of writing topics, which allows for limited depth of expression. In addition, speaking can be challenging for language learners without an authentic target language learning environment. However, 360-degree video technology can help address these limitations, e.g., create an authentic and immersive learning context, which may help learners gain a deeper perception and understanding of writing topics and provide them with a realistic communication environment. For example, Chen and Hwang (2020) adopted 360-degree video technology to provide a realistic sociocultural environment for English-speaking practice. Yang et al. (2021) proposed a system based on 360-degree video technology to set a simulated environment enabling learners to have in-depth perception in descriptive paper writing.

Language input skills also attracted the interest of researchers. In terms of vocabulary learning, 360-degree videos can support language learners to recognize new words through multimodal information. In Repetto et al. (2021), objects and verbs from 360-degree video scenes were listed as target words, and there were voice descriptions in videos to guide students to pay attention to them. With respect to reading skills, Abd Majid et al. (2020) developed reading lessons and questions based on 360-degree video materials. Students were asked to complete a set of reading comprehension activities after viewing video materials. For listening skills development, Ji et al. (2019) provided EFL learners with 360-degree videos to watch and then made them practice their listening skills.

We also found that reviewed studies focused on such abilities as intercultural communicative competence or intercultural knowledge. Perhaps, educators and researchers have considered applications of 360-degree video technology in such intercultural learning contexts because of its technological affordances. For example, in Shadiev et al. (2021b), Chinese students learned English and they were partnered with Indonesian students to practice language skills. 360-degree video technology was used by students to record learning content related to their culture and traditions in English. Students from two countries communicated with each other through exchanging created 360-degree video content. In this way, students were engaged in learning English and culture. Therefore, based on our results, it is suggested that 360-degree video technology can be used to assist language skills development. However, scholars should not limit their focus to language skills only, other abilities can be developed too, e.g., intercultural communicative competence or intercultural knowledge because they are closely related to language skills.

### 4.3. Theories and pedagogical approaches

The results showed that embodied cognition, experiential learning, situated learning, cognitive theory of multimedia learning, and the hypothetical model of immersive cognition were the most frequently used theories. Scholars based their research on these theories. According to the embodied cognition, cognitive processes are based on sensory-motor experiences (Barsalou, 2008). Repetto et al. (2021) used immersive 360-degree video as learning material where the target words were presented along with visual, auditory and motor inputs, which offered an embodied experience to learners. Experiential learning theory views learning as "the process whereby knowledge is created through the transformation of experience" (Kolb, 1984, p. 38). Scholars argued that 360-degree video technology has the potential to enable experiential learning due to its ability to create authentic and immersive experience and that students' knowledge aroused when they had the transformation of experience (Huang et al., 2020; Yang et al., 2021). According to situated learning theory (Lave & Wenger, 1991), knowledge should be learned through full participation in sociocultural practices. In Chen and Hwang (2020), learners could practice speaking in real sociocultural contexts created by the interactive virtual reality environment. For this reason, scholars based their research on the situated learning theory (Chen & Hwang, 2020; Xie et al., 2019).

It is worth noting that scholars in eleven studies did not indicate what theories they used, accounting for nearly half of all review studies. A relevant and sound theory could provide scientific guidance for the application of 360-degree video technology in language learning. It is suggested that researchers indicate theories used in their articles, as this may help readers to better understand the relationship between theoretical foundation, methodology, and results.

The findings showed that several pedagogical approaches were employed. Task-based learning was the most frequently used pedagogical approach. Completing language-related tasks in the real context created by 360-degree video technology can facilitate the development of language skills. Task-based learning emphasizes learning by doing and means that students need to use the target language to complete tasks assigned by teachers. For example, students in Xie et al. (2021) were required to act as museum guides and to introduce a famous attraction in Chinese with the assistance of Google Expeditions. Experiential learning approach was employed in two studies. This approach enables knowledge construction through four learning modes: concrete experience, reflective observation, abstract conceptualization, and active experimentation (Kolb, 1984). Huang et al. (2020)

and Yang et al. (2021) developed 360-degree video systems for descriptive paper writing based on the four stages. No pedagogical approaches were identified in three reviewed studies. It is suggested that researchers indicate such information explicitly in the future, as the pedagogical approaches can help readers understand the context and reasons for the design of learning activities.

The use of emerging technologies to assist language learning does not necessarily lead to effective learning; what is more important is that educators incorporate effective pedagogical approaches. As 360-degree video technology is an emerging technology, it may take some time to explore how its application can be integrated with pedagogical approaches. Therefore, educators can refer to the aforementioned pedagogical approaches, and adapt them to learning situations. Furthermore, in the future, researchers and educators can explore more appropriate pedagogical approaches supported by technology to facilitate the development of students' language skills.

### 4.4. Methodology

In terms of the methods employed, experiment/quasi-experiment and mixed methods were the most frequently used ones. Experiment/quasi-experiment involves control of certain factors according to research objectives and hypotheses in order to investigate correlations or cause-effect relationships among research variables. Experiment/quasi-experiments were often used in reviewed studies to examine the effects of learning intervention on language learning outcomes, learners' motivation, cognitive load, or language anxiety. For example, in Chien et al. (2020), a peer-assessment approach was proposed using 360-degree video technology and an experiment was conducted to evaluate the effectiveness of the approach on learning outcomes. The results showed that the peer-assessment approach had better effects on learning outcomes (e.g., speaking performance, learning motivation, critical thinking, and decrease of learning anxiety) compared to the non-peer-assessment approach. Mixed methods usually combine multiple data collection methods, e.g., quantitative and qualitative. Mixed methods can provide a more reliable basis for research findings and explore the reasons behind the findings. In addition, the shortcomings of each single method can be avoided. Lin and Wang (2021) adopted mixed methods research to investigate how the VR creative project might have influenced learners' creative self-efficacy and intrinsic motivation toward VR technology. Quantitative data were collected through the intrinsic motivation inventory and the creative self-efficacy student scale, while qualitative data were mainly collected via an open-ended survey. It is suggested that researchers should choose their research methods properly based on the purpose of their research.

Our results revealed that participants of different numbers and academic levels participated in the reviewed studies. Most studies were carried out with less than fifty participants. This may be due to some objective constraints (e.g., tool or participants' availability or experimental space). University students were frequently recruited in reviewed research. Perhaps this is because such participants have better experience in language learning and they are more experienced in using technology. Furthermore, relatively few studies involved participants from other academic levels, e.g., primary or junior school. This could be due to their limited experience in both language learning and technology usage. Based on our findings, we suggest that future research may consider focusing on younger age participants more. The instructor may consider designing learning activities appropriately. For example, the instructor may create 360-degree videos and provide them for young learners (instead of asking them to create their own videos or searching them online) so that they can learn through experience and observation. Also, young participants need to be instructed and constantly guided by educators or researchers to achieve better language learning outcomes.

The majority of the reviewed studies were conducted for more than one month. One possible reason is that learners need systematic training before using 360-degree video technology, especially in terms of creating and editing 360-degree videos/images. For example, in Yeh et al. (2021), the instructor taught students how to create panoramas and add interactive elements in the first several weeks. Then students watched these videos/images and so training and learning activities increased the length of the study. Another reason is that some of the studies had many language learning activities, which included, for example, both students watching 360-degree instructional videos and students making their own videos. In Chen and Hwang (2020), the participants were required to ask and answer directions with the help of a 360-degree video. After that, they were asked to plan a trip itinerary using Tour Creator, and finally give an oral presentation of their work. What's more, the development of language abilities can't be achieved in a short period of time. As a result, educators and researchers should consider longer periods for their research in order to ensure that students have adequate technical training and language practice.

We found that researchers used a variety of data collection methods. Questionnaires/scales, interviews and tests were the most commonly used ones. Questionnaires/scales are often used in studies because they can facilitate the rapid collection of large amounts of data from lots of participants. For example, Chen et al. (2021b) adopted several questionnaires to measure multiple factors, including those related to students' motivation, self-efficacy, degrees of anxiety, and cognitive load. Interviews were used to collect qualitative data that can support quantitative results subjectively. In many studies, interview data were mainly used to substantiate and explain the quantitative findings. For instance, in Xie et al. (2021), interview content was transcribed and analyzed to provide possible explanations for the quantitative data. Another frequently used data collection method was a test. Usually, tests were used to assess whether learners' performance improved in a particular aspect after the implementation of the instructional intervention, and thus, generally included pre-tests and post-tests for comparison. For example, Shadiev et al. (2021b) conducted a pre-test and a post-test of English at the beginning and end of the semester respectively to investigate the effects of learning activities supported by 360-degree video technology on students' English performance.

Language skills were mainly measured using scales or tests. For speaking, usually, learners' oral performance or speaking training was recorded and then scored by the instructor or their peers according to the speaking rating scale (Chen & Hwang, 2020; Chien et al., 2020; Ebadi & Ebadijalal, 2020). With regard to the ability to write, learners were asked to write essays on a particular topic, which were then scored by the instructor based on an essay assessment scale (Huang et al., 2020; Dolgunsöz et al., 2018; Yang et al., 2021). In addition, for vocabulary measurement, researchers adopted a bilingual translation test to check learners' level of vocabulary mastery (Repetto et al., 2021). Regarding reading skills, learners were required to take a pre-test and post-test of reading questions (Abd Majid et al., 2020). With respect to listening, a listening comprehension test was adopted to measure the learners' comprehension of the authentic material in the experiment (Ji et al., 2019).

Most studies collected data from multiple sources. The reason for this is that such an approach can make results more rigorous and robust. For instance, in Xie et al. (2019), class observations and the audio-recorded discussions between researchers were used to triangulate the data. Therefore, future studies may consider the data collection sources relevant to their research questions or hypotheses. Some data collection techniques that received little attention in reviewed studies can also be considered (i.e., observation, recordings or student reflection). In addition, physiological data can also be collected because it can reflect the objective physiological condition of learners during the language learning process. Finally, future studies may focus on multiple sources of data to make their findings and conclusions more robust.

### 4.5. Applications of 360-degree video technology to assist language learning

Our review revealed two different ways to use 360-degree video technology in language learning, i.e., students view 360-degree videos/images and students create 360-degree videos/images. Some studies had only one way to use this technology (e.g., Huang et al., 2020), while others included both (e.g., Chen & Hwang, 2020). During viewing process, interacting with video content was a very common strategy and it appeared in different forms, e.g., learners could click on interactive elements in the video/image to get more detailed information. For example, in Monteiro and Ribeiro (2020), after clicking on a text icon, the viewer saw one targeted word appear in context, together with a related image, which could facilitate students' vocabulary learning. Oral presentation strategies were also used. For instance, students in Chen et al. (2021b) were asked to introduce the museum. Question and answer strategy (i.e., a student answering questions asked by his partner) was also used. For example, a learner viewed the 360-degree video with HMD and answered questions about a destination asked by his partner (Lin et al., 2021). In Chen et al. (2021b), students took on the roles of tutors and tutees, with the tutors asking questions and the tutees answering them.

After viewing the 360-degree videos, students were asked to discuss the content of the videos and share their thoughts about it (Abd Majid et al., 2020; Chen et al., 2021a). In addition, participants were asked to complete relevant tasks, such as completing a writing assignment, answering a set of reading comprehension questions, etc., to assist students' learning and test their learning effectiveness. Furthermore, a peer assessment strategy was found in reviewed studies. In Chien et al. (2020), both groups were asked to interact in the VR environment and talk to virtual characters in English. Students' voices and viewed content were recorded. Afterwards, participants in the experimental group viewed the peers' recorded content and gave feedback about their speaking performance. Based on the results, it is suggested that when implementing technology-assisted language learning activities, educators can develop interactive 360-degree VR content as teaching resources, allowing students to learn through experience and interaction. In addition, when designing learning activities, educators can refer to the above-mentioned strategies and choose the appropriate ones for effective instruction and learning.

Another way to use 360-degree video technology is to create content in which language learners can use target languages to describe it. Some steps for creating content may include: drafting scripts, evaluating and revising scripts, creating or editing 360-degree video/image, presenting 360-degree video/image to others, viewing others' works and giving feedback. In some reviewed studies, students were asked to write narrative texts for theirs 360-degree videos in advance (drafting scripts). After that, scripts were submitted to their instructor for evaluation (evaluating scripts). Then, students revised their drafts based on the instructor's comments (revising scripts). Alternatively, students discussed their scripts in groups and then refined them based on discussion and provided feedback. After scripts were finalized, students used a 360-degree camera to shoot videos (creating content). They also used some tools to edit their own 360-degree content, and to add some interactive elements, e.g., text, image or voice (editing content). Finally, students presented their works to others (presenting content). In other cases, students watched their peers' videos and gave feedback (viewing others' works and giving feedback). For example, in Chen and Hwang (2020), students created their VR tours using their own 360-degree photos or images from Google Street View. After that, they published it in Poly and presented content to peers. For future studies, we suggest that students should be trained on how to make and edit 360-degree VR content in advance and provided with a complete production guide. This will help their production process and avoid some technical problems.

### 4.6. Findings

According to the results, learning outcomes, perceptions of using 360-degree video technology and learning motivation were the three aspects that researchers most often focused on. We found that many studies reported positive language learning outcomes. That is, learners had better performance in certain language skills or other abilities with the intervention of 360-degree video technology or certain learning mode. For example, Repetto et al. (2021) reported that students who received the training with 360-degree videos learned more words than those in the control group. In Chien et al. (2020), students, who learned with the VR-based peer-assessment approach, performed better on English speaking tests than those who learned with the VR-based non-peer-assessment approach.

With respect to perceptions, language learners generally had positive perceptions of using 360-degree video technology. The technology promoted writing immersion (Yang et al., 2021) and enabled learners to practice speaking in an authentic environment (Ebadi & Ebadijalal, 2020) and to experience the sense of "being there" (Lin et al., 2021). The real-life view provided by 360-degree video technology sparked an interest to learn learning material (Xie et al., 2019).

In addition, some studies examined learners' motivation in the context of interventions with different learning modes, and they reported different results. For instance, according to Chen and Hwang (2020), VR increased the learning motivation of the field independent learners more than those of the field dependent learners. Huang et al. (2020) found that the VR learning approach in the descriptive article writing course could not stimulate learners' intrinsic or extrinsic motivation. In Chen et al. (2021b), no significant difference in motivation between students who learned under progressive question prompt-based peer-tutoring and conventional question prompt-based peer-tutoring approaches was found.

These findings provide a reference for future related research. Future research can explore the effects of different learning modes on students' language learning performance. In addition, when considering learners' learning experiences, researchers could use not only questionnaires or scales, but also some physiological measurement instruments to obtain objective data on students' psychological states during learning.

In language learning, affordances of 360-degree video technology are defined as the "application possibilities" due to its qualities or properties. According to the reviewed studies, we found that the affordances of this technology could support language learning. 360-degree video technology can create an authentic environment for language learners to practice their language skills because it can provide scenes filmed in the real world. For example, in Chen and Hwang (2020), the instructor created the campus tour using 360-degree campus images (e.g., campus main entrance or library) from Google Street View's extensive library, allowing learners to experience real-life scenarios. Learners may feel immersed in the environment when wearing HMD. Meanwhile, this technology may provide learners with a sense of presence (Lin et al., 2021), making them feel like they are actually there. In support of language instruction, 360-degree videos can be developed as learning materials and provide learning content for learners. For instance, Repetto et al. (2021) downloaded and edited several 360-degree videos, and extracted the target words to be learned. Thus, learners could learn the target words by watching the video materials. Developers can add interactive information to 360-degree video materials, which may increase interaction between students and learning content (Chen & Hwang, 2020; Chien et al., 2020;

Huang et al., 2020; Yang et al., 2021). When learners click icons in the 360-degree VR environment, some interactive information will pop up, which enables learners to gain additional knowledge. Learners can also interact with people in the 360-degree videos (Berns et al., 2018; Chien et al., 2020; Song, 2019).

Scholars reported about positive effects of integrating 360-degree video technology on language acquisition. For example, Monteiro and Ribeiro (2020) claimed that 360-degree video technology can contribute to foreign language vocabulary learning. Scholars found that learning with 360-degree videos can enhance students' learning motivation (Abd Majid et al., 2020; Chen & Hwang, 2020; Nobrega & Rozenfeld, 2019) and reduce their speaking anxiety (Chen & Hwang, 2020; Chien et al., 2020; Xie et al., 2019). In addition, Huang et al. (2020) reported that the Chinese writing VR learning system can promote students' writing self-efficacy and creative tendency. Therefore, educators and researchers should continue to explore the potential of 360-degree video technology in language teaching and learning based on its affordances.

### 4.7. Reported problems

The reported problems are related to methodology, technology implementation, and learning process. In terms of methodology, small sample size, data collection strategy, short period, and lack of a control group were the most frequently reported ones. Scholars claimed that the number of participants in the experiment was small, which could affect the generalizability of the results (Chen et al., 2020; Ebadi & Ebadijalal, 2020; Lin & Wang, 2021). There were also problems with the data collection strategy, such as a single source of data (Chen et al., 2021a) and measurement instruments that needed to be improved (Chien et al., 2020; Lin et al., 2021). Such problems may limit the possibility of obtaining more complete experimental results. In addition, since the development of language skills requires a long time and some studies lasted short-term, there could be different findings about learning outcomes from studies that are carried out for a longer time (Chien et al., 2020; Ebadi & Ebadijalal, 2020; Yang et al., 2021). Scholars also argued that there was a lack of a control group (Chen et al., 2021a; Lin & Wang, 2021; Shadiev et al., 2021b; Xie et al., 2019) to compare the differences in learning outcomes in settings with and without the intervention.

The most frequently mentioned technology implementation-related problems were physical discomfort and technical difficulties. In the reviewed studies, scholars reported that some participants experienced physical discomfort when they watched 360-degree videos using HMD (e.g., dizziness, fatigue, and nausea) (Dolgunsöz et al., 2018; Monteiro & Ribeiro, 2020; Lin et al., 2021). Technical difficulties included the lack of internet connectivity, incompatibility of software with specific mobile phones or difficulty in editing videos (Monteiro & Ribeiro, 2020; Chien et al., 2020; DeWitt et al., 2022). These issues negatively affected participants' learning experience and outcomes.

The lack of adequate feedback from the instructor was the most mentioned problem in the learning process. Scholars claimed that the instructor did not provide feedback regularly and timely due to time constraints and big number of students (Ebadi & Ebadijalal, 2020; Xie et al., 2021). In this case, possible problems with students' learning performance were not corrected in time.

These problems should be considered and addressed in future studies. Researchers should expand the sample size, carefully define the data collection strategy, and arrange learning activities to last for longer time. To help learners overcome physical discomfort, it is also possible to expose those learners to 360-degree video before the learning activity. Using HMDs for longer time and more frequently can help them get used to the VR learning environment. Participants should be trained in advance and given guidelines for solving potential technical problems. It is also recommended that researchers and educators integrate feedback in language learning and instruction as well as encourage peer feedback and assistance.

### 4.8. Similarities and differences with related review studies

Compared with other review studies, some results of the present study are similar and some are different. For example, related studies explored tools that the instructors and students used for watching 360-degree videos (Parmaxi, 2023; Peixoto et al., 2021; Pirker & Dengel, 2021; Shadiev et al., 2021a). However, the present study explored not only tools for viewing 360-degree videos, but also tools for creating, editing and obtaining 360-degree videos. Shadiev et al. (2021a) reported that the most commonly employed theory was situated learning theory, while the present study found that embodied cognition theory was used most frequently in 360-degree video technology-assisted language learning research. In addition, no review studies explored pedagogical approaches used in reviewed studies. Thus, our study fills the gap in some missing aspects, e.g., theoretical foundation or pedagogical approaches, and extends the current knowledge of this field.

In contrast to the present study, Parmaxi (2023), Peixoto et al. (2021), Pirker and Dengel (2021) and Shadiev et al. (2021a) paid no attention to the research methods used in reviewed studies. Our results showed that the majority of participants were from university, which is in line with the findings of Shadiev et al. (2021a) and Pirker and Dengel (2021). In terms of the duration of the study, we found that most studies were conducted over a month, which is inconsistent with other review studies. For example, Shadiev et al. (2021a) reported that most studies were conducted in less than one day, and Parmaxi (2023) found that the majority of the studies employed VR for about 1-10 tasks or sessions. We found that questionnaires/scales, interviews and tests were the most commonly used data collection method, which was similar to Shadiev et al. (2021a). Parmaxi (2023), Peixoto et al. (2021), and Pirker and Dengel (2021) did not explore how the data was collected in reviewed studies. Therefore, our results can help future researchers determine research methods, numbers and academic level of participants, study duration and data collection methods. Earlier review studies (Parmaxi, 2023; Peixoto et al., 2021; Pirker & Dengel, 2021; Shadiev et al., 2021a) did not focus on applications of 360-degree video technology to assist language learning. Therefore, the present study can provide educators and researchers with some references on strategies that can be adopted in language learning activities.

Previous review studies found that most of the studies reported the potential and benefits of technology in education (i.e., using VR in language education or using 360-degree video technology in education). Our study focused on using 360-degree video technology in language learning specifically. Also, we explored the affordances of 360-degree video technology to assist language learning, which can be a reference for educators and researchers in designing technology-assisted language learning activities.

Regarding the problems, Parmaxi (2023), Pirker and Dengel (2021) and Shadiev et al. (2021a) mentioned some problems or disadvantages related to the application of the technology, whereas this study explored problems not only about technology implementation but also about the methodology of the studies and the learning process. Therefore, the problems reported in our study can help educators to avoid possible issues in language teaching and learning, and can also help future researchers to better conduct relevant research.

## 5. Conclusion

In this study, we reviewed articles on the applications of 360-degree video technology in language learning based on seven dimensions. A general agreement was found on the potential and effectiveness of using 360-degree video technology for language learning. Therefore, future researchers and educators can refer to our results when designing their language learning activities assisted by 360-degree video technology. In addition, we provide several educational tips for educators and researchers with respect to the seven dimensions covered in the present research that can be helpful in future teaching and learning practices supported by 360-degree video technology. First, the specific tools associated with the 360-degree video technology need to be determined based on the designed language learning activities. The tools can be used for language teaching and learning in such diverse ways as creating, editing, obtaining, and viewing 360-degree videos/images. Therefore, educators and researchers need to decide how their language learners are going to use them based on their teaching goals and objectives, technological capacity, users' skills to use them, and available time. In the future, studies may focus on English and Chinese as well as other languages that received little attention. Language learning activities can focus on diverse skills including language output and input. Furthermore, other closely related skills (e.g., intercultural communicative competence) can be considered too. Theories need to be used to frame future research, thus making it more scientific. As for pedagogical approaches, in addition to task-based learning, educators can make the most out of the affordances of 360-degree video technology to explore new language teaching approaches, such as experiential learning, collaborative learning or problem-based learning. Educators and researchers need to choose appropriate methods based on their research purpose, try to collect multiple data, adopt a larger sample size, and consider their study lasts longer to make it more robust. As 360-degree video technology is an emerging technology, researchers and educators can consider the above strategies and steps in the process of applying it to assist language learning. Future research can explore the effects of different learning modes that incorporate 360-degree video technology on learners' language learning outcomes. Finally, the problems reported by scholars deserve attention in the design of future language learning activities. That is, educators and researchers may get acquainted with reported problems before they design their language learning activities supported by 360-degree video technology. For example, they may first try to understand what are potential problems and how to avoid or address them more efficiently.

The research field on the use of 360-degree video technology in language learning is still emerging. For this reason, we were able to find only twenty-four articles that met our inclusion criteria. This is one important sign for educators and researchers to actively do research in the field to enrich it with theoretical, technological, and pedagogical knowledge. On the other hand, even though our study reviewed only twenty-four articles, it offers current knowledge of the field which may guide educators and researchers in the design and implementation of their future studies.

Three limitations need to be acknowledged regarding our review study: (1) Technologies are emerging and developing very fast, and at the same time, some technologies are outdating or are no longer supported. One example is the Google Expeditions platform which was the most frequently used technology in the reviewed studies. However, it is no longer supported by Google company since 2021. Perhaps, some other tools that are listed in the present review study will be discontinued or outdated in a few years. Some technologies that were listed in the present review study can be robust and affordable enough to last much longer. On the other hand, some new tools may also emerge very quickly and they are not included in this review study. (2) It is possible that using other data sources and inclusion/exclusion criteria for selecting related studies would yield a different number of articles. (3) Finally, some reviewed articles did not include important information related to the aspects covered in the present study, e.g., tools used, theories, pedagogical approaches or methodological details. Therefore, we were unable to collect and report all necessary information. Educators and researchers need to consider these issues in their future research. Furthermore, educators and researchers need to explore the use of 360-degree video technology in other domains of knowledge as there is a huge gap in the literature.

## References


*Abd Majid, S. N., Ismail, R., & Bakar, A. H. A. (2020). Learners' perspective of using 360-degree video in reading classroom. *Malaysian Journal of Social Sciences and Humanities, 5*(10), 238-246. https://doi.org/10.47405/mjssh.v5i10.501

Barsalou, L. W. (2008). Grounded cognition. *Annual review of psychology, 59*(1), 617-645. https://doi.org/10.1146/annurev.psych.59.103006.093639

*Berns, A., Mota, J. M., Ruiz-Rube, I., & Dodero, J. M. (2018). Exploring the potential of a 360° video application for foreign language learning. In *Proceedings of the sixth international conference on technological ecosystems for enhancing multiculturality* (pp. 776-780). https://doi.org/10.1145/3284179.3284309

Caseiro, N., & Santos, D. (Eds.). (2018). *Smart specialization strategies and the role of entrepreneurial universities*. IGI Global. https://www.igi-global.com/book/smart-specialization-strategies-role-entrepreneurial/197442 https://doi.org/10.4018/978-1-5225-6152-1

*Chen, M., Chai, C. S., Jong, M. S. Y., & Chao, G. C. N. (2021a). Modeling learners' self-concept in Chinese descriptive writing based on the affordances of a virtual reality-supported environment. *Education and Information Technologies, 26*(5), 6013-6032. https://doi.org/10.1007/s10639-021-10582-4

*Chen, C. Y., Chang, S. C., Hwang, G. J., & Zou, D. (2021b). Facilitating EFL learners' active behaviors in speaking: A Progressive question prompt-based peer-tutoring approach with VR contexts. *Interactive Learning Environments*, 1-20. https://doi.org/10.1080/10494820.2021.1878232

*Chen, C. H., Hung, H. T., & Yeh, H. C. (2021c). Virtual reality in problem-based learning contexts: Effects on the problem-solving performance, vocabulary acquisition and motivation of English language learners. *Journal of Computer Assisted Learning, 37*(3), 851-860. https://doi.org/10.1111/jcal.12528

*Chen, M. R. A., & Hwang, G. J. (2020). Effects of experiencing authentic contexts on English speaking performances, anxiety and motivation of EFL students with different cognitive styles. *Interactive Learning Environments*, 1-21. https://doi.org/10.1080/10494820.2020.1734626

*Chien, S. Y., Hwang, G. J., & Jong, M. S. Y. (2020). Effects of peer assessment within the context of spherical video-based virtual reality on EFL students' English-Speaking performance and learning perceptions. *Computers & Education, 146,* 103751. https://doi.org/10.1016/j.compedu.2019.103751

Concannon, B. J., Esmail, S., & Roberts, M. R. (2019). Head-mounted display virtual reality in post-secondary education and skill training. *Frontiers in Education*, *4*(80), 1–23. https://doi.org/10.3389/feduc.2019.00080

Dhimolea, T. K., Kaplan-Rakowski, R., & Lin, L. (2022). A Systematic review of research on high-immersion virtual reality for language learning. *TechTrends, 66*(5), 810-824. https://doi.org/10.1007/s11528-022-00717-w

*DeWitt, D., Chan, S. F., & Loban, R. (2022). Virtual reality for developing intercultural communication competence in Mandarin as a Foreign language. *Educational Technology Research and Development*, *70*, 615–638. https://doi.org/10.1007/s11423-021-10074-9

[*]Dolgunsöz, E., Yildirim, G., & Yildirim, S. (2018). The Effect of virtual reality on EFL writing performance. *Journal of Language and Linguistic Studies, 14*(1), 278-292.

[*]Ebadi, S., & Ebadijalal, M. (2020). The Effect of Google Expeditions virtual reality on EFL learners' willingness to communicate and oral proficiency. *Computer Assisted Language Learning,* 1-25. https://doi.org/10.1080/09588221.2020.1854311

Fukuta, J., Gill, N., Rooney, R., Coombs, A., & Murphy, D. (2021). Use of 360 video for a virtual operating theatre orientation for medical students. *Journal of Surgical Education*, *78*(2), 391-393. https://doi.org/10.1016/j.jsurg.2020.08.014

Geng, J., Chai, C.-S., Jong, M. S.-Y., & Luk, E. T.-H. (2021). Understanding the pedagogical potential of interactive spherical video-based virtual reality from the teachers' perspective through the ACE framework. *Interactive Learning Environments, 29*(4), 618-633. https://doi.org/10.1080/10494820.2019.1593200

Herault, R. C., Lincke, A., Milrad, M., Forsgärde, E.-S., & Elmqvist, C. (2018). Using 360-degrees interactive videos in patient trauma treatment education: Design, development and evaluation aspects. *Smart Learning Environments*, *5*(1), 1–15. https://doi.org/10.1186/s40561-018-0074-x

[*]Huang, H. L., Hwang, G. J., & Chang, C. Y. (2020). Learning to be a writer: A Spherical video-based virtual reality approach to supporting descriptive article writing in high school Chinese courses. *British Journal of Educational Technology, 51*(4), 1386-1405. https://doi.org/10.1111/bjet.12893

[*]Ji, S., Li, K., & Zou, L. (2019). The Effect of 360-degree video authentic materials on EFL learners' listening comprehension. In *2019 International Joint Conference on Information, Media and Engineering (IJCIME)* (pp. 288-293). IEEE. https://doi.org/10.1109/IJCIME49369.2019.00065

Kim, H., So, H. J., & Park, J. Y. (2022). Examining the effect of socially engaged art education with virtual reality on creative problem solving. *Educational Technology & Society, 25*(2), 117-129.

Kittel, A., Larkin, P., Elsworthy, N., & Spittle, M. (2019). Using 360° virtual reality as a decision making assessment tool in sport. *Journal of Science and Medicine in Sport*, *22*(9), 1049–1053. https://doi.org/10.1016/j.jsams.2019.03.012

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development.* Prentice-Hall.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge university press. https://doi.org/10.1017/CBO9780511815355

Li, H., Chan, G., & Skitmore, M. (2012). Multiuser virtual safety training system for tower crane dismantlement. *Journal of Computing in Civil Engineering*, *26*(5), 638–647. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000170

Liu, Q., Homma, R., & Iki, K. (2020). Evaluating cyclists' perception of satisfaction using 360° videos. *Transportation Research Part A: Policy and Practice, 132*, 205–213. http://doi: 10.1016/j.tra.2019.11.008

[*]Lin, V., Barrett, N. E., Liu, G. Z., Chen, N. S., & Jong, M. S. Y. (2021). Supporting dyadic learning of English for tourism purposes with scenery-based virtual reality. *Computer Assisted Language Learning,* 1-37. https://doi.org/10.1080/09588221.2021.1954663

[*]Lin, Y. J., & Wang, H. C. (2021). Using virtual reality to facilitate learners' creative self-efficacy and intrinsic motivation in an EFL classroom. *Education and Information Technologies, 26*(4), 4487-4505. https://doi.org/10.1007/s10639-021-10472-9

[*]Monteiro, A. M. V., & Ribeiro, P. N. D. S. (2020). Virtual reality in English vocabulary teaching: An Exploratory study on affect in the use of technology. *Trabalhos em Linguística Aplicada, 59*, 1310-1338. https://doi.org/10.1590/010318137569316200716

[*]Nobrega, F. A., & Rozenfeld, C. C. D. F. (2019). Virtual reality in the teaching of FLE in a Brazilian public school. *Languages, 4*(2), 36. https://doi.org/10.3390/languages4020036

Parmaxi, A. (2023). Virtual reality in language learning: A Systematic review and implications for research and practice. *Interactive learning environments, 31*(1), 172-184. https://doi.org/10.1080/10494820.2020.1765392

Peixoto, B., Pinto, R., Melo, M., Cabral, L., & Bessa, M. (2021). Immersive virtual reality for foreign language education: A PRISMA systematic review. *IEEE Access, 9*, 48952-48962. https://doi.org/10.1109/ACCESS.2021.3068858

Pirker, J., & Dengel, A. (2021). The Potential of 360° virtual reality videos and real VR for education—A Literature review. *IEEE Computer Graphics and Applications, 41*(4), 76-89. https://doi.org/10.1109/MCG.2021.3067999

[*]Repetto, C., Di Natale, A. F., Villani, D., Triberti, S., Germagnoli, S., & Riva, G. (2021). The Use of immersive 360 videos for foreign language learning: A Study on usage and efficacy among high-school students. *Interactive Learning Environments*, 1-16. https://doi.org/10.1080/10494820.2020.1863234

Rupp, M. A., Odette, K. L., Kozachuk, J., Michaelis, J. R., Smither, J. A., & McConnell, D. S. (2019). Investigating learning outcomes and subjective experiences in 360-degree videos. *Computers & Education, 128*, 256-268. https://doi.org/10.1016/j.compedu.2018.09.015

*Smith, H., & Townsend, S. D. (2021). Investigating the efficacy of utilizing 360° camera technology as a language teaching and learning tool in a science content based EFL classroom. *Innovation in Language Learning and Teaching, 15*(2), 143-155. https://doi.org/10.1080/17501229.2019.1710515

Snelson, C., & Hsu, Y. C. (2020). Educational 360-degree videos in virtual reality: A Scoping review of the emerging research. *TechTrends, 64*(3), 404-412. https://doi.org/10.1007/s11528-019-00474-3

*Song, J. (2019). Enhancing intercultural competence with 360-degree virtual reality videos. *The Korean Language in America, 23*(1), 85-98. https://doi.org/10.5325/korelangamer.23.1.0085

Shadiev, R., & Feng, Y. (2023). Using automated corrective feedback tools in language learning: A Review study. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2022.2153145

Shadiev, R. & Li, D. (2023). A Review study on eye-tracking technology usage in immersive virtual reality learning environments. *Computers & Education, 196*, 104681. https://doi.org/10.1016/j.compedu.2022.104681

Shadiev, R., & Wang, X. (2022). A Review of research on technology-supported language learning and 21st century skills. *Frontiers in Psychology, 13*, 897689. https://doi.org/10.3389/fpsyg.2022.897689

Shadiev, R., Yang, L., & Huang, Y. M. (2021a). A Review of research on 360-degree video and its applications to education. *Journal of Research on Technology in Education*, 1-16. https://doi.org/10.1080/15391523.2021.1928572

*Shadiev, R., Yu, J., & Sintawati, W. (2021b). Exploring the impact of learning activities supported by 360-degree video technology on language learning, intercultural communicative competence development, and knowledge sharing. *Frontiers in Psychology, 12*, 766924. https://doi.org/10.3389/fpsyg.2021.766924

Shadiev, R., & Yu, J. T. (2023). Review of research on computer-assisted language learning with a focus on intercultural education. *Computer Assisted Language Learning*. https://doi.org/10.1080/09588221.2022.2056616

Wu, J., Guo, R., Wang, Z., & Zeng, R. (2021). Integrating spherical video-based virtual reality into elementary school students' scientific inquiry instruction: Effects on their problem-solving performance. *Interactive Learning Environments, 29*(3), 496-509. https://doi.org/10.1080/10494820.2019.1587469

*Xie, Y., Chen, Y., & Ryder, L. H. (2021). Effects of using mobile-based virtual reality on Chinese L2 students' oral proficiency. *Computer Assisted Language Learning, 34*(3), 225-245. https://doi.org/10.1080/09588221.2019.1604551

*Xie, Y., Ryder, L., & Chen, Y. (2019). Using interactive virtual reality tools in an advanced Chinese language class: A Case study. *TechTrends, 63*(3), 251-259. https://doi.org/10.1007/s11528-019-00389-z

*Yang, G., Chen, Y. T., Zheng, X. L., & Hwang, G. J. (2021). From experiencing to expressing: A Virtual reality approach to facilitating pupils' descriptive paper writing performance and learning behavior engagement. *British Journal of Educational Technology, 52*(2), 807-823. https://doi.org/10.1111/bjet.13056

Ye, X., Liu, P. F., Lee, X. Z., Zhang, Y. Q., & Chiu, C. K. (2021). Classroom misbehaviour management: An SVVR-based training system for preservice teachers. *Interactive Learning Environments, 29*(1), 112-129. https://doi.org/10.1080/10494820.2019.1579235

*Yeh, H. C., Tseng, S. S., & Heng, L. (2021). Enhancing EFL students' intracultural learning through virtual reality. *Interactive Learning Environments*, 1-10. https://doi.org/10.1080/10494820.2020.1734625

Zulkiewicz, B. A., Boudewyns, V., Gupta, C., Kirschenbaum, A., & Lewis, M. A. (2020). Using 360-degree video as a research stimulus in digital health studies: Lessons learned. *JMIR Serious Games, 8*(1), e15422. https://doi.org/10.2196/15422

# Appendix

*Appendix 1.* Tools

| Category | Tool | Frequency | Reference |
|---|---|---|---|
| ***I. Creating or editing 360-degree videos/images with:*** | | | |
| 360-degree Camera | Insta 360 | 2 | Shadiev et al. (2021b), Song (2019) |
| | Samsung Gear 360 | 2 | Ji et al. (2019), Song (2019) |
| | LG 360 CAM | 1 | Smith & Townsend (2021) |
| | Unspecified model | 1 | Chien et al. (2020) |
| Google Tour Creator | - | 6 | Chen et al. (2021c), Chen & Hwang (2020), DeWitt et al. (2022), Huang et al. (2020), Lin & Wang (2021), Nobrega & Rozenfeld (2019) |
| EduVenture® | - | 6 | Chen et al. (2021b), Chen & Hwang (2020), Chien et al. (2020), Lin et al. (2021), Yang et al. (2021), Yeh et al. (2021) |
| ***II. Obtaining 360-degree videos/images from other sources such as:*** | | | |
| | Google Expeditions | 4 | Ebadi & Ebadijalal (2020), Monteiro & Ribeiro (2020), Xie et al. (2021), Xie et al. (2019) |
| | YouTube | 2 | Abd Majid et al. (2020), Dolgunsöz et al. (2018) |
| | Youku | 1 | Yang et al. (2021) |
| | Let's date! | 1 | Berns et al. (2018) |
| | Unspecified source | 1 | Repetto et al. (2021) |
| ***III. Viewing 360-degree videos/images with:*** | | | |
| Head-Mounted Display | Google Cardboard | 9 | Abd Majid et al. (2020), Chen et al. (2021c), Chen & Hwang (2020), Huang et al. (2020), Lin & Wang (2021), Monteiro & Ribeiro (2020), Song (2019), Xie et al. (2021), Xie et al. (2019) |
| | Samsung Gear VR | 1 | Dolgunsöz et al. (2018) |
| | iHarbort ® VR-G | 1 | Repetto et al. (2021) |
| | VR BOSS Z5 headsets | 1 | Nobrega & Rozenfeld (2019) |
| | Oculus VR | 1 | Shadiev et al. (2021b) |
| | MI VR | 1 | Shadiev et al. (2021b) |
| | VIOTEK goggle | 1 | Song (2019) |
| | Unspecified model | 7 | Berns et al. (2018), Chen et al. (2021a), Chen et al. (2021b), DeWitt et al. (2022), Lin et al. (2021), Smith & Townsend (2021), Yang et al. (2021) |
| Mobile phone | - | 16 | Abd Majid et al. (2020), Berns et al. (2018), Chen & Hwang (2020), Chien et al. (2020), DeWitt et al. (2022), Dolgunsöz et al. (2018), Ebadi & Ebadijalal (2020), Ji et al. (2019), Lin et al. (2021), Lin & Wang (2021), Monteiro & Ribeiro (2020), Repetto et al. (2021), Smith & Townsend (2021), Xie et al. (2021), Xie et al. (2019), Yeh et al. (2021) |
| Tablet | - | 2 | Chien et al. (2020), Xie et al. (2019) |

*Appendix 2.* Languages

| Language | Frequency | Reference |
|---|---|---|
| English | 15 | Abd Majid et al. (2020), Chen et al. (2021b), Chen et al. (2021c), Chen & Hwang (2020), Chien et al. (2020), Dolgunsöz et al. (2018), Ebadi & Ebadijalal (2020), Ji et al. (2019), Lin et al. (2021), Lin & Wang (2021), Monteiro & Ribeiro (2020), Repetto et al. (2021), Shadiev et al. (2021b), Smith & Townsend (2021), Yeh et al. (2021) |
| Chinese/Mandarin | 6 | Chen et al. (2021a), DeWitt et al. (2022), Huang et al. (2020), Xie et al. (2021), Xie et al. (2019), Yang et al. (2021) |
| French | 1 | Nobrega & Rozenfeld (2019) |
| German | 1 | Berns et al. (2018) |
| Korean | 1 | Song (2019） |

*Appendix 3.* Skills

| Skill | Frequency | Reference |
|---|---|---|
| ***I. Language skills such as:*** | | |
| Speaking | 9 | Berns et al. (2018), Chen et al. (2021b), Chen & Hwang (2020), Chien et al. (2020), Ebadi & Ebadijalal (2020), Lin et al. (2021), Nobrega & Rozenfeld (2019), Xie et al. (2021), Xie et al. (2019) |
| Writing | 6 | Chen et al. (2021a), Dolgunsöz et al. (2018), Huang et al. (2020), Lin et al. (2021), Nobrega & Rozenfeld (2019), Yang et al. (2021) |
| Vocabulary | 3 | Chen et al. (2021c), Monteiro & Ribeiro (2020), Repetto et al. (2021) |
| Listening | 2 | Berns et al. (2018), Ji et al. (2019) |
| Reading | 1 | Abd Majid et al. (2020) |
| ***II. Other skills:*** | | |
| Intercultural Competence | 3 | DeWitt et al. (2022), Shadiev et al. (2021b), Song (2019) |
| Intracultural knowledge | 1 | Yeh et al. (2021) |
| ***III. Unspecified*** | | |
| - | 1 | Smith & Townsend (2021) |

*Appendix 4.* Theories

| Theory | Frequency | Reference |
|---|---|---|
| Embodied cognition theory | 3 | Chen et al. (2021c), Repetto et al. (2021), Xie et al. (2019) |
| Experiential learning theory | 2 | Huang et al. (2020), Yang et al. (2021) |
| Situated learning theory | 2 | Chen & Hwang (2020), Xie et al. (2019) |
| Cognitive Theory of Multimedia Learning | 2 | Lin et al. (2021), Monteiro & Ribeiro (2020) |
| The Hypothetical Model of Immersive Cognition | 2 | Xie et al. (2021), Xie et al. (2019) |
| The Dual Coding Theory | 1 | Lin et al. (2021) |
| Cognitive load theory | 1 | Ji et al. (2019) |
| Social learning theory | 1 | Chien et al. (2020) |
| Sociocultural theory | 1 | Lin et al. (2021) |
| Constructivist learning principle | 1 | Berns et al. (2018) |
| Engagement theory | 1 | Ebadi & Ebadijalal (2020) |
| Self-determination theory | 1 | Chen et al. (2021a) |
| Unspecified | 10 | Abd Majid et al. (2020), Chen et al. (2021b), DeWitt et al. (2022), Dolgunsöz et al. (2018), Lin & Wang (2021), Nobrega & Rozenfeld (2019), Shadiev et al. (2021b), Smith & Townsend (2021), Song (2019), Yeh et al. (2021) |

*Appendix 5.* Pedagogical approaches

| Pedagogical approach | Frequency | Reference |
|---|---|---|
| Task-based learning | 19 | Abd Majid et al. (2020), Chen et al. (2021a), Chen et al. (2021b), Chen et al. (2021c), Chen & Hwang (2020), Chien et al. (2020), DeWitt et al. (2022), Dolgunsöz et al. (2018), Ebadi & Ebadijalal (2020), Huang et al. (2020), Lin et al. (2021), Lin & Wang (2021), Nobrega & Rozenfeld (2019), Shadiev et al. (2021b), Song (2019), Xie et al. (2021), Xie et al. (2019), Yang et al. (2021), Yeh et al. (2021) |
| Experiential learning approach | 2 | Huang et al. (2020), Yang et al. (2021) |
| PQP-PTVR learning (a progressive question prompt-based peer tutoring approach to VR-enhanced learning) | 1 | Chen et al. (2021b) |
| Problem-based learning | 1 | Chen et al. (2021c) |
| Dyadic learning | 1 | Lin et al. (2021) |
| Collaborative learning | 1 | Smith & Townsend (2021) |
| Implicit and explicit strategies | 1 | Monteiro & Ribeiro (2020) |
| Unspecified | 3 | Berns et al. (2018), Ji et al. (2019), Repetto et al. (2021) |

*Appendix 6*. Method

| Method | Frequency | Reference |
|---|---|---|
| Experiment/ Quasi-experiment | 10 | Chen et al. (2021c), Chen et al. (2021b), Chen & Hwang (2020), Chien et al. (2020), DeWitt et al. (2022), Huang et al. (2020), Ji et al. (2019), Lin et al. (2021), Repetto et al. (2021), Yang et al. (2021) |
| Mixed-method | 7 | Abd Majid et al. (2020), Dolgunsöz et al. (2018), Ebadi & Ebadijalal (2020), Lin & Wang (2021), Monteiro & Ribeiro (2020), Shadiev et al. (2021b), Xie et al. (2021) |
| Quantitative research | 3 | Berns et al. (2018), Chen et al. (2021a), Yeh et al. (2021) |
| Qualitative research | 1 | Xie et al. (2019) |
| Action research | 1 | Nobrega & Rozenfeld (2019) |
| Unspecified | 2 | Smith & Townsend (2021), Song (2019) |

*Appendix 7*. Number of participants

| Number | Frequency | Reference |
|---|---|---|
| 1-50 | 15 | Abd Majid et al. (2020), Berns et al. (2018), Chen et al. (2021b), DeWitt et al. (2022), Dolgunsöz et al. (2018), Ebadi & Ebadijalal (2020), Lin & Wang (2022), Lin et al. (2021), Monteiro & Ribeiro (2020), Nobrega & Rozenfeld (2019), Shadiev et al. (2021b), Song (2019), Xie et al. (2021), Xie et al. (2019), Yang et al. (2021) |
| 51-100 | 6 | Chen et al. (2021c), Chen & Hwang (2020), Chien et al. (2020), Huang et al. (2020), Ji et al. (2019), Yeh et al. (2021) |
| > 100 | 2 | Chen et al. (2021a), Repetto et al. (2021) |
| Unspecified | 1 | Smith & Townsend (2021) |

*Appendix 8*. Academic level of the participants

| Academic Level | Frequency | Reference |
|---|---|---|
| College students | 17 | Abd Majid et al. (2020), Berns et al. (2018), Chen et al. (2021c), Chen et al. (2021b), Chen & Hwang (2020), DeWitt et al. (2022), Dolgunsöz et al. (2018), Ji et al. (2019), Lin & Wang (2021), Lin et al. (2021), Monteiro & Ribeiro (2020), Shadiev et al. (2021b), Smith & Townsend (2021), Song (2019), Xie et al. (2021), Xie et al. (2019), Yeh et al. (2021) |
| High school students | 3 | Chien et al. (2020), Huang et al. (2020), Repetto et al. (2021) |
| Junior school students | 1 | Chen et al. (2021a) |
| Primary school students | 1 | Yang et al. (2021) |
| Vocational school students | 1 | Shadiev et al. (2021b) |
| Unspecified | 2 | Ebadi & Ebadijalal (2020), Nobrega & Rozenfeld (2019) |

*Appendix 9*. Study duration

| Duration | Frequency | Reference |
|---|---|---|
| >1 month | 11 | Chen & Hwang (2020), Dolgunsöz et al. (2018), Ebadi & Ebadijalal (2020), Lin & Wang (2021), Lin et al. (2021), Nobrega & Rozenfeld (2019), Shadiev et al. (2021b), Smith & Townsend (2021), Xie et al. (2021), Xie et al. (2019), Yeh et al. (2021) |
| 1 week -1 month | 5 | Chien et al. (2020), DeWitt et al. (2022), Huang et al. (2020), Repetto et al. (2021), Yang et al. (2021) |
| < 1 day | 1 | Chen et al. (2021b) |
| Unspecified | 7 | Abd Majid et al. (2020), Berns et al. (2018), Chen et al. (2021a), Chen et al. (2021c), Ji et al. (2019), Monteiro & Ribeiro (2020), Song (2019) |

*Appendix 10.* Data collection

| Evaluation | Frequency | Reference |
|---|---|---|
| Questionnaire/Scale | 18 | Berns et al. (2018), Chen et al. (2021a), Chen et al. (2021b), Chen et al. (2021c), Chen & Hwang (2020), Chien et al. (2020), DeWitt et al. (2022), Ebadi & Ebadijalal (2020), Huang et al. (2020), Ji et al. (2019), Lin et al. (2021), Lin & Wang (2021), Monteiro & Ribeiro (2020), Nobrega & Rozenfeld (2019), Repetto et al. (2021), Shadiev et al. (2021b), Yang et al. (2021), Yeh et al. (2021) |
| Interview | 14 | Abd Majid et al. (2020), Chen et al. (2021b), Chen et al. (2021c), Chien et al. (2020), DeWitt et al. (2022), Dolgunsöz et al. (2018), Ebadi & Ebadijalal (2020), Huang et al. (2020), Nobrega & Rozenfeld (2019), Shadiev et al. (2021b), Song (2019), Xie et al. (2021), Xie et al. (2019), Yang et al. (2021) |
| Test | 13 | Abd Majid et al. (2020), Chen et al. (2021b), Chen et al. (2021c), Chen & Hwang (2020), Chien et al. (2020), Dolgunsöz et al. (2018), Huang et al. (2020), Ji et al. (2019), Lin et al. (2021), Monteiro & Ribeiro (2020), Repetto et al. (2021), Shadiev et al. (2021b), Yang et al. (2021) |
| Observation | 6 | Ebadi & Ebadijalal (2020), Monteiro & Ribeiro (2020), Nobrega & Rozenfeld (2019), Shadiev et al. (2021b), Xie et al. (2021), Xie et al. (2019) |
| Recordings (audio or video) | 5 | Chen & Hwang (2020), Ebadi & Ebadijalal (2020), Nobrega & Rozenfeld (2019), Shadiev et al. (2021b), Xie et al. (2021) |
| Student reflection | 4 | Lin et al. (2021), Song (2019), Xie et al. (2021), Xie et al. (2019) |
| Unspecified | 1 | Smith & Townsend (2021) |

*Appendix 11.* Applications of 360-degree video technology to assist language learning

| Category | Example | Frequency | Reference |
|---|---|---|---|
| ***I. Strategies for viewing 360-degree videos or images*** | | | |
| Interact with video content | Interact with the pop-ups in VR environment by pressing the action button on HMD, click the record button to answer the questions presented by the system, interact with foreign teachers and local students by replying to people | 8 | Berns et al. (2018), Chen & Hwang (2020), Chien et al. (2020), Huang et al. (2020), Monterio & Ribeiro (2020), Nobrega & Rozenfeld (2019), Song (2019), Yang et al. (2021) |
| Complete relevant exercise tasks | Complete the related exercise tasks, e.g., complete a writing task or answer a set of reading comprehension after viewing | 5 | Abd Majid et al. (2020), Chen et al. (2021a), Dolgunsöz et al. (2018), Huang et al. (2020), Yang et al. (2021) |
| Discussion | Discuss the observations after viewing | 4 | Abd Majid et al. (2020), Chen et al (2021c), Shadiev et al. (2021b), Song (2019) |
| Oral presentation | Introduce a specific location in target language with the assistance of 360-degree virtual tour, e.g., museum, a famous attraction | 4 | Chen et al. (2021b), Ebadi & Ebadijalal (2020), Xie et al. (2021), Xie et al. (2019) |
| Question and answer | Participants who were watching the 360-degree video practiced giving destination information as they answered detailed questions posed by their partner | 3 | Chen et al. (2021b), Chen & Hwang (2020), Lin et al. (2021) |
| Collaboration | Collaborate in pairs to complete learning tasks | 3 | Chen et al. (2021b), Chen & Hwang (2020), Lin et al. (2021) |
| Peer assessment | Watch peers’ films and conduct peer assessment | 1 | Chien et al. (2020) |
| ***II. General steps for students to create 360-degree videos or images*** | | | |
| Draft scripts | Draft oral scripts | 3 | Lin & Wang (2021), Shadiev et al. (2021b), Smith & Townsend (2021) |
| Evaluate and revise scripts | Submit the drafts to the instructor for editing and evaluation, discuss scripts’ content in groups | 2 | Lin & Wang (2021), Shadiev et al. (2021b) |

| | | | |
|---|---|---|---|
| Create or edit 360-degree video/image | Use 360-degree camera to shoot 360-degree video, use Google Tour Creator or EduVenture to edit virtual tour | 8 | Chen et al. (2021c), Chen & Hwang (2020), DeWitt et al. (2022), Lin & Wang (2021), Nobrega & Rozenfeld (2019), Shadiev et al. (2021b), Smith & Townsend (2021), Yeh et al. (2021) |
| Present 360-degree video/image to others | Give an oral presentation | 2 | Chen & Hwang (2020), Nobrega & Rozenfeld (2019) |
| View others' works | Watch peers' 360-degree VR content, watch partner's video | 3 | Lin & Wang (2021), Shadiev et al. (2021b), Yeh et al. (2021) |
| Give feedback | Evaluate peers' videos | 1 | Yeh et al. (2021) |

*Appendix 12.* Findings

| Category | Code | Example | Frequency | Reference |
|---|---|---|---|---|
| Learning outcomes | Speaking performance | The content and vocabulary of participants' oral presentations when using VR tools scored significantly higher than when not using VR tools. | 7 | Chen et al. (2021b), Chen & Hwang (2020), Chien et al. (2020), Ebadi & Ebadijalal (2020), Lin et al. (2021), Nobrega & Rozenfeld (2019), Xie et al. (2021) |
| | Writing performance | The VR system could promote students' writing performance for content and appearance. | 5 | Dolgunsöz et al. (2018), Huang et al. (2020), Lin et al. (2021), Nobrega & Rozenfeld (2019), Yang et al. (2021) |
| | Vocabulary performance | Students who underwent the training with 360-degree videos learned more words. | 3 | Chen et al. (2021c), Monteiro & Ribeiro (2020), Repetto et al. (2021) |
| | Intercultural communicative competence | 360-degree video technology -supported intercultural learning activities improved students' intercultural communicative competence. | 2 | DeWitt et al. (2022), Shadiev et al. (2021b) |
| | Listening performance | The learners who watched 360-degree video journalism did worse in English listening comprehension test. | 1 | Ji et al. (2019) |
| | Reading performance | The use of 360-degree video in reading lesson help students understand the reading text better. | 1 | Abd Majid et al. (2020) |
| | Intracultural knowledge | Students developed better intracultural awareness through the features of VR technology including panorama, audio, interaction, and structuring. | 1 | Yeh et al. (2021) |
| | Problem-solving performance | Exposing students to PBL contexts can develop students' problem-solving performance. | 1 | Chen et al. (2021c) |
| Perceptions of using 360-degree video technology | - | The real-life view VR tools offered an authentic context for Chinese language learning, sparked interest in the virtually presented locales, and encouraged students to further explore the target culture. | 12 | Abd Majid et al. (2020), Berns et al. (2018), Dolgunsöz et al. (2018), Ebadi & Ebadijalal (2020), Huang et al. (2020), Lin et al. (2021), Lin & Wang (2021), Monteiro & |

| | | | | |
|---|---|---|---|---|
| | | | | Ribeiro (2020), Repetto et al. (2021), Shadiev et al. (2021b), Xie et al. (2019), Yang et al. (2021) |
| Learning motivation | - | The use of 360-degree video in reading lesson increased students' motivation. | 7 | Chen et al. (2021b), Chen et al. (2021c), Chen & Hwang (2020), Chien et al. (2020), Huang et al. (2020), Lin & Wang (2021), Nobrega & Rozenfeld (2019) |
| Cognitive load | - | The EFL learners who watched 360-degree video journalism had higher cognitive load. | 3 | Chen et al. (2021b), Huang et al. (2020), Ji et al. (2019) |
| Self-efficacy | - | VR learning system can enhance students' descriptive article writing self-efficacy. | 3 | Chen et al. (2021b), Huang et al. (2020), Lin & Wang (2021) |
| Language Anxiety | - | The peer-assessment-based on 360-degree video technology can reduce students' English learning anxiety. | 3 | Chen et al. (2021b), Chen & Hwang (2020), Chien et al. (2020) |
| Learning behaviors | Learning behavior engagement | The degree of learning behavior engagement did not show any difference between the experimental and control groups. | 1 | Yang et al. (2021) |
| | Learning behavioral patterns | The tutors and tutees in the experimental group had more interactions and more meaningful communication. | 1 | Chen et al. (2021b) |
| Thinking skills | Creative thinking tendency | The learning system can promote students' higher order creativity tendency. | 1 | Huang et al. (2020) |
| | Critical thinking skills | The peer-assessment based on 360-degree video technology can enhance students' critical thinking skills. | 1 | Chien et al. (2020) |

*Appendix 13.* Affordances of 360-degree video technology

| Affordance | Frequency | Reference |
|---|---|---|
| Create authentic context | 22 | Abd Majid et al. (2020), Berns et al. (2018), Chen et al. (2021a), Chen et al. (2021b), Chen & Hwang (2020), Chien et al. (2020), DeWitt et al. (2022), Dolgunsöz et al. (2018), Ebadi & Ebadijalal (2020), Huang et al. (2020), Ji et al. (2019), Lin et al. (2021), Lin & Wang (2021), Monteiro & Ribeiro (2020), Nobrega & Rozenfeld (2019), Repetto et al. (2021), Shadiev et al. (2021b), Song (2019), Xie et al. (2021), Xie et al. (2019), Yang et al. (2021), Yeh et al. (2021) |
| Provide immersive experience | 20 | Abd Majid et al. (2020), Berns et al. (2018), Chen et al. (2021a), Chen et al. (2021b), Chen et al. (2021c), Chen & Hwang (2020), DeWitt et al. (2022), Dolgunsöz et al. (2018), Huang et al. (2020), Ji et al. (2019), Lin et al. (2021), Lin & Wang (2021), Monteiro & Ribeiro (2020), Nobrega & Rozenfeld (2019), Repetto et al. (2021), Shadiev et al. (2021b), Song (2019), Xie et al. (2021), Xie et al. (2019), Yang et al. (2021) |
| Facilitate language learning | 12 | Abd Majid et al. (2020), Chen et al. (2021c), Chen & Hwang (2020), Ebadi & Ebadijalal (2020), Huang et al. (2020), Lin et al. (2021), Monteiro & Ribeiro (2020), Nobrega & Rozenfeld (2019), Repetto et al. (2021), Shadiev et al. (2021b), Xie et al. (2021), Yang et al. (2021) |
| Provide learning content /material | 11 | Abd Majid et al. (2020), Chen et al. (2021b), Chen & Hwang (2020), Dolgunsöz et al. (2018), Huang et al. (2020), Ji et al. (2019), Lin et al. (2021), Monteiro & Ribeiro (2020), Repetto et al. (2021), Song (2019), Yang et al. (2021) |
| Give a sense of presence | 10 | Abd Majid et al. (2020), DeWitt et al. (2022), Dolgunsöz et al. (2018), Ebadi & Ebadijalal (2020), Huang et al. (2020), Lin et al. (2021), Monteiro & Ribeiro (2020), Xie et al. (2021), Xie et al. (2019), Yang et |

| | | |
|---|---|---|
| | | al. (2021) |
| Enable interaction with learning content | 10 | Berns et al. (2018), Chen & Hwang (2020), Chien et al. (2020), Huang et al. (2020), Lin et al. (2021), Monteiro & Ribeiro (2020), Nobrega & Rozenfeld (2019), Song (2019), Yang et al. (2021), Yeh et al. (2021) |
| Increase motivation | 5 | Abd Majid et al. (2020), Chen et al. (2021c), Chen & Hwang (2020), Nobrega & Rozenfeld (2019), Yang et al. (2021) |
| Reduce speaking anxiety | 3 | Chen & Hwang (2020), Chien et al. (2020), Xie et al. (2019) |
| Improve self-efficacy | 1 | Huang et al. (2020) |
| Promote creative tendency | 1 | Huang et al. (2020) |

*Appendix 14.* Problems related to methodology

| Problems | Description | Frequency | Reference |
|---|---|---|---|
| Small sample size | The number of participants is small | 7 | Chen et al. (2021c), Ebadi & Ebadijalal (2020), Lin & Wang (2021), Nobrega & Rozenfeld (2019), Xie et al. (2021), Xie et al. (2019), Yang et al. (2021) |
| Data collection strategy | Insufficient data for analysis or measure instruments used need to be improved, etc. | 6 | Chen et al. (2021c), Chen et al. (2021b), Chen et al. (2021a), Chien et al. (2020), Lin et al. (2021), Xie et al. (2019) |
| Short period | The time of the learning activities is short | 5 | Chien et al. (2020), Ebadi & Ebadijalal (2020), Huang et al. (2020), Lin et al. (2021), Yang et al. (2021) |
| Lack of a control group | There was only one experimental group in the study | 4 | Chen et al. (2021a), Lin & Wang (2021), Shadiev et al. (2021b), Xie et al. (2019) |
| Problems related to participants | Single group of participants such as single major of participants, single educational level of participants | 2 | Chen et al. (2021c), Dolgunsöz et al. (2018) |
| Experimental design | Such as an unbalanced partnership in intercultural learning activity | 1 | Shadiev et al. (2021b) |
| Lack of formal assessment of English proficiency | students have not been formally assessed on English proficiency prior to the beginning of the study | 1 | Repetto et al. (2021) |

*Appendix 15.* Problems related to technology implementation

| Problems | Description | Frequency | Reference |
|---|---|---|---|
| Physical discomforts | Such as dizziness, eye fatigue and nausea | 9 | Chen & Hwang (2020), DeWitt et al. (2022), Dolgunsöz et al. (2018), Lin et al. (2021), Abd Majid et al. (2020), Monteiro & Ribeiro (2020), Repetto et al. (2021), Song (2019), Xie et al. (2019) |
| Technical difficulties | Such as lack of internet connectivity, slow internet speed, incompatibility of software with specific mobile phones, difficulty to edit videos and different volumes of software | 7 | Chien et al. (2020), DeWitt et al. (2022), Ebadi & Ebadijalal (2020), Monteiro & Ribeiro (2020), Song (2019), Xie et al. (2021), Xie et al. (2019) |
| Low quality of videos | The quality of videos is low and thus participants encountered visibility issue | 2 | Dolgunsöz et al. (2018), Abd Majid et al. (2020) |
| Length of videos | The length of videos should be shorter | 2 | Chien et al. (2020), Shadiev et al. (2021b) |
| A small number | The number of devices (e.g., cameras) is | 2 | Nobrega & Rozenfeld (2019), |

| | | | |
|---|---|---|---|
| of devices | small | | Smith & Townsend (2021) |
| Novelty effect | The higher questionnaire score for the experimental group might be due to participants' novelty for new technology | 2 | DeWitt et al. (2022), Huang et al. (2020) |
| Unfamiliarity with the technology | Participants' insufficient knowledge about VR led to difficulties in using the technology | 1 | Lin & Wang (2021) |
| Missing the text shown | Participants may miss the text shown because they are focused on another part of the video | 1 | Abd Majid et al. (2020) |
| Unsuitable VR goggle size | The size of VR goggle is unsuitable | 1 | Dolgunsöz et al. (2018) |
| Increased cognitive load | The high cognitive load in the VR environments | 1 | Ji et al. (2019) |

*Appendix 16.* Problems related to learning process

| Problems | Description | Frequency | Reference |
|---|---|---|---|
| Lack of adequate feedback from instructor | The instructor was not able to provide a lot of feedback to participants during their presentations due to the constraint of class time | 2 | Ebadi & Ebadijalal (2020), Xie et al. (2021) |
| Lack of attention to the learning status of participants | The researchers needed to pay attention to the users' learning status during the learning process | 1 | Chen et al. (2021b) |
| Lack of consideration of participants' technical competency | Some factors of learners' technology competency were not adequately considered | 1 | Chen & Hwang (2020) |
| Participants' insufficient understanding of the project instruction | Participants did not fully understand the project instructions | 1 | Lin & Wang (2021) |
| Participants' inaccurate pronunciation or unfamiliar vocabulary | Inaccurate pronunciation or unfamiliar vocabulary hindered followers' interactions with the presenters | 1 | Xie et al. (2019) |
| Distraction | Participants become absorbed in the technology at the expense of language use | 1 | Ebadi & Ebadijalal (2020) |

Hsu, I.-Y., & Tsai, F.-H. (2023). Development and Evaluation of a Physical Computing Game-Design Project for Students' Computational Thinking. *Educational Technology & Society*, *26*(4), 38-50. https://doi.org/10.30191/ETS.202310_26(4).0003

# Development and Evaluation of a Physical Computing Game-Design Project for Students' Computational Thinking

**I-Ying Hsu[1] and Fu-Hsing Tsai[2*]**

[1]The Affiliated Senior High School of National Kaohsiung Normal University, Taiwan // [2]Teacher Education Center, National Chiayi University, Taiwan // tosca@tea.nknush.kh.edu.tw // fhtsai@mail.ncyu.edu.tw

[*]Corresponding author

(Submitted July 26, 2022; Revised December 8, 2022; Accepted January 30, 2023)

**ABSTRACT:** This study developed a physical computing game-design project that incorporates block-based programming, physical computing, and computer game design for Taiwan's high school technology education curriculum to strengthen students' computational thinking. The project asked students to develop a somatosensory computer game using a block-based programming language and physical computing devices. This study also attempted to enhance students' attitudes toward programming, technology, and engineering, and to explore the effectiveness differences between students with different majors. The research findings indicate that the project may improve students' computational thinking concepts, but did not improve students' attitudes toward programming, technology, and engineering. While participating science major students' perceptions and attitudes toward technology and engineering were significantly higher than those of social science majors, this study also found that students' performance on their project product showed no significant difference between the different groups of majors. These results imply that the application of this project could be feasible and may be beneficial to deepen science majors' interest in technology and engineering.

**Keywords:** Physical computing, Game design, Computational thinking, Project-based learning

## 1. Introduction

Computational thinking refers to the skills, cognitive procedures, and concepts that computer scientists and engineers use to operate computers and solve problems (Anderson, 2016). Wing (2006) noted that computational thinking should not be a skill exclusive to computer engineers; rather, it should be a fundamental skill that everyone should learn at an early age because such skills can also be used in daily life. This argument has prompted rapid development in computational thinking capabilities in education systems worldwide and led to computational thinking being viewed as a crucial skill for the 21st century. Many countries, such as the United States, the United Kingdom, the Netherlands, Australia, Poland, and South Korea, have prioritized cultivating computational thinking skills through their national K–12 education systems (Hsu et al., 2018; Voogt et al., 2015). Taiwan also includes computational thinking as a vital educational goal for the technology education curriculum of the national 12-year basic education program (Ministry of Education, 2018). Computational thinking has rapidly become a skill to be cultivated in every student from an early age; accordingly, the proper methods to do so have become a focus of research on education.

As programming education is an effective method of strengthening students' computational thinking capabilities (Shute et al., 2017), many countries, such as the United Kingdom, Japan, and Finland, have begun to offer programming courses in primary schools (Seow et al., 2019). This has contributed to a gradual expansion of computer science education at the primary school level. However, programming is extremely difficult to learn for beginners, especially the complex traditional text-based programming syntax (Lu et al., 2020). To make programming easy to learn and interesting, visual/block-based programming languages are considered suitable for novice programmers (Price & Barnes, 2015). For example, Scratch, Makecode, and App Inventor have been widely used in computer programming courses in primary and secondary schools as they enable novice programmers to program by manipulating block-based languages without syntax. Taiwan's technology education curriculum also recommends that students begin learning block-based programming languages in the third and fourth grades (Ministry of Education, 2020). Many studies have noted that using block-based programming languages can strengthen students' computational thinking skills (Rodríguez-Martínez et al., 2020; Sáez-López et al., 2016).

Block-based programming languages facilitate both programming and computer game design activities; for example, popular programming software Scratch uses a block-based programming language that enables students to develop game projects (Maloney et al., 2010). As students in the era of digital natives enjoy playing video games, having them design games using block-based programming languages can encourage them to engage in programming (Mladenović et al., 2018). In addition, game production helps students learn design rules, generate and test creative ideas, and collaborate with others (Ke, 2014). Thus, computer game design is a form of learning by design (Robertson & Howells, 2008). Many studies also indicated that game design can contribute to the development of computational thinking skills (Garneli et al., 2015; Troiano et al., 2019).

Using physical computing devices to help students learn programming has also become a common way to increase students' interest in learning programming. Physical computing devices are programmable devices for which users write programs to control and develop creative, practical, and interactive hardware. Physical computing enables students to understand how programs work in physical objects and to touch the final product, which increases students' interest, unlike traditional programming, where the results are displayed on a monitor (Hodges et al., 2013). In addition, physical programming can provide students with hands-on and building experience. Thus, physical programming is appropriate for technology and engineering education courses that specifically emphasize hands-on activities. Taiwan's new technology education curriculum also recommends exposing elementary students to physical computing during the fifth and sixth grades (Ministry of Education, 2020). With the rise of the maker movement, small and inexpensive physical computing devices, such as microcontrollers, have increasingly been used in programming and technology education; for example, BBC micro:bit is an inexpensive microcontroller with many built-in sensors. As it can be used with block-based programming software, such as MakeCode and Scratch, which are easy for novice programmers to use, the BBC micro:bit has been widely used in programming education in the United Kingdom since 2016 (Ball et al., 2016). Some studies have demonstrated that using BBC micro:bit for physical computing can strengthen students' computational thinking skills (Song et al., 2020; Wu & Su, 2021).

Therefore, to strengthen students' computational thinking skills, learning block-based programming languages and applying them to game design or physical computing are common and feasible teaching strategies that make programming easy to learn and interesting. However, few studies have examined the effects of combining block-based programming with physical computing and computer game design. The BBC micro:bit microcontroller has a built-in accelerometer sensor, which enables it to detect forces in three dimensions, as well as a radio antenna to communicate wirelessly with other micro:bits. These features enable programmers to turn a device into a handheld controller for somatosensory games by applying block-based programming. Thus, block-based programming languages, physical computing devices, and computer game design can be combined with a somatosensory game.

In 2019, Taiwan began to implement a new technology education curriculum, which combines living technology education and information education that originally focused on hands-on practice and information technology, respectively. Thus, in addition to computational thinking, design thinking is a primary educational goal in the new curriculum (Ministry of Education, 2018). As the curriculum prioritizes computational thinking and hands-on practice, providing students with opportunities to engage in physical computing design projects is a good choice for Taiwan's new technology education curriculum. Therefore, this study proposed a physical computing game-design project that incorporates the block-based programming, physical computing, and computer game design for the high school technology education curriculum in Taiwan. That is, this project aimed to guide students to develop a somatosensory computer game using Scratch and BBC micro:bit as the programming language and the game controller, respectively. It also aimed to explore whether it can improve learning outcomes in terms of computational thinking concepts and attitudes toward computer programming, technology, and engineering.

Numerous studies have explored gender differences in game design (Hsu, 2013), programming, and computational thinking skills (Mouza et al., 2020; Wu & Su, 2021). However, few studies have explored the differences of learning effectiveness in computational thinking among students with different majors. High school students in Taiwan are grouped into science or social science majors in 11th grade based on aptitude. Science majors study in depth on mathematics, physics, and chemistry, while social science majors focus on social science courses, such as geography and history. However, as the technology education curriculum is compulsory for Taiwan's high school students, this study also explored the differences in the effects of the project among different majors. Specifically, the objectives of this study were, as follows.

- Explore the changes and differences of students with different majors regarding computational thinking concepts and attitudes toward computer programming, technology, and engineering after participating in the physical computing game-design project.
- Explore the differences of students with different majors regarding perceptions and project results after taking part in the physical computing game-design project.

## 2. Literature review

### 2.1. Visual / block-based programming languages

To date, many people advocate that providing programming courses is an important way to cultivate students' computational thinking (Shute et al., 2017; Voogt et al., 2015). To facilitate learning, visual and block-based programming languages are used as the primary tool in introductory programming courses for K-12 students (Portelance et al., 2016). Visual programming languages provide a shortcut for producing programming code by using icons and graphical objects to represent instructions (Myers, 1990). They are suitable for young students in the era of digital natives because they enable students to solve programming problems through trial and error (Mladenović et al., 2018).

Scratch, which is a visual programming software, was developed by MIT Media Lab to increase young students' interest in programming (Maloney et al., 2010). Students program by dragging and dropping visual blocks, similar to building blocks, which makes it a block-based programming language. The ability to stack blocks is perfect for novice programmers (Price & Barnes, 2015) as it eliminates the restrictions of traditional textual languages (João et al., 2019). According to Brennan and Resnick (2012), Scratch involves seven concepts of computational thinking: sequences, loops, events, parallelism, conditionals, operators, and data, and students can transfer these concepts to other programming or non-programming tasks. Empirical studies have noted that Scratch facilitates the development of students' computational thinking (Rodríguez-Martínez et al., 2020; Sáez-López, 2016).

Scratch can also shift the focus of programming activities from math problems to game design (Mladenović et al., 2018). As it simplifies game design, Scratch is widely used to design computer games to help students learn programming and develop computational thinking (Garneli et al., 2015). Kafai and Burke (2015) analyzed 55 studies on game design-based learning and discovered that game production helped the students learn computing concepts. Zur-Bargury et al. (2013) used Scratch in a middle school's computer science course and discovered that it helped the students learn the computational concept of loops. Mladenović et al. (2018) also discovered that using Scratch to design games was more effective than textual language-based software as it prevented misunderstanding of the computational concept of loops.

### 2.2. Physical computing

The term physical computing was derived from O'Sullivan and Igoe (2004) and refers to an interactive application that combines virtual and physical worlds through computer programming, sensors, microcontrollers, and tangible materials. It is also known as digital making or tangible programming (Kotsopoulos et al., 2017). Physical computing helps students learn about hardware, software, and design, and makes abstract programming concepts concrete (Kotsopoulos et al., 2017). It also encourages students to use their creativity to invent physical interactive devices (Przybylla & Romeike, 2014). As physical computing involves programming, embedded systems, and electronic engineering, physical computing education was previously only offered in university education; however, with the advances in physical devices and easily used visual programming, young students can now engage in physical programming (Jang et al., 2016). The maker movement has also facilitated the development of physical computing, as several making activities (e.g., creating robots) involve physical computing.

The physical computing device, BBC micro:bit, is an inexpensive and programmable microcontroller. It has a variety of on-board modules, including LEDs, alight sensor, compass, accelerometer, and programmable buttons. The United Kingdom has used this device for programming education since 2016, and students in more than 50 countries have used this device to learn programming and create physical products (Austin et al., 2020). In some studies, students have used the micro:bit to create paper-cutting lamps (Lu et al., 2021), headbands, stuffed animals (Klimová, 2020), and wireless remote-control car (Austin et al., 2020). Most students indicated that the micro:bit is easy to use and increased their motivation to learn programming (Gibson & Bradley, 2017). Some studies have also indicated that applying the micro:bit to physical computing activities improves students' computational thinking skills (Song et al., 2020; Wu & Su, 2021).

## 3. Methods

### 3.1. Physical computing game-design project

This physical computing game-design project was implemented in an 11th-grade technology education course (two 50-min classes per week for 20 weeks). The first half of the project lasted 9 weeks. The teacher taught the students how to use Scratch and bDesigner, which provide more block-based instructions for the micro:bit, to control micro:bit with different sensing modules (buttons, accelerometer, buzzers, and LEDs). The students also learned to wirelessly transmit information between micro:bits to create a simple somatosensory game and used the IFTTT web service to upload the game scores to the cloud services.

The second half of the project lasted nine weeks. The students worked in groups of three to five, depending on their preferences, and used Scratch and micro:bits to design a somatosensory game with any theme. Each group was provided with two mciro:bits. The first micro:bit was used to detect the player's movement for the game controller. Each group was required to use various materials, such as cardboard, to create a game controller according to their game theme, and to affix the first micro:bit in the self-developed game controller. The second micro:bit was connected with a computer as the receiver and transmitter for receiving the movement information from the first micro:bit and transmitting the messages to the self-developed Scratch game. Finally, each group was asked to present their final products in the last week.

### 3.2. Research procedure

This experiment was conducted in a high school's technology education course for 20 weeks. To explore the learning outcomes, during the first and last weeks, the students took pre-tests and post-tests on computational thinking concepts and completed attitude scales regarding computer programming, technology, and engineering. The participants were also required to indicate their perceptions of the project on a scale during the last week. Besides, students participated in the learning sessions from week 2 to week 10, and their project designs from week 11 to week 19.

### 3.3. Research participants

This study recruited 11th-grade students from a high school in southern Taiwan as these students accepted the new compulsory technology education curriculum. To explore the learning effectiveness differences between majors, one class of science majors (45 students) and two classes of social science majors (54 students) were selected. As four science majors did not participate in the entire experiment, 41 science majors and 54 social science majors remained, and they were divided into 10 and 14 groups, respectively, during the project. Ethical approval for this study was waived by the Taiwan Centers for Disease Control Policy # 1010265075, as this study was conducted in a general teaching environment for educational purposes, and all participants were provided the same teaching procedure and activities regardless of participants' major. This study collected no data that could identify specific individuals.

### 3.4. Research instruments

#### *3.4.1. Computational thinking test*

This study developed a computational thinking test including 10 items related to Scratch-based visual programming problems, which encompassed the seven computational thinking concepts of sequences, loops, events, parallelism, conditionals, operators, and data, as proposed by Brennan and Resnick (2012). Items 1–7 tested a single concept each, and Items 8–10 tested multiple concepts simultaneously, and the highest total score was 100 points. Figure 1 presents an example item testing the computational thinking concept of the data. Figure 2 presents an example item testing multiple concepts, namely data, operators, conditionals, and loops. The test was reviewed by a technology education teacher to ensure its appropriateness. According to a pilot study with 134 high school students, the Kuder-Richardson reliability was .73.

*Figure 1.* Computational thinking test on the concept of data

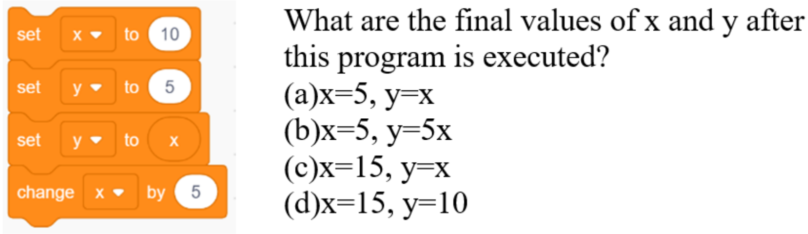
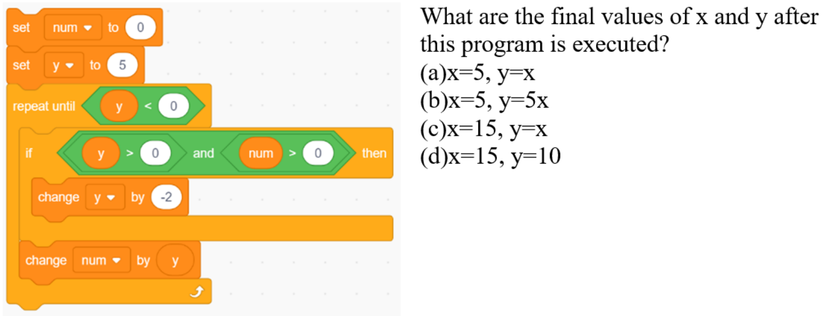



*Figure 2.* Computational thinking test on the concept of data, operators, conditionals, and loops




### *3.4.2. Computer programming attitude scale*

This study referenced Korkmaz and Altun (2013) to create a computer programming attitude scale comprising three dimensions: confidence, preference, and usefulness. Each dimension contained four items, and each of which was answered on a 5-point scale. Items in the confidence and preference dimensions included "I am confident that I can learn computer programming" and "I like computer programming", respectively. The item in the usefulness dimensions included "I don't think programming will be useful in my life." According to a pilot study with 134 high school students, the Cronbach's α of this scale was .92.

### *3.4.3. Technology and engineering attitude scale*

This study used the 9-item subscale of technology and engineering attitude in the STEM attitude scale, as developed by Faber et al. (2013), to evaluate the students' attitude changes toward technology and engineering after the project. A 5-point scale was used for scoring the items, such as "I like to imagine creating new products," "I am good at building and fixing things," and "If I learn engineering, then I can improve things that people use every day." According to a pilot study with 153 high school students, the Cronbach's α of this scale was .91.

### *3.4.4. The rubric of project assessment*

This study referenced the Creativity Product Analysis Matrix of Besemer (1998) to develop a project assessment rubric (Table 1). The students' projects were evaluated in terms of two dimensions. The first was resolution, which evaluated the game's logic and comprehensibility, playability, and innovativeness; the highest score was 40 points. The second dimension was elaboration, which evaluated the game's basic operation and the quality of the game's software and hardware; the highest score was 60 points. Table 1 presents the scoring standards for each indicator. The evaluation standards were provided to the students as a reference before the project. All the final project products, as developed by the science and social science majors, were scored by two scorers using the abovementioned rubric, and the average scores were their final product scores. The scorer reliability was .94.

### *3.4.5. Participation perception scale*

This study developed a 5-point scale with 11 items to understand the participants' perceptions after the project. The questions assessed the participants' feelings towards the project, as well as their knowledge acquisition, such as "I thought this project was interesting," "I acquired knowledge and skills about micro:bit through this project," and "I felt a sense of achievement when we finished our somatosensory game." An open-ended question was included at the end of this scale, which enabled students to further provide their perceptions and suggestions. The Cronbach's α of this scale was .92.

*Table 1.* Project assessment indicators

| Dimension | Indicators | Scoring standards | Scores |
|---|---|---|---|
| 1. Resolution | 1.1. Logical and understandable | • Is the game understandable and logical?<br>• Does the game have a name, goal, and rules?<br>• Is the game based on somatosensory principles? | 0–20 |
| | 1.2. Playable | • Is the game style interesting?<br>• Can the game attract players to play constantly? | 0–10 |
| | 1.3. Innovative | • Is the game original and creative as a somatosensory game? | 0–10 |
| 2. Elaboration | 2.1. Basic operation | • Does the game work on Scratch?<br>• Can the game be controlled using a wireless game controller?<br>• Does the game run smoothly? | 0–20 |
| | 2.2. Quality of software | • Are the game's graphics and art refined?<br>• Is the game's software optimized? Does it have bugs? | 0–20 |
| | 2.3. Quality of hardware | • Is the game controller well produced?<br>• During the game, can the micro:bit be fixed firmly in the game controller without affecting the game performance? | 0–20 |

## 4. Results

### 4.1. Analysis of students' computational thinking concepts

To determine how the students' understanding of concepts in computational thinking was changed by participating in the project, their pre-test and post-test scores in the computational thinking test were analyzed through paired sample *t*-test, and Table 2 presents the results. The science major group's post-test scores ($M$ = 77.56, $SD$ = 17.72) after the project were significantly higher than their pre-test scores ($M$ = 68.29, $SD$ = 18.01), $t(40) = 3.53$, $p < .05$. The social science major group's post-test scores ($M$ = 67.96, $SD$ = 22.1) were also significantly higher than their pre-test scores ($M$ = 54.07, $SD$ = 19.18), $t(53) = 4.69$, $p < .05$. The results indicate that the proposed project may improve students' computational thinking concepts, no matter the major.

Students' scores on different questions regarding computational thinking concepts were also explored, including seven single-concept questions (10 points per concept) and three multiple-concept questions (30 points). All students performed well in the computational thinking concepts of sequences, parallelism, and conditionals; however, the pre-test scores were low for the concepts of data, operators, loops, and multiple concepts, as shown in Table 2. After the project, the science major group's post-test scores for the concepts of data ($t(40) = 3.13$, $p < .05$) and loops ($t(40) = 2.5$, $p < .05$) were significantly higher than their pre-test scores. The social science major group's post-test scores for the concepts of data ($t(53) = 2.7$, $p < .05$), loops ($t(53) = 4.01$, $p < .05$), events ($t(53) = 3.23$, $p < .05$), and multiple-concepts ($t(53) = 2.72$, $p < .05$) were also significantly higher than their pre-test scores. In other words, the proposed project could help most students improve their computational thinking concepts regarding data and loops, and additionally improve social science major students' computational thinking concepts regarding events and multiple concepts.

To identify differences in the effectiveness of the project between students with different majors, this study used the total scores in the computational thinking pre-test and post-test scores as the covariable and dependent variable, respectively, to conduct a one-way analysis of covariance (ANCOVA). Before ANCOVA, the assumption of homogeneity of regression was conducted ($F(1, 91) = 0.02$, $p > .05$) and was not violated. The results of the ANCOVA were $F(1, 92) = 0.27$, $p > .05$, which indicated that after the effects of the covariates were excluded, and the factor of the student's major did not significantly affect the post-test scores. The post hoc comparison using the least significant difference (LSD) also revealed no significant difference in post-test scores between the science ($M$ = 73.26) and social science majors ($M$ = 71.23). Moreover, this study used the pre-test and post-test scores of each concept and multiple-concept scores in computational thinking as the covariable and dependent variable, respectively, to conduct one-way ANCOVA, and the results indicated that the post-test scores did not differ significantly between majors. In other words, although the science majors' pre-test scores were significantly higher than those of the social science majors scores in the computational thinking test ($t(93) = 3.67$, $p < .05$), after the effects of prerequisite abilities were excluded, the learning effectiveness for students with different majors to acquire computational thinking concepts had no significant difference. Therefore, the proposed project could help students improve their understanding of computational thinking concepts, and the improvement efficiency had no significant difference between students with different majors.

*Table 2.* Results of Computational Thinking Test

| | Science major group | | | Social science major group | | | Science and social science major groups |
|---|---|---|---|---|---|---|---|
| | Pre-test | Post-test | $t$ | Pre-test | Post-test | $t$ | ANCOVA |
| Total | 68.29 | 77.56 | 3.53* | 54.07 | 67.96 | 4.69* | .27 |
| Sequences | 10.00 | 10.00 | 0.00 | 9.44 | 9.81 | 1.43 | .00 |
| Events | 8.04 | 8.75 | 1.00 | 5.93 | 8.33 | 3.23* | .01 |
| Parallelism | 10.00 | 9.76 | -1.00 | 8.89 | 9.44 | 1.14 | .16 |
| Conditionals | 8.78 | 9.27 | 0.81 | 8.15 | 8.15 | 0.00 | 1.98 |
| Data | 3.90 | 6.59 | 3.13* | 2.04 | 4.26 | 2.70* | 3.30 |
| Operators | 6.83 | 6.83 | 0.00 | 6.11 | 6.48 | 0.42 | .04 |
| Loops | 3.90 | 6.34 | 2.50* | 1.85 | 5.00 | 4.01* | .74 |
| Multiple | 16.83 | 20.00 | 1.92 | 11.67 | 16.48 | 2.72* | .88 |

*Note.* *$p < .05$.

### 4.2. Analysis of students' computer programming attitudes

To determine how the project changed the students' attitudes toward computer programming, this study analyzed the students' responses to the computer programming attitude scale before and after the experiment. To facilitate statistical analysis, the answers were converted into points, with strongly agree, agree, neither agree nor disagree, disagree, and strongly disagree corresponding to the scores of 5–1 points, respectively, and Table 3 presents the statistical results. The scores indicate that the attitudes of the science majors both before and after the project were overall positive. However, the paired sample $t$-test results indicate that the mean scores before ($M = 3.51$, $SD = 0.84$) and after the project ($M = 3.46$, $SD = 0.99$) did not differ significantly ($t(40) = -0.88$, $p > .05$). The social science majors' attitudes were less positive before the project ($M = 2.44$, $SD = 0.76$), and the total mean score after the project still indicated a negative attitude ($M = 2.49$, $SD = 0.80$) without significant improvement ($t(53) = 0.69$, $p > .05$). These results suggest that before the project, the science majors had significantly more positive attitudes than the social science majors, and further analysis revealed a significant difference between the groups' attitudes before the project ($t(93) = 6.50$, $p < .05$). However, after the project, students' attitudes toward computer programming had no significant change, no matter the major. This result indicated that the proposed project may not improve students' overall computer programming attitudes.

This study also explored students' scores in the subscales. The responses to the subscales indicated that despite a low score in confidence, the science majors had a positive attitude toward programming before the project. No significant change in any of the three attitude categories was observed after the project, as the paired sample $t$-test results for the three sub-items indicated no significant difference (Table 3). The confidence and preference of social science major students toward programming before the experiment were low, and further analysis revealed that their initial confidence ($t(93) = 6.53$, $p < .05$) and preference scores ($t(93) = 6.95$, $p < .05$) differed significantly from those of the science majors. This is the primary reason that the social science majors had lower scores for overall attitudes toward computer programming before the project. However, the subscale results indicated that they found programming useful. After the project, the social science majors did not exhibit significant changes in any of the three attitude categories, and their confidence and preference toward computer programming still did not improve significantly, indicating a negative attitude. In other words, the project did not help most students improve their programming attitude regarding confidence, preference, and usefulness, regardless of major. The findings also indicated that the social science majors initially had a negative attitude toward programming in terms of their confidence and preference, which differed significantly from those of the science majors.

To determine whether the major factor affected the improvement of students' attitudes toward programming, the total mean pre-test scores of the programming attitude were used as a covariate, and the total mean post-test scores were used as a dependent variable, to conduct a one-way ANCOVA. Before ANCOVA, the assumption of homogeneity of regression was conducted ($F(1, 91) = 1.08$, $p > .05$) and was not violated. The results of the ANCOVA were $F(1, 92) = 0.20$, $p > .05$, indicating that after the effects of the covariates were excluded, the factor of the student's major did not significantly affect the post-test scores. The LSD post hoc comparison also revealed no significant difference in mean post-test scores between the science ($M$ = 2.95) and social science majors ($M$ = 2.88). Moreover, this study used the scores of each subscale in the programming attitude pre-test and post-test as the covariable and dependent variable, respectively, to conduct a one-way ANCOVA. The results also indicated that the post-test scores did not differ significantly between majors (Table 3). Therefore, the project may not have significantly improved the students' attitudes toward programming in the short term. No difference in improvement efficiency was observed between the groups.

*Table 3.* Results of computer programming attitude

| | Science major group | | | Social science major group | | | Science and social science major groups |
|---|---|---|---|---|---|---|---|
| | Pre-test | Post-test | $t$ | Pre-test | Post-test | $t$ | ANCOVA |
| Mean | 3.51 | 3.46 | -.88 | 2.44 | 2.49 | .69 | .20 |
| Confidence | 3.17 | 3.16 | -.05 | 1.94 | 2.05 | 1.07 | 1.43 |
| Preference | 3.41 | 3.40 | -.06 | 2.00 | 2.12 | 1.17 | .33 |
| Usefulness | 3.96 | 3.82 | -1.00 | 3.37 | 3.30 | -.62 | .70 |

### 4.3. Analysis of students' technology and engineering attitudes

To determine whether the project changed the students' attitudes toward technology and engineering, this study analyzed the students' responses to the technology and engineering attitude scale, which ranged from strongly agree to strongly disagree, corresponding to scores of 5–1 points, respectively. The mean score of the science majors before the project was higher than 3 points ($M$ = 3.67, $SD$ = 0.76), and their attitude toward technology and engineering was positive. After the project, despite a slight improvement in their mean score ($M$ = 3.76, $SD$ = 0.62), no significant difference was observed by paired sample $t$-test ($t(40) = 1.11$, $p > .05$). The mean score of the social science majors before the project was lower than 3 points ($M$ = 2.51, $SD$ = 0.64), indicating that their attitude toward engineering and technology was negative. Their mean score after the project was also similar to that before the project ($M$ = 2.52, $SD$ = 0.67), indicating no significant change by paired sample $t$-test ($t(53)$ = 0.04, $p > .05$). Further analysis revealed a significant difference in mean scores between the groups before the project ($t(93) = 10.94$, $p < .05$). These results indicated that the project might not have significantly improved the students' attitudes toward technology and engineering. The science majors maintained a positive attitude toward technology and engineering, whereas the social science majors still had a negative attitude.

This study also performed a one-way ANCOVA using the mean pre-test and post-test scores as the covariate and dependent variable, respectively, to determine whether the factor of student's major significantly affected their improvement in attitudes toward technology and engineering. Before ANCOVA, the assumption of homogeneity of regression was conducted ($F(1, 91) = 0.06$, $p > .05$) and was not violated. The results of the ANCOVA were $F(1, 92) = 16.95$, $p < .05$, $\eta^2 = .16$, indicating that after the effects of the covariates were excluded, the factor of student's major significantly affected the post-test scores. The LSD post hoc comparison also revealed a significant difference in the adjusted mean post-test scores between the science (M = 3.37) and social science majors ($M$ = 2.81), which means that the science majors exhibited a significantly larger change in technology and engineering attitudes than did the social science majors. Therefore, although the proposed project could not help all students significantly improve their attitudes toward technology and engineering, the improvement efficiency had a significant difference between students with different majors, thus, the project may improve the science majors' attitudes more than those of the social science majors.

### 4.4. Analysis of students' project products

Regardless of the students' majors, all project teams accomplished different somatosensory games according to the project task. Figure 3 presents examples of the final products. One group of social science majors transformed the micro:bit into a dumbbell that controlled a jumping game when lifted. Another group of science majors combined the micro:bit with a door handle and created a somatosensory game involving unlocking doors.

Students' project products were scored according to the rubric shown in Table 1. To explore the differences in project performances between different majors, independent sample $t$-test was conducted on the students' project scores, and Table 4 presents the analysis results of students' project products. Although the social science majors had slightly higher total scores ($M$ = 77.21, $SD$ = 7.71) than the science majors ($M$ =71.90, $SD$ =11.95), the difference was nonsignificant ($t(22) = -1.33$, $p > .05$), indicating that their performance did not differ significantly. Regarding the scores on the dimension of resolution, the science ($M$ = 29.90, $SD$ = 4.31) and social science majors ($M$ = 30.21, $SD$ = 2.91) had similar scores that did not differ significantly ($t(22) = -0.21$, $p > .05$). No significant differences in the scores for the three indicators in this dimension were observed between groups (Table 4). Thus, regardless of their major, most students produced logical games, and their games' playability and innovativeness performance did not differ significantly between groups. Regarding the scores of the dimension of elaboration, although the total quality scores of the social science majors ($M$ = 47, $SD$ = 5.04) were slightly higher than those of the science majors ($M$ = 42.00, $SD$ = 7.96), the difference was nonsignificant. In addition, the scores of social science majors in each of the three indicators of the elaboration dimension were slightly higher than those of the science majors; however, no significant difference was observed. This suggests that no significant difference in outcome was observed between the groups. Therefore, although the social science majors exhibited poor performance in terms of their understanding of concepts in computational thinking, and they had a negative attitude toward computer programming before the project, their performance did not differ significantly from that of the science majors, and they all accomplished the project task.

*Figure 3.* Students' somatosensory games

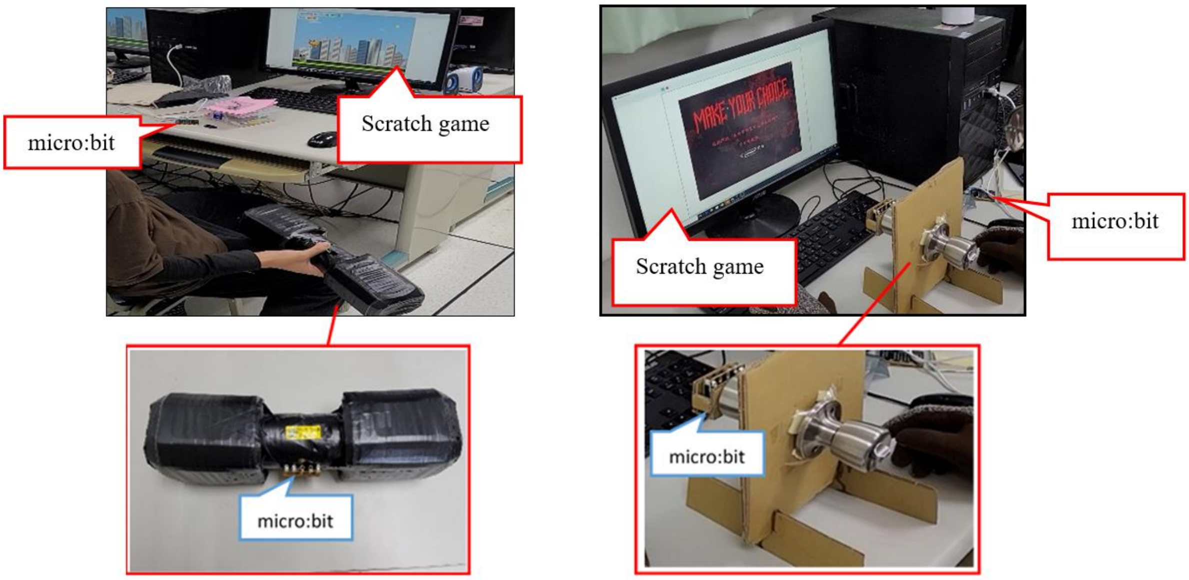



*Table 4.* Analysis of final products

| | Science major group ($n$ = 10) | Social Science major group ($n$ = 14) | $t$ |
|---|---|---|---|
| Total score | 71.90 | 77.21 | -1.33 |
| 1. Resolution | 29.90 | 30.21 | -.21 |
| 1.1: Logical and understandable | 15.90 | 16.64 | -1.15 |
| 1.2: Playable | 7.10 | 7.07 | .05 |
| 1.3: Innovative | 6.90 | 6.50 | .56 |
| 2. Elaboration | 42.00 | 47.00 | -1.89 |
| 2.1: Basic operation | 14.70 | 16.50 | -2.06 |
| 2.2: Quality of software | 14.70 | 15.43 | -.86 |
| 2.3: Quality of hardware | 12.60 | 15.07 | -1.72 |

### 4.5. Analysis of students' participating perception

To identify the differences in students' participating perceptions toward the project between the different majors, independent sample $t$-test was performed on the students' responses to the perception scale. To facilitate statistical analysis, responses were converted to the scores of 1 (*strongly disagree*) to 5 points (*strongly agree*). While the participating perceptions of both the science ($M$ = 3.84) and social science majors ($M$ = 3.55) were positive, the $t$-test results were $t(93) = 2.13$, $p < .05$, indicating that the mean score of the science majors was significantly higher than that of the social science majors. This suggests that the science majors' impressions of the project were more positive than those of the social science majors in terms of satisfaction and gain.

Analysis of the students' responses to the open-ended questions on the scale indicated that most students were satisfied with the project. Social science majors stated, "Designing a somatosensory game on my own was a unique experience," "Now I know how to apply programming for a range of tasks," and "I learned how to create new games with my team members." Science majors stated, "The project helped me understand Scratch and micro:bit more," "This was my first experience using a microcontroller to execute a program I wrote, and it was really exciting and might help me in the future," and "I learned a lot about micro:bit."

## 5. Discussion

This study found that both science and social science major groups' post-test scores on computational thinking tests were significantly higher than those of their respective pre-test scores after the project, and the results of ANCOVA showed no significant difference between the groups. In other words, students significantly improved their scores in computational thinking, and the improvement efficiency did not differ from their initial concepts of computational thinking after participating in the project. Thus, this project helped all students strengthen their computational thinking concepts. The findings demonstrate that the combination of block-based programming, physical computing, and computer game design may be beneficial to students' computational thinking. The findings also support those of other studies regarding the use of Scratch, physical computing, and game design to improve students' computational thinking skills (Mladenović et al., 2018; Rodríguez-Martínez et al., 2020; Sáez-López, 2016; Wu & Su, 2021; Zur-Bargury et al., 2013).

Moreover, the analysis results of the students' computational thinking test also indicated that the changes in their scores were mainly due to the significant increase in all students' computational thinking concepts regarding data and loops, and social science majors' events concept and multiple concepts. However, the students' computational thinking concepts of data, operators, loops, and multiple concepts can still be improved. These results are consistent with those of other studies, in which novice programmers exhibited poor performance in terms of their understanding of data, operators, and loops concepts (Grover & Basu, 2017).

Regarding students' attitudes toward computer programming, technology, and engineering, this study found that students with different majors differed significantly in their attitudes before the project; the social science majors had negative attitudes, lacked confidence and preference for programming, and were uninterested in technology and engineering, whereas the science majors exhibited positive attitudes. After the project, the students' attitudes did not differ significantly from those before the project. The social science majors' confidence and preference for programming did not increase significantly, and their attitudes toward technology and engineering remained similar, as did those of the science majors. This suggests that the project did not significantly improve the students' attitudes toward computer programming, technology, or engineering, which may be because the project was too short to result in immediate changes in attitude or because the students had already established career goals related to their majors. Nonetheless, the project could still be effective to influence science major students' attitudes toward technology and engineering. The ANCOVA analysis results of the technology and engineering attitude scale suggests that the project improved the science majors' technology and engineering attitudes more than the social science majors' attitudes.

The science and social science majors did not significantly differ in terms of the scores for their project products, which indicates that although science majors initially exhibited a deeper understanding of programming concepts and more positive attitudes toward programming and technology creation than the social science majors, their somatosensory games were not significantly superior to those of social science majors. In addition, although the social science majors exhibited poor performance in terms of their initial understanding of the concepts of computational thinking and lacked interest in programming and technological creation, they completed the project successfully after a semester of study and activities. In other words, the difficulty level of the proposed project was appropriate, and the approach may be appropriate for all students. Therefore, this project could continue to be implemented in Taiwan's technology education curriculum.

Regarding the students' perceptions toward the project, this study found that most students were satisfied and indicated that it helped them acquire knowledge. This supports the above discussion, indicates that the difficulty of the project was appropriate, and explains the finding that social science majors can produce the same quality products as the science majors and improve their computational thinking. The analysis results also revealed that science majors' scores for participating perception were significantly higher than those of the social science majors, indicating that the science majors enjoyed the project more than the social science majors; however, this result might be related to the students' differences in attitudes toward programming and technology between majors.

## 6. Conclusion

This study combined block-based programming, physical computing, and game design in a project for Taiwan's high school technology education curriculum to increase interest in computer programming and technology and strengthen computational thinking. This study also explored differences in the project's effects between majors, and the results indicate that the project may improve students' computational thinking concepts; however, the project may not improve students' attitudes toward programming, technology, and engineering. Regarding the comparisons of different majors, this study found science majors exhibited better performance in their computational thinking test, and they had significantly more positive attitudes toward programming and technology than the social science majors before the project. Nevertheless, no significant differences were observed in the computational thinking concepts, programming and technology attitudes, and the final product scores after the project between the groups. These findings imply that the proposed project could be beneficial to all students and feasible for a compulsory technology education curriculum. Moreover, this study also found that science majors' scores regarding their participation perception and technology and engineering attitude were significantly higher than those of the social science majors, which implies that this project may be beneficial to deepen science majors' interest in technology and engineering.

Therefore, this proposed project can continue to be promoted in Taiwan's technology education curriculum; however, future research should increase the number of participants to further verify the effectiveness between different majors. In addition, more rigorous methods can be conducted in the future to assess students' projects.

## Acknowledgment

This research was financially supported by the Ministry of Science and Technology, Taiwan, under Grant no. MOST 109-2511-H-415-008-MY3.

## References

Anderson, N. D. (2016). A Call for computational thinking in undergraduate psychology. *Psychology Learning & Teaching, 15*(3), 226–234. https://doi.org/10.1177/1475725716659252

Austin, J., Baker, H., Ball, T., Devine, J., Finney, J., De Halleux, P., Hodges, S., Moskal, M., & Stockdale, G. (2020). The BBC micro: bit: from the UK to the world. *Communications of the ACM, 63*(3), 62-69. https://doi.org/10.1145/3368856

Ball, T., Protzenko, J., Bishop, J., Moskal, M., De Halleux, J., Braun, M., Hodges, S., & Riley, C. (2016). Microsoft touch develop and the BBC Micro: bit. In *2016 IEEE/ACM 38th International Conference on Software Engineering Companion (ICSE-C)* (pp. 637-640). IEEE.

Besemer, S. P. (1998). Creative product analysis matrix: Testing the modelstructure and a comparison among products--Three novel chairs. *Creativity Research Journal, 11*(4), 333–346. https://doi.org/10.1207/s15326934crj1104_7

Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. In *Proceedings of the 2012 annual meeting of the American educational research association, Vancouver, Canada* (Vol. 1, pp. 25).

Faber, M., Unfried, A., Wiebe, E. N., Corn, J., Townsend, L. W., & Collins, T. L. (2013). Student attitudes toward STEM: The development of upper elementary school and middle/high school student surveys. In *2013 ASEE Annual Conference & Exposition* (pp. 23-1094). https://doi.org/10.18260/1-2--22479

Garneli, V., Giannakos, M. N., & Chorianopoulos, K. (2015). Computing education in K-12 schools: A Review of the literature. In *2015 IEEE Global Engineering Education Conference (EDUCON)* (pp. 543-551). IEEE.

Gibson, S., & Bradley, P. (2017). A Study of Northern Ireland key stage 2 pupils' perceptions of using the BBC Micro: bit in STEM education. *The STeP Journal, 4*(1), 15-41.

Grover, S., & Basu, S. (2017). Measuring student learning in introductory block-based programming: Examining misconceptions of loops, variables, and boolean logic. In *Proceedings of the 2017 ACM SIGCSE technical symposium on computer science education* (pp. 267-272). https://doi.org/10.1145/3017680.3017723

Hodges, S., Scott, J., Sentance, S., Miller, C., Villar, N., Schwiderski-Grosche, S., Hammil, K., & Johnston, S. (2013). . NET Gadgeteer: A New platform for K-12 computer science education. In *Proceeding of the 44th ACM technical symposium on Computer science education* (pp. 391-396). https://doi.org/10.1145/2445196.2445315

Hsu, H. M. J. (2013). Gender differences in elementary school students' game design preferences. *International Journal of Information and Education Technology, 3*(2), 172-176.

Hsu, T. C., Chang, S. C., & Hung, Y. T. (2018). How to learn and how to teach computational thinking: Suggestions based on a review of the literature. *Computers & Education, 126*, 296-310. https://doi.org/10.1016/j.compedu.2018.07.004

Jang, Y., Lee, W., & Kim, J. (2016). The Changes of middle school students' perception and achievement based on the teaching method in physical computing education. *Indian Journal of Science and Technology, 9*(24), 1-8.

João, P., Nuno, D., Fábio, S. F., & Ana, P. (2019). A Cross-analysis of block-based and visual programming apps with computer science student-teachers. *Education Sciences, 9*(3), 181. https://doi.org/10.3390/educsci9030181

Kafai, Y. B., & Burke, Q. (2015). Constructionist gaming: Understanding the benefits of making games for learning. *Educational psychologist*, *50*(4), 313-334. https://doi.org/10.1080/00461520.2015.1124022

Ke, F. (2014). An Implementation of design-based learning through creating educational computer games: A Case study on mathematics learning during design and computing. *Computers & Education*, *73*, 26-39. https://doi.org/10.1016/j.compedu.2013.12.010

Klimová, N. (2020). Wearables: Educational projects made with the BBC micro: bit. In *2020 18th International Conference on Emerging eLearning Technologies and Applications (ICETA)* (pp. 323-329). IEEE.

Korkmaz, Ö., & Altun, H. (2013). Engineering and ceit student's attitude towards learning computer programming. *International Journal of Social Science, 6*(2), 1169–1185.

Kotsopoulos, D., Floyd, L., Khan, S., Namukasa, I. K., Somanath, S., Weber, J., & Yiu, C. (2017). A Pedagogical framework for computational thinking. *Digital Experiences in Mathematics Education, 3*, 154-171.

Lu, C. C., Hong, J. C., Chen, F. F., & Ma, S. Y. (2020). Elementary school students learn Arduino programming to assemble sensory-controlled works. *Int. J. Inf. Educ. Technol, 10*(4), 265-270.

Lu, S. Y., Lo, C. C., & Syu, J. Y. (2021). Project-based learning oriented STEAM: The Case of micro–bit paper-cutting lamp. *International Journal of Technology and Design Education*, *32*, 2553–2575. https://doi.org/10.1007/s10798-021-09714-1

Maloney, J., Resnick, M., Rusk, N., Silverman, B., & Eastmond, E. (2010). The Scratch programming language and environment. *ACM Transactions on Computing Education (TOCE), 10*(4), 1-15. https://doi.org/10.1145/1868358.1868363

Ministry of Education (2018). *Curriculum guidelines of 12-year basic education for Technology domain*. https://www.k12ea.gov.tw/files/class_schema/課綱/13-科技/13-1/十二年國民基本教育課程綱要國民中學暨普通型高級中等學校—科技領域.pdf

Ministry of Education (2020). *The Reference instructions of curriculum development for the primary living technology and information technology education*. https://www.naer.edu.tw/upload/1/9/doc/280/1090611_國民小學科技教育及資訊教育課程發展參考說明.pdf

Mladenović, M., Boljat, I., & Žanko, Ž. (2018). Comparing loops misconceptions in block-based and text-based programming languages at the K-12 level. *Education and Information Technologies, 23*, 1483-1500. https://doi.org/10.1007/s10639-017-9673-3

Mouza, C., Pan, Y. C., Yang, H., & Pollock, L. (2020). A Multiyear investigation of student computational thinking concepts, practices, and perspectives in an after-school computing program. *Journal of Educational Computing Research, 58*(5), 1029-1056. https://doi.org/10.1177/07356331209056

Myers, B. A. (1990). Taxonomies of visual programming and program visualization. *Journal of Visual Languages & Computing, 1*(1), 97-123. https://doi.org/10.1016/S1045-926X(05)80036-9

O'Sullivan, D., & Igoe, T. (2004). *Physical computing: Sensing and controlling the physical world with computers*. Course Technology Press.

Portelance, D. J., Strawhacker, A. L., & Bers, M. U. (2016). Constructing the ScratchJr programming language in the early childhood classroom. *International Journal of Technology and Design Education, 26*, 489-504. https://doi.org/10.1007/s10798-015-9325-0

Price, T. W., & Barnes, T. (2015). Comparing textual and block interfaces in a novice programming environment. In *Proceedings of the eleventh annual international conference on international computing education research* (pp. 91-99). https://doi.org/10.1145/2787622.2787712

Przybylla, M., & Romeike, R. (2014). Physical computing and its scope--Towards a constructionist computer science curriculum with physical computing. *Informatics in Education, 13*(2), 241-254.

Robertson, J., & Howells, C. (2008). Computer game design: Opportunities for successful learning. *Computers & Education, 50*(2), 559-578. https://doi.org/10.1016/j.compedu.2007.09.020

Rodríguez-Martínez, J. A., González-Calero, J. A., & Sáez-López, J. M. (2020). Computational thinking and mathematics using Scratch: an experiment with sixth-grade students. *Interactive Learning Environments*, *28*(3), 316-327. https://doi.org/10.1080/10494820.2019.1612448

Sáez-López, J. M., Román-González, M., & Vázquez-Cano, E. (2016). Visual programming languages integrated across the curriculum in elementary school: A Two year case study using "Scratch" in five schools. *Computers & Education, 97*, 129-141. https://doi.org/10.1016/j.compedu.2016.03.003

Seow, P., Looi, C. K., How, M. L., Wadhwa, B., & Wu, L. K. (2019). Educational policy and implementation of computational thinking and programming: Case study of Singapore. In *Computational thinking education* (pp. 345-361). Springer. https://doi.org/10.1007/978-981-13-6528-7

Shute, V., Sun, C., Asbell-Clarke, J. (2017). Demystifying computational thinking. *Educational Research Review, 22*, 142–158. https://doi.org/10.1016/j.edurev.2017.09.003

Song, O. J., Park, E. K., & Bae, J. M. (2020). The Effect of software education using Micro: bit on computational thinking of elementary school students. *Journal of Knowledge Information Technology and Systems (JKITS)*, *15*(1), 37-46.

Troiano, G. M., Snodgrass, S., Argımak, E., Robles, G., Smith, G., Cassidy, M., Tucker-Raymond, E., Puttick, G., & Harteveld, C. (2019). Is my game OK Dr. Scratch? Exploring programming and computational thinking development via metrics in student-designed serious games for STEM. In *Proceedings of the 18th ACM international conference on interaction design and children* (pp. 208-219). https://doi.org/10.1145/3311927.3323152

Voogt, J., Fisser, P., Good, J., Mishra, P., & Yadav, A. (2015). Computational thinking in compulsory education: Towards an agenda for research and practice. *Education and Information Technologies, 20*, 715–728. https://doi.org/10.1007/s10639-015-9412-6

Wing, J. M. (2006). Computational thinking. *Communications of the ACM, 49*(3), 33–35. https://doi.org/10.1145/1118178.1118215

Wu, S. Y., & Su, Y. S. (2021). Visual programming environments and computational thinking performance of fifth-and sixth-grade students. *Journal of Educational Computing Research*, *59*(6), 1075-1092. https://doi.org/10.1177/07356331209888

Zur-Bargury, I., Pârv, B., & Lanzberg, D. (2013). A Nationwide exam as a tool for improving a new curriculum. In *Proceedings of the 18th ACM conference on Innovation and technology in computer science education* (pp. 267-272). ACM Press. https://doi.org/10.1145/2462476.2462479

Cheng, Y.-H. (2023). Exploring the Effects of Tool-Assisted Paraphrasing Strategy Instruction on EFL Learners' Paraphrasing Performance. *Educational Technology & Society*, *26*(4), 51-68. https://doi.org/10.30191/ETS.202310_26(4).0004

# Exploring the Effects of Tool-Assisted Paraphrasing Strategy Instruction on EFL Learners' Paraphrasing Performance

**Ying-Hsueh Cheng**
National Chung Cheng University, Taiwan // folyhc@alum.ccu.edu.tw

(Submitted April 29, 2022; Revised February 25, 2023; Accepted March 6, 2023)

**ABSTRACT:** Many students who study English as a foreign language (EFL) often find it challenging to paraphrase while writing from source texts. Lacking such an ability can lead to different meanings as well as copying another person's ideas, words or work. However, little research has been done to integrate tool consultation to assist students in paraphrasing. To address this gap, this study explored whether guided tool consultation with paraphrasing strategy instruction can help students improve their overall paraphrasing performance. During an 18-week course, a class of students were trained to use three different e-tools to find synonyms: *Microsoft Word* thesaurus, *Oxford Living Dictionaries* synonyms (now called *Thesaurus.com*), and *Linggle*. Adopting a mixed-method approach, data sets included: pre-posttest drafts (summary writing), surveys, screen recordings, and interviews. The results showed significant differences between the pre- and post-tests. The majority of the lexical and phrasal paraphrases was suitable, while only a few were inaccurate. The students demonstrated the ability to consult the tools for changing synonyms and were able to apply taught strategies to restructure and restate the original sentences. Although students revealed different perceptions of the usefulness of the three tools for finding synonyms, they generally agreed that paraphrasing strategies combined with tool training were beneficial for learning. Pedagogical implications and research suggestions are provided based on the findings.

**Keywords:** EFL, Paraphrasing, Tool consultation, Second language writing, Instructional design

## 1. Introduction

Paraphrasing is one of the essential skills for successful academic writing (Shi, 2012). It is a cognitively demanding process since it involves both reading and writing skills when integrating source information into a new text (Hirvela & Du, 2013). Many English as a foreign language (EFL) writers often have difficulties using their own words to restate certain parts of the ideas from original texts (Li & Casanave, 2012). Their inappropriate use of source texts and evident instances of textual borrowing are frequently considered plagiarism by disciplinary professors and writing instructors (Pecorari & Shaw, 2012). To help these learners move beyond reliance on copying from source texts, many writing educators have provided pedagogical solutions such as explicit teaching, concept building, and paraphrasing strategies (e.g., using synonyms, changing the sentence structure) (Liu et al., 2016).

With the emergence of reference resources, a growing number of researchers have suggested that teachers provide learners with training on how to use these e-tools (Yoon, 2016). Research has shown the benefits of data-driven learning (DDL), "where learners examine naturally-occurring language and discover patterns on their own" (Boulton, 2010, p. 2). That is, through DDL, learners can utilize corpora or concordancers to check if a word is frequently used in a certain context by observing its patterns. However, some researchers have argued that DDL may be too demanding for EFL learners to understand the monolingual concordance lines (Mizumoto & Chujo, 2016). When learners write from source texts, they may turn to web resources (e.g., Google, dictionaries, or concordancers) to find synonymous words/phrases for replacement. Despite the increase in DDL studies, little research has been conducted to examine how EFL students consult concordancers when paraphrasing. To what extent DDL can be introduced to assist synonym finding during paraphrasing remains unclear. Addressing this gap, the current research aimed to introduce a variety of e-tools, including concordancers and non-concordancers, to assist EFL learners' synonym finding. Moreover, it integrated paraphrasing strategy instruction to help students improve their overall paraphrasing skills.

## 2. Literature review

### 2.1. EFL paraphrasing at the college level

Studies have reported that EFL students' difficulty in paraphrasing is mainly due to a lack of lexical and grammatical knowledge. Loh (2013) found that when Malaysian EFL students were asked to paraphrase two quotations, they made more linguistic errors (grammar, syntax, and lexis) than writing and paraphrasing conventions errors and semantic errors (content of message). Na et al. (2017) found that Vietnamese EFL students encountered difficulties understanding the source text, and had insufficient vocabulary to use when paraphrasing. These studies imply that there is room to help EFL writers tackle linguistic challenges while paraphrasing.

Factors regarding the paraphrasing difficulties among EFL students have also been discussed. First, students' paraphrasing ability may be related to their English proficiency levels. Liao and Tseng (2010) found that undergraduates copied more excerpts from source texts than graduate students did. Liou (2016) found that higher-proficiency students produced longer essays with better quality and advanced vocabulary. The students also showed better paraphrasing skills than the low-proficiency group. These studies suggest that the lower the students' proficiency levels are, the more likely they are to struggle when writing from source texts. Second, readability of source texts may influence EFL students' paraphrasing performance. Sun (2012) revealed that college students tended to do more substantial paraphrasing from the higher-readability texts than from the low-readability texts. This indicates that when providing paraphrasing training, teachers need to take readability of source texts into consideration. Last, limited instructional time and lack of explicit teaching may deter students from practicing paraphrasing (Marr, 2019). The general assumption is that students can learn how to paraphrase through online resources or from one-shot instructional practice. However, much has been underestimated in terms of the situations where the teachers are not there to help students. This is also the case in many educational contexts in Taiwan.

Some have suggested providing concept-building and strategy training since learning of paraphrasing requires a "host of interconnected subskills" (Hirvela & Du, 2013, p. 88). Developing an online writing tutorial system (DWright), Liu et al. (2018) found that Taiwanese EFL students' paraphrasing quality improved together with their plagiarism avoidance knowledge and citation abilities. Escudero et al. (2019) recommended teaching strategies such as learning synonyms, changing word order and word class, changing sentence structure, enhancing reading strategies, and developing lexical fluency to enhance students' paraphrasing skills. These studies indicate that step-by-step instruction is needed to help students cite the original source accurately and to reshape the original source using different grammar structures and vocabulary while maintaining the original meaning. Based on this, this study included the teaching of paraphrasing strategies prior to tool training.

### 2.2. e-tools for paraphrasing

Evaluation of automated paraphrasing tools and their paraphrasing quality has been discussed in the literature. Fitria (2021) suggested that automated paraphrasing tools (e.g., *QuillBot*) can be employed by EFL students to practice paraphrasing because they have various features such as allowing users to choose synonyms, change the word order, change the word form, and change between active and passive voice. However, Prentice and Kinden (2018) contended that these paraphrasing tools may not be suitable because they tend to rely on synonym substitution without changing the syntactical structure of the sentence and thus often produce inappropriate synonyms and incomprehensible texts. The application of word matching software such as *Turnitin* was also investigated. Kostka and Maliborska (2016) noted that the originality checker feature in *Turnitin* could identify whether the student only substituted synonyms without changing the structure of the original sentence. However, Rogerson and McCarthy (2017) cautioned that *Turnitin* only "detects some but not all cases of synonym replacement" (p. 6) and it also fails to identify the similarities between the original source materials and the output generated by automated paraphrasing tools. Prentice and Kinden (2018) compared the output of automated paraphrasing tools with those produced by *Google Translate* and found that paraphrasing tools tend to use inaccurate synonyms to replace accepted medical terminology, whereas *Google Translate* can preserve these terms intact. In addition to the aforementioned tools, Bailey and Withers (2018) investigated how L1 and L2 students utilized *Microsoft Word* default settings (i.e., spell check, grammar check, and synonym finder) and reported that synonym finder was the most frequently used tool during a paraphrasing task.

## 2.3. DDL for paraphrasing

In L2 writing research, learners' direct use of corpora or reference tools (i.e., concordancing) has been examined. This line of research is called *data-driven learning* (DDL). Some studies have shown that DDL can facilitate L2 writing due to its inductive nature and authentic linguistic examples that provide learners with opportunities to experience discovery learning (Lee & Lin, 2019). However, others have revealed that the abundant information presented in corpora/concordancers often make it difficult for learners to observe the patterns from the examples and to extract relevant information (Mueller & Jacobson, 2015). Thus far, DDL has been investigated with various learning purposes: collocations (Li, 2017), error correction (Cheng, 2021), paraphrasing (Han & Shin, 2017), and thesis writing (Crosthwaite et al., 2019). Of these, studies have shown positive effects on DDL for paraphrasing. For instance, Chen et al. (2015) developed a Chinese-English corpus-based paraphrasing system called *PREFER* (PREFabricated Expression Recognizer) and compared its effectiveness with a dictionary (*Longman English Dictionary Online*) and a thesaurus (*Theasarus.com*). Their findings showed that PREFER helped students make significant progress in the paraphrasing task, and most students felt satisfied with the paraphrases generated by PREFER. In another study, Han and Shin (2017) taught Korean EFL students how to use *Google* for paraphrasing with a focus on using quotation marks (" ") and a wildcard (*). After training, students found it easier to find synonyms, although they still had difficulty paraphrasing. These studies showed the potential of DDL for paraphrasing. They also indicated that a variety of e-tools, not only concordancers, can be introduced to learners so that they can conduct lexical searches and identify appropriate synonyms effectively while paraphrasing.

## 2.4. Research gaps

Based on the above, some gaps in the research can be found. First, empirical studies on tool consultation for paraphrasing are scant. Only a handful of studies were identified. However, they were conducted either as a one-shot task or within a short period of time. Except for Han and Shin's (2017) research, the studies did not provide step-by-step instructional procedures regarding using e-tools to find synonyms and then restating and restructuring the original sentences after tool consultation. Second, existing DDL studies tended to adopt only one tool to help students paraphrase. Few explored how different tools can complement each other and the fact that learners may have different learning styles during tool consultation (Cheng, 2021). For example, some learners might consider exploring the concordances too much work, while others favor DDL and tend to make more use of the corpus/concordancer (Bridle, 2019). The fact that DDL may not be suitable for all learners should be considered (Boulton, 2009).

To fill these gaps, this study aimed to integrate tool consultation with paraphrasing strategy instruction. In this paper, multiple e-tools were introduced to facilitate synonym finding in order to help learners with different needs to make effective lexical modifications. Also, paraphrasing strategies were provided to increase students' fluency of restating so that they could make syntactical changes successfully.

The rationales for choosing the e-tools include: (1) They must allow users to find synonyms with abundant search results; (2) one of them must be a concordancer while the others can be non-concordancers suggested by previous studies or based on learners' preferences; (3) they should be freely accessible and easy to use without requiring log-in. Based on these criteria, the current research introduced three different e-tools (i.e., a concordancer, a word processor, and a dictionary): (1) *Linggle (Synonyms)*, (2) *Microsoft Word* thesaurus, and (3) *Oxford Living Dictionaries (Synonyms)*. Specific reasons regarding why these tools were selected are explained as follows.

**Linggle (Synonyms)** (hereafter Linggle) was introduced because it contains naturally-occurring language data for learners who prefer DDL to explore concordance lines and discover patterns on their own. Linggle was developed by National Tsing Hua University in Taiwan (Boisson et al., 2013). It retrieves lexical bundles in response to a given query which can contain synonyms, keywords, wildcards, and wild parts of speech. By typing in a wave symbol (~) in front of a word, results with synonymous words will show up in a list for users to select. More crucially, it provides frequency counts of a word and allows multi-word input. These features were noted in prior studies (Lai & Chang, 2020; Zhu, 2015) and thus were considered useful for the purpose of this study.

**Microsoft Word Thesaurus** (hereafter MW) was adopted because students can simply click their mouse on the word whose synonyms they wish to search for and select the most suitable one by clicking on the drop-down menu. MW allows them to compare multiple options of synonyms without having to open a webpage for further searches. These features were noted by Bailey and Withers (2018) who found that students frequently consulted MW synonym finders in a paraphrasing task. Unlike Linggle, MW does not consist of concordances and nor does it provide words in context. It only presents a list of synonymous suggestions which can be considered convenient for learners who do not like DDL and prefer to view synonyms through a list.

**Oxford Living Dictionaries (Synonyms)** (hereafter OLD) was selected because learners can type not only a word but also a phrase to find its synonyms. Also, it offers definitions and sentences to help learners identify suitable synonyms. For learners who prefer consulting online dictionaries, OLD should be considered user-friendly. Similar to MW, OLD does not require users to observe the concordance lines as is required in Linggle. Differing from MW, OLD provides multiple example sentences which allow learners to differentiate how a certain word is used in different contexts, and thus helps them to identify suitable synonyms for substitution. (Note that OLD has undergone drastic change and has been called *Dictionary.com* since 2019. The current interface is different from the one used in this study. *Thesaurus.com* is the database used for finding synonyms.)

In sum, three tools were included in this study to allow learners to cross-check the search results. These tools have varying functions that are considered complementary and should be sufficient for students to query, compare, and identify appropriate synonyms for substitution and ultimately make syntactical changes during paraphrasing.

### 2.5. Research questions

The present study was guided by the following research questions:

- What are the effects of tool-assisted paraphrasing strategy instruction on EFL students' paraphrasing performance?
- How do the students use the three tools during paraphrasing?
- What are the students' perceptions of the tool-assisted paraphrasing strategy instruction?

## 3. Methodology

### 3.1. Settings and participants

This study was conducted in an intermediate-level EFL writing class, which is required for all English-major students studying at the university in southern Taiwan. Adopting a single-group design, an intact class of 21 students in their second year of study was invited to participate in this research. These students had taken two required writing courses in their first year of study. According to a pre-course survey, all the participants reported that they had web consultation experience during L2 writing but had little experience of tool consultation or concordancing for paraphrasing.

### 3.2. Instructional procedures

In this study, two stages of practices were included: (1) paraphrasing strategies, and (2) tool consultation. In the first stage, *paraphrasing strategies*, students were taught about the concepts of paraphrasing, summarizing, citing, and quoting. To help the students build these concepts, they were guided to do exercises in their textbook (Unit 3 of *Great Writing 5* by Folse & Pugh, 2015). The students learned to distinguish the best paraphrasing, those that were too similar, and those that had different meanings or wrong information. Then, how paraphrasing is different from summarizing was discussed. Moreover, handouts extracted from *Engaging Writing 2* (Fitzpatrick, 2011) were provided with the focus on seven paraphrasing strategies. These included: (1) changing the synonym, (2) changing the form of words, (3) changing the subject of the sentence, (4) changing the connectors, (5) changing parts of the sentence, (6) combining or dividing sentences, and (7) omitting unnecessary words. Exercises were offered for students to discuss and compare strategies used in different paraphrases. These seven strategies were later employed to guide the students in the tool training sessions so as to remind them to make not only lexical but also syntactical changes during paraphrasing. To assist students in practicing the first strategy (i.e., changing the synonym), three tools were introduced. See below for details.

During the second stage (i.e., *tool consultation*), three tool training sessions and a combined review session of the three tools were offered to teach students how to use the tools for finding synonyms. Each session took place for about 90 minutes in a computer lab. One trained assistant was present to provide immediate help. See Figures 1 to 3 for the three tools. The students were first guided to use Microsoft Word thesaurus (Figure 1), and then Oxford Living Dictionaries (Figure 2) for finding synonyms. Following these, they then experienced DDL with the use of Linggle (Synonyms) (Figure 3). When they utilized Linggle for finding synonyms, they were told to examine authentic language examples to identify appropriate synonyms.

*Figure 1*. Microsoft Word Thesaurus

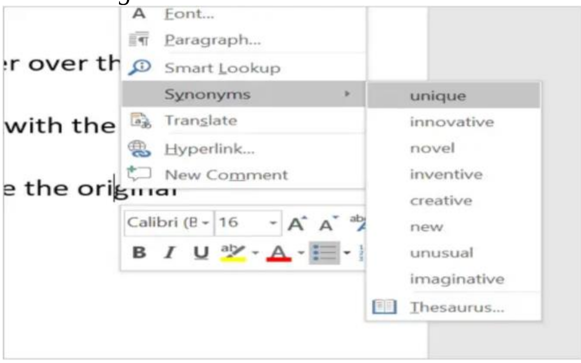



*Figure 2*. Oxford Living Dictionaries (synonyms)

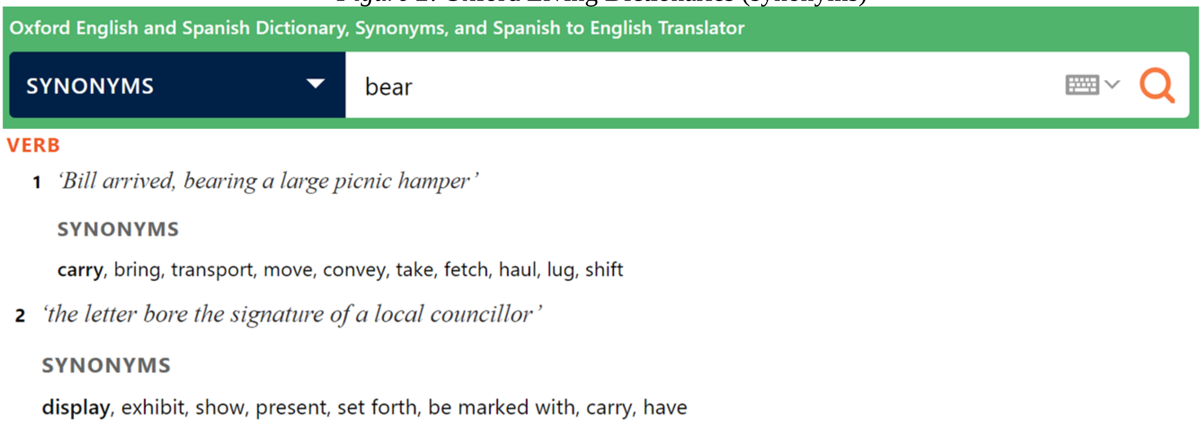



*Figure 3*. Linggle (synonyms)

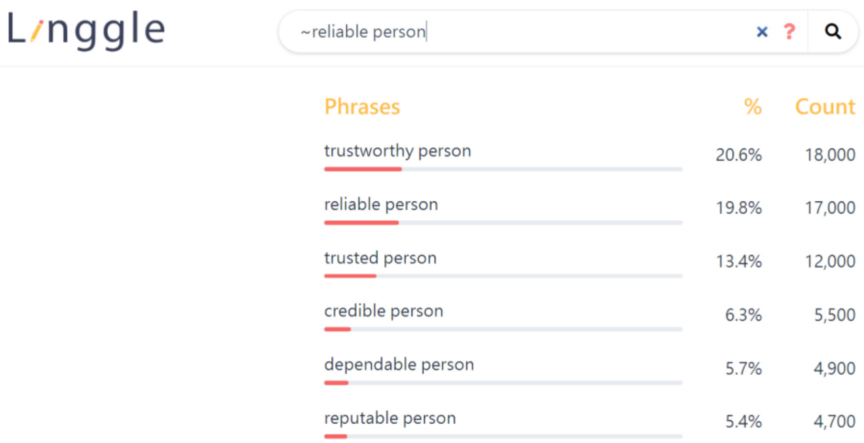



*Table 1*. Instructional design

| Weeks | Content | Tasks |
|---|---|---|
| 1~3 | Course orientation & Process essay | Essay writing |
| 4~6 | What is Paraphrasing and how? | Concept building |
| 7~9 | Introducing 7 paraphrasing strategies: Quoting, citing sources, summarizing | Paraphrasing strategies |
| 10 | Tool training 1: *Microsoft Word Thesaurus* (Worksheet 1) | Tool consultation for finding synonyms (students were reminded to practice using paraphrasing strategies) |
| 12 | Tool training 2: *Oxford Living Dictionaries-Synonyms* (Worksheet 2) | |
| 14 | Tool training 3: *Linggle-Synonyms* (Worksheet 3) | |
| 16 | Combined review of 3 tools (Worksheet 4) | |

Table 1 shows the content and tasks included in the 2-stage instructional procedures. Of note is that process essay writing was taught at the beginning of the semester. The students were asked to write about a hunting process of an animal. It was assumed that the learners would need to do some research on the Internet and some might do verbatim copying for composing their essay drafts. To help students increase their paraphrasing awareness, the course was designed as presented in Table 1.

*Figure 4*. Three-step instructions in worksheets 1 and 2

1. Decide what words you want to search for synonyms. Highlight the words you searched for synonyms.
2. Write down **equivalent synonyms** you found (one synonym for each highlighted word).
3. Restate and restructure the sentence. (When you change the sentence structure, try to use the 7 strategies[1] you have learned in the handout of *Engaging Writing 2*, pp. 119-201).

Example

Step 1

**Original sentence**: *Selling a product successfully in another country often requires changes in the product.*

Step 2

**Synonyms found**: *Exporting / merchandise / needs*

**Sentence re-structured**: Oftentimes, changes are needed when a company wants to succeed in exporting merchandise in another country.

Step 3

*Figure 5*. Four-step instructions in worksheets 3 and 4

1. Highlight the words you searched for synonyms. (You can decide what words you want to search for synonyms.)
2. Write down the suitable synonym for replacing the original word (one synonym for each highlighted word).
3. Restate and restructure the sentence. (When you change the sentence structure, try to use the 7 strategies[1] you have learned in the handout of *Engaging Writing 2*, pp. 119-201).
4. Include the author(s)' names and year of publication in the restructured sentence

**Example**

Step 1: Highlight the words you searched

It is quite obvious that healthy people are happier than unhealthy people. What is now becoming increasingly evident through study is that the reverse is also true: happy people are healthier than unhappy people. It appears that happiness, which simply means having happy thoughts most of the time, causes biochemical changes in the brain that in turn have profoundly beneficial effects on the body's physiology.
Chopra, Deepak. *Creating Health: How to Wake Up the Body's Intelligence*. Boston: Houghton Mifflin, 1987. (Original text from Engaging Writing, p. 208)

Step 2: Synonyms searched and found:

| Synonyms searched | Tool used | Synonyms found |
|---|---|---|
| it ~appears that | Linggle | reveals |
| happy ~thoughts | Linggle | thinking |
| have ~profoundly beneficial | Linggle | enormously |
| have profoundly ~beneficial | Linggle | positive |

Step 3: Restate and restructure the sentences:

*we all recognize that people with good health are happier than those whose health is not as good. Research is revealing that the opposite is true as well: Happy individuals enjoy better health than unhappy individuals. Happiness, or the quality of thinking positively most of the time, probably stimulates biochemical reactions in the brain that protect one's physical health.*

Step 4: Paraphrased texts and included author's name and year of publication:

*Deepak Chopra (1987) wrote that ….*

In total, four worksheets were created for the three tool training sessions and a combined review session. The development of the training worksheets was based on Han and Shin's (2017) directions for paraphrasing. In the first and second worksheets for MW and OLD, only single sentences were provided for practicing searching for synonyms when paraphrasing. From the third to the fourth tool training worksheets (i.e., Linggle and combined review of the three tools), however, paragraphs consisting of two to three sentences were provided to enhance students' fluency of rewriting.

In the first two training worksheets, three-step instructions were introduced: (1) highlight the words you searched for, (2) write down synonyms found, and (3) restate and restructure the sentence (see Figure 4). In the last two worksheets, the students not only had to follow the above three-step instructions, but also needed to include one more step, i.e., citing the name(s) of author(s) and year of publication of the paraphrased text (see Figure 5).

### 3.3. Data collection

To understand whether and how these three tools influenced the students' paraphrasing performance and their perceptions of using these tools, both quantitative and qualitative data were collected. These included: (1) pre- and post-test drafts, (2) surveys completed after the four tool training sessions, (3) screen recordings from the post-test, and (4) interviews conducted near the end of the course.

#### *3.3.1. Pre- and posttests*

In the pre- and posttests, summary writing tasks were administered. Summary writing was adopted for two reasons. First, a summary task requires students to focus on the main ideas discussed in one source text rather than including a personal opinion like writing an argument. Thus, it is more likely to "elicit paraphrases of specific excerpts of the source text than other writing tasks" (Keck, 2006, p. 264). Moreover, summary writing was chosen since it is close to the actual practice where EFL students read multiple source texts and paraphrase sentences when writing their assignments. Although students sometimes read Chinese sources, they tend to locate more English than Chinese source texts. In light of these considerations, summary writing was thus adopted.

Following the summary writing tasks described in Keck's study (2014), one source text published by *Time Magazine* in 2015 was selected (number of words: 667; title: *Can Brain Games Keep My Mind Young?*). That means, the same source text (an argumentative text) was used in the pre- and posttests to avoid differences among text length, topic familiarity, and writing outcomes (Zhang, 2013). Being mindful of the impact of text difficulty and topical familiarity on the production of paraphrasing (Sun, 2012), two methods were employed to ensure text readability: a readability checker (Flesch Reading Ease; score: 57.7) and consultation with one instructor from the same department. Based on the evaluation, the text was considered suitable for the participants in terms of its topic, content, and comprehensibility.

During the pre- and posttests, the students were given 45 minutes to read and write a one-paragraph summary (with a word limit of 200). The students were guided with these instructions in their pre- and post-tests: "Explain the most important main ideas (or arguments) of the essay in your own words" and "do not include your personal opinion" (Keck, 2014, p. 8). Of note is that during the posttest, participants were encouraged to use the three tools for finding synonyms: (1) *Microsoft WORD Thesaurus*, (2) *Oxford Living Dictionary (synonyms)*, and (3) *Linggle (~)*. To avoid students' familiarity with the same source text, the pre-test was provided in Week 1 while the post-test was carried out in Week 17. Moreover, no indication was made regarding the post-test to prevent learners from researching the source text online.

#### *3.3.2. Surveys*

A total of four online surveys were distributed in this study. After each tool training session, the students were required to fill out one survey. The first three surveys were related to the three respective tool training sessions, while the last survey was used to gather students' perceptions of the combined tool review session. The development of the first three surveys was based on Chen et al.'s (2015) study, whereas the last survey was newly created. In these surveys, both closed-ended and open-ended items were included. Chen et al.'s (2015) survey items were adapted because the researchers developed a paraphrase tool to improve Taiwanese EFL college students' writing skills, which was similar to the purpose of the current study. By doing so, it could help us understand Taiwanese EFL learners' perceptions of using tools for paraphrasing.

#### *3.3.3. Screen recordings*

The screen capture software called *scre.io* (https://scre.io/) was employed. It is easy to use and freely available. Users can choose to record online or download it for use. When the recording is completed, the user needs to download a video (.webm). In this study, before the posttest began, the students were instructed to do a test recording for a few minutes to learn to download their own video recordings.

#### *3.3.4. Interviews*

All students were invited to take part in a 15-minute interview after the completion of the post-test. The participants could use either Chinese or English during the interviews. The interviews were audio-recorded and later transcribed for analysis and comparison. The interview questions included students' perceptions of the tools, satisfaction with the training sessions, and experiences of using these tools for paraphrasing practice.

### 3.4. Data analysis

#### *3.4.1. Test data*

To examine whether students' paraphrasing performance was improved through tool consultation, the two summary writing drafts in the pre- and posttests were constantly compared to identify the changes and evaluate the quality of paraphrasing. To ensure high inter-coder reliability, two experienced EFL writing teachers were asked to score the students' summary writing drafts based on the scoring rubric developed by Yamanishi et al. (2019). This scoring rubric was chosen because it included two paraphrasing dimensions: *paraphrase (quantity),* and *paraphrase (quality).* It also contained three other dimensions for assessing summary writing: *content, language use*, and *overall*. According to Yamanishi et al. (2019), this five-dimensional rubric featured analytic and holistic assessments which were based on the teacher raters' evaluation and comparison with the rubric developed by the Educational Testing Service (ETS). This rubric was considered appropriate for evaluating EFL students' paraphrasing since it could be used to assess L2 summaries with a focus on paraphrasing skills. In terms of the scores, each dimension ranged from 1 to 4 points (*poor* to *very good*), with a total score of 20. For convenience of analysis, all the points were converted to 100. Specific descriptions were provided for each dimension (see the Appendix).

An orientation was offered for the two raters to ensure that they understood each specific description. After the rating process, disagreements were resolved through multiple discussions. To answer research question 1, descriptive statistics together with a paired sample *t* test were computed to understand whether the students made improvement in the posttest.

#### *3.4.2. Survey data*

Pre- and post-surveys were tabulated first. For the analysis of the Likert-type data, the mean scores and *t* test of the students' responses were computed relative to the 5-point scale. The internal reliability of the instrument was also checked using Cronbach's alpha. Other questions such as multiple-choice questions or open-ended questions were computed separately for comparisons of the Likert-type data.

#### *3.4.3. Screen recordings*

Since one of the purposes of the study was to understand the type of information the students looked for when using the tools, screen recordings of students' tool consultation processes were coded and triangulated with their summary writing drafts in the pre- and posttests.

#### *3.4.4. Interview data*

Once all the interviews were fully transcribed, notes and indexes were made in the margins. The interviews explored the reasons underlying students' perceptions of tool consultation and their self-assessments of the value of tool consultation in paraphrasing. Constant comparisons among the three tools were also made.

## 4. Findings and discussion

### 4.1. What are the effects of tool-assisted paraphrasing strategy instruction on EFL students' paraphrasing performance?

The participants' pre- and post-test of summary writing drafts were analyzed to answer this question. The overall average mean scores in the posttest ($M$ = 14.74) were higher than those in the pre-test ($M$ = 11.86). In other words, most of the students made progress in their paraphrasing abilities with an average of 2.88 gained in the post-test.

A paired sample $t$ test showed a significant difference between the pre- and post-test scores with the use of the three tools for paraphrasing. Table 2 shows the mean scores and standard deviation of students' writing performance in the pretest and posttest. (The minimum point value in each dimension is 1; the maximum point value is 4 (* $p < .05$, ** $p < .01$.) Results of average mean scores revealed that the students progressed in the posttest across the five dimensions of summary writing performance: content (+0.52), paraphrase (quantity) (+0.64), paraphrase (quality) (+0.83), language use (+0.48), and overall (+0.43). The results of the paired sample $t$ test were significant across these five dimensions: content ($t$ = -3.13, $p < .01$), paraphrase (quantity) ($t$ = -4.37, $p < .01$), paraphrase (quality) ($t$ = -5.80, $p < .01$), language use ($t$ = -4.48, $p < .01$), and overall ($t$ = -2.76, $p < .05$). In sum, 19 of the 21 students made progress in the post-test. Among them, S6 and S8 made the most progress with 7 points on average. This indicates that with the use of the three e-tools plus strategy instruction, students made significant progress in paraphrasing. These findings lend support to the assertion that tool consultation positively affected students' searches for synonyms and increased their paraphrasing performance, which were also reported in prior studies (Bailey & Withers, 2018; Chen et al., 2015). Moreover, the results showed that teaching paraphrasing strategies such as using synonyms, changing word order, and changing sentence structure could enhance students' paraphrasing performance as was noted by Escudero et al. (2019). Altogether, the findings highlight the possibility of offering guided tool consultation with strategy instruction for effective paraphrasing. By consulting e-tools for synonym searches, students can take the first step to make lexical substitutions rather than relying on copying from the source text. Through strategy instruction, they can be guided to practice a "host of interconnected subskills" (Hirvela & Du, 2013, p. 88) so as to increase their overall paraphrasing ability.

*Table 2*. Pre- and post-test results of summary writing ($N$ = 21)

| Dimensions | Test | Mean | *SD* | *df* | *t* | Sig. (2-tailed) |
|---|---|---|---|---|---|---|
| Content | Pre-test | 2.40 | 0.63 | 20 | -3.13 | .005** |
| | Post-test | 2.92 | 0.58 | 20 | | |
| Paraphrase (Quantity) | Pre-test | 2.50 | 0.71 | 20 | -4.37 | .000** |
| | Post-test | 3.14 | 0.51 | 20 | | |
| Paraphrase (Quality) | Pre-test | 2.36 | 0.83 | 20 | -5.80 | .000** |
| | Post-test | 3.19 | 0.64 | 20 | | |
| Language Use | Pre-test | 2.40 | 0.61 | 20 | -4.48 | .000** |
| | Post-test | 2.88 | 0.43 | 20 | | |
| Overall | Pre-test | 2.19 | 0.61 | 20 | -2.76 | .012* |
| | Post-test | 2.62 | 0.53 | 20 | | |

*Note.* $^{*}p < .05$; $^{**}p < .01$.

To gain further information, S8, Cathy (pseudonym), was chosen as a representative case based on her scores, frequency of consulting the tools, and paraphrasing outcomes. As shown in Table 3, her scores increased in all five dimensions, with the highest scores in the dimension of content, followed by the dimension of paraphrase (quantity).

*Table 3*. Cathy's scores in the pretest and posttest

| Dimensions | Content | Paraphrase (quantity) | Paraphrase (quality) | Language use | Overall | Total scores |
|---|---|---|---|---|---|---|
| Pretest | 2 | 1 | 2 | 2 | 2 | 9 |
| Posttest | 4 | 3 | 3 | 3 | 3 | 16 |

As can be seen in Figure 6 as shown below, Cathy demonstrated improved paraphrasing skills in her posttest with longer text and better paraphrasing quality. Analysis of screen recordings also revealed that Cathy revisited the original source text, highlighted some words, pasted chunks of source text information, and then changed the original sentences. Cathy consulted all the three introduced tools (a total of 14 times), and mainly used MW for most of the searches (13 times). She was able to search and replace most of the selected words with synonyms. She substituted 11 words based on MW search results (see the blue highlights in Figure 6). She made changes not only at the lexical level but also at the syntactical level. She adopted strategies taught in class for making syntactical changes. As shown in the second example in Table 4, the original sentence in the source text, "*There's just no solid **evidence***" was changed into "*no solid **proof** supports that brain games can…*" (Strategy: changing the subject of the sentence). In instance No. 6, she combined two sentences and produced a new paraphrased sentence without changing the original meaning. The original sentence, "*…a product that helps people **improve** cognitive abilities…online-based brain training can **improve** thinking*" was changed to "*brain games **advance** people's mental abilities and thinking*" (Strategies: sentence combining; omitting unnecessary words). Taken together, from these examples, we could learn that Cathy was able to identify appropriate synonyms based on tool consultation, apply the strategies to restructure the sentence, and produce a comprehensible text different from the original text.

*Figure 6*. Cathy's drafts in pre- and post-tests
(Original source text: https://time.com/3706689/can-brain-games-keep-my-mind-young/)

| Pretest (95 words) |
|---|
| Companies of brain games try to let people believe brain games keep our aging brain nimble. However, there is no solid evidence can show the benefit of brain games. Playing games over and over again let us get better at it, but it doesn't mean our brain becomes more capable. We are just good at what we practicing many times. Michael Scanlon, brain-game designer disagrees the opinion. He thinks brain game products help people improve cognitive abilities and thinking. Many scientists claim that the brain game company exaggerated and misleading the benefit of brain game |

| Posttest (188 words) |
|---|
| Brain game companies advocate that brain games can improve our memories and slow the declining of our mental functions. But Randall W. Engle says that there's just no solid proof supports that brain games can improving aspects of mental capability. Research has shown that brain games lack of "transfer." Therefore, People get better at the game only because of repeating it several times. Ursula Staudinger, director of the Butler Columbia Aging Center at Columbia University says brain game players' brain become more skilled of practicing just the tasks." However, Michael Scanlon, chief scientific officer at Lumosity, a large brain-game company, says brain games advance people's mental abilities and thinking. Over 70 important brain scientists and psychologists argued brain games' benefit about improvement of cognition were exaggerated and misleading. But the advantage of entertainment of brain games was admitted. In 2013, there were 50 million brain game players. Opportunity cost of playing brain games is the topic discussed by most scientists, because factors like healthy diet, regular meditation, and learning new things have been proved by research that they also can help aging brains instead of playing brain games. |

*Table 4*. Cathy's MW consultation records in the posttest

| | Original sentence | Paraphrased sentence in the posttest |
|---|---|---|
| 1. | ***sharpen*** your memory and slow the inexorable decline of your mental functions. | brain games ***improve*** our memories |
| 2. | There's just no solid ***evidence***. | no solid ***proof*** supports that brain games can… |
| 3. | …sharpen your memory and slow the inexorable decline of your mental ***functions*** | brain games can improv[e] aspects of mental ***capability*** |
| 4. | "It has become more ***capable*** of doing exactly the tasks it was practicing." | brain game players' brain become more ***skilled*** of practicing just the tasks |
| 5. | It has become more capable of doing ***exactly*** the tasks it was practicing. | brain game players' brain become more skilled of practicing ***just*** the tasks |
| 6. | …a product that helps people ***improve*** cognitive abilities…online-based brain training can ***improve*** thinking. | brain games ***advance*** people's mental abilities and thinking |
| 7. | …a product that helps people improve ***cognitive*** abilities…online-based brain training can improve thinking. | brain games advance people's ***mental*** abilities and thinking |
| 8. | ***More than*** 70 prominent brain scientists | ***Over*** 70 important brain scientists |
| 9. | ***prominent*** brain scientists | Over 70 ***important*** brain scientists |
| 10. | brain games do have the ***benefit*** of being fun | the ***advantage*** of entertainment of brain games was admitted |
| 11. | The ***issue*** most scientists have with people playing the games frequently is the opportunity cost | Opportunity cost of playing brain games is the ***topic*** discussed by most scientists |

An analysis of Cathy's paraphrasing outcomes revealed her recursive practice of reading, searching, and revising which demonstrated her ability to apply paraphrasing strategies with tool consultation. Although some changes remained at the lexical level by substituting synonyms, Cathy attempted to make syntactical changes by adopting strategies such as combining sentences, changing the subject, and omitting unnecessary words. This finding indicates that by providing strategy-based tool consultation training, students can learn to paraphrase not only at the lexical level but also at the syntactical level, which has rarely been discussed in previous studies related to tool consultation for paraphrasing.

### 4.2 How do the students use the three tools while paraphrasing?

To answer this research question, four sets of data were analyzed: the students' posttest summary writing drafts, screen recordings, surveys, and interviews. Analysis of the screen recordings showed that the total number of times the students used the three tools was: MW ($N$ = 72), OLD ($N$ = 60), and Linggle ($N$ = 13). This indicates that most students tended to use MW and OLD, while few used Linggle for finding synonyms. Moreover, it was found that the majority ($N$ = 15) used more than one tool to search for the synonyms they found. The students seemed to start with one tool to search for synonyms and if they could not locate suitable synonyms, they would continue to use other tools for better search results.

Regarding the use of MW thesaurus, many students noted in the posttest survey and the interviews that MW was convenient because they did not have to open a webpage to find synonyms. They could simply search for synonyms by clicking on words in the source text. Four students also mentioned that they used the built-in thesaurus on MW to find more synonyms for replacement since it provides more options for them to choose. As for those who used OLD frequently, they indicated that OLD offers synonyms based on different meanings so that they were able to select the synonyms they wanted. Although few students used Linggle in the posttest, three students commented that Linggle was easy to use since it provides frequency of word usage and its interface is simple and clean, making it efficient to scroll up and down for the search results.

Among the total number of times ($N$ = 145) the students used the tools for searching for synonyms, the majority of the lexical and phrasal paraphrases were suitable ($N$ = 135) with only 10 examples identified as unsuitable and inaccurate. This could mean that the participants were fairly discriminating when choosing synonyms. Table 5 presents successful examples from participants' drafts using the three tools.

*Table 5*. Successful examples of using the three tools

| Tool used | Frequency | Original word | Replaced word |
|---|---|---|---|
| MW | **67** out of 72 (93%) | **sharpen** your memory | improve |
| | | **benefits** of brain games | advantages |
| | | **concrete** proof | tangible |
| | | **improve** thinking | advance |
| | | a **worthwhile** way | valuable |
| | | the industry has **grown** | expanded |
| | | cognitive **ability** | capability |
| | | our **intention** was to… | purpose |
| | | convince them of their **merit** | value |
| OLD | **58** out of 60 (98%) | were frequently **exaggerated** | overstated |
| | | **prominent** brain scientists | leading |
| | | the primary **investigator** | researcher |
| | | the **primary** investigator | chief |
| | | no **solid** evidence | dependable |
| | | **improve** cognitive abilities | enhance |
| | | the **issue** most scientists have… | matter |
| | | **argued** that claims on … | stated |
| | | claims … were **misleading** | deceptive |
| Linggle | **10** out of 13 (77%) | **sharpen** your memory | improve |
| | | no **solid** evidence | strong |

While the above showed suitable replacement through the use of the three tools, results of unsuitable replacements ($N$ = 10) are shown below (see Table 6).

*Table 6*. Unsuccessful examples with the three tools

| Tool used | Frequency | Original word | Unsuitable replacement |
|---|---|---|---|
| MW | **5** out of 72 (6.9%) | Brain games can only make people …because they lack **transfer**. | transmission |
| | | for people of any **age** | oldness |
| | | unstoppable **decline** of your mental functions will be much slower | failure |
| | | **sharpen** your memory | refine |
| | | According to a **report** from… | convey |
| OLD | **2** out of 60 (3.3%) | slow the **inexorable** decline of your mental functions | relentless |
| | | in a way that **necessarily** slows aging | automatically |
| Linggle | **3** out of 13 (23.1%) | **slow** aging | accelerate |
| | | **cognitive** functions | theological |
| | | **More than** 70 brain scientists… | Better than |

It was found that some words provided by MW seemed unable to ensure good quality of paraphrasing ($N$ = 5). For example, *transmission* for *transfer*, *oldness* for *age*, and *failure* for *decline*. Moreover, *refine* was not suitable for replacing *sharpen*, and *convey* (as a verb) is not an accurate replacement for *report* (as a noun). Noticing these errors, the participants mentioned in the interviews that although MW provided numerous synonyms, they often struggled to identify appropriate results. Similarly, in terms of the use of OLD, we also observed two instances where synonym selection appeared to be undiscriminating and unnatural (i.e., *relentless* for *inexorable*; *automatically* for *necessarily*). The difference in the meanings of these words is subtle, and so it could be difficult for participants to differentiate. Additional efforts were needed by either consulting other tools or relying on personal knowledge to avoid unsuitable replacement. Regarding the use of Linggle, three search results seemed misleading (i.e., *accelerate* for *slow*; *theological* for *cognitive*; *better than* for *more than*), causing the participants to choose the wrong words. In one instance, S10 failed to select a suitable word due to the misleading results provided by Linggle. She began with OLD and learned the synonymous word, *discernment*. She later searched *discernment* with Linggle, but ended up selecting *theological*, which was listed as the top search result. Mistakenly, she replaced the phrase, *cognitive functions* with *theological functions*. Altogether, these instances indicate that learners need to discriminate between suitable and unsuitable synonyms so that the replacement can fit the context of the sentences.

To answer the second research question, the above findings revealed that a few of the synonymous replacements were inaccurate and unsuitable, indicating the issue of synonym context when students explore with the tools (Bailey & Withers, 2018). Although these inappropriate collocations only occurred for 10 out of 145 synonyms, it may be useful for students to explore the unsuccessful examples in Table 6 and discuss why they do not work as appropriate synonyms. Unlike native English speakers, EFL students are often not able to discard inappropriate suggestions automatically based on intuition. Through finding and comparing synonyms with different tools, they can learn to be selective and critical.

Moreover, analysis of screen recordings showed that MW was used most frequently, followed by OLD, and then Linggle. MW was commonly used because the participants could simply right-click on a word in the source text and did not have to open a new webpage to find synonyms. This was also found in the study conducted by Bailey and Withers (2018). In their study, only one student used the MW built-in thesaurus, while the rest of the students tended to right-click on the word to access the synonym finder. In the current study, however, five students used the built-in thesaurus when they could not retrieve suitable results after the right-click, while the rest of the participants only accessed the synonym finders with the right-click. This indicates individual differences regarding the functions they preferred to use as well as how they searched, consistent with previous studies (e.g., Cheng, 2021). Bridle (2019) investigated corpus use and learner types, and found that some learners ("reflectors") might consider exploring the concordances too much work, while others ("pragmatists") tended to make more use of the tool. Liou (2019) indicated that learners' writing development may be influenced by individual and contextual factors. If writing instructors can take these into consideration, students' motivation for tool consultation can be increased and their anxiety can be reduced.

The synonym function of OLD was considered helpful since it provides example sentences for users to locate the most appropriate synonyms close to the context. EFL students' preferences for online dictionaries were also discussed (Lai & Chen, 2015). In this study, although most students acknowledged the benefits of the OLD synonym finder, some mentioned that OLD did not provide bilingual definitions and thus they needed to spend time browsing the example sentences. These students also noted that they would turn to Cambridge Dictionary since it offers both English-Chinese and Chinese-English definitions.

Linggle was used least frequently ($N$ = 13) among the three tools. In this study, some participants noted that Linggle was useful since it provides percentage of usage and it allows users to search for synonyms in multiple-word phrases (e.g., *it ~appears that*; *have ~profoundly beneficial*). As a concordancer, Linggle provides naturally-occurring language which may benefit learners who favor DDL. However, in this study, the participants who used Linggle did not indicate difficulties of exploring monolingual concordances. Instead, they noted that they found it inconvenient to add the wave (~) symbol in the search queries, and indicated that the search results were oftentimes either irrelevant (e.g., *theological* for *cognitive*) or unsuitable (e.g., *accelerate* for *slow*) which thus discouraged them from using it. Based on this result, it is difficult to conclude whether DDL should be recommended for teaching paraphrasing. Although Linggle was less consulted by the students, it could provide synonym searches just like the other two tools could offer. To make Linggle a useful tool, perhaps more guidance is needed so that more relevant search results can be generated.

### 4.3. Perceptions of the tool-assisted paraphrasing strategy instruction

Overall, the students felt positive about the course design which combined tool consultation and paraphrasing strategies for assisting paraphrasing. See Table 7. (Note for the item response scale: 1: strongly disagree; 2: disagree; 3: somewhat disagree; 4: agree; 5: strongly agree.)

*Table 7*. Survey responses regarding the course design

| Item | Mean | *SD* |
|---|---|---|
| 1. After receiving the paraphrasing tool training, I feel more confident that I can search for synonyms to replace the words/phrases in the source text. | 4.1 | 1.02 |
| 2. The paraphrasing strategies combined with the 3-tool training increased my paraphrasing ability. | 4.0 | 0.97 |
| 3. In the future, when I paraphrase, I will apply the paraphrasing strategies and the 3 tools introduced in this course to vary my expressions in writing. | 4.1 | 0.99 |

*Table 8*. Tool training 1 survey results (MW)

| Items | Mean | *SD* |
|---|---|---|
| 1. The training helped me to find synonyms and vary my expressions. | 4.3 | 0.50 |
| 2. The training time was enough to learn "Microsoft WORD Thesaurus." | 4.5 | 0.71 |
| 3. I will try to use "Microsoft WORD Thesaurus" to vary my expressions in the future. | 4.6 | 0.57 |
| 4. Overall, I learned a lot from today's training. | 4.6 | 0.60 |

*Table 9*. Tool training 2 survey results (OLD)

| Items | Mean | *SD* |
|---|---|---|
| 1. I can use "Oxford Living Dictionary-Synonyms" to find synonyms and vary expressions. | 4.3 | 0.61 |
| 2. The training time was enough to learn "Oxford Living Dictionary-Synonyms." | 4.1 | 0.80 |
| 3. I will try to use "Oxford Living Dictionary-Synonyms" to vary my expressions in the future. | 4.3 | 0.85 |
| 4. Overall, I learned a lot from today's training. | 4.3 | 0.60 |

*Table 10*. Tool training 3 survey results (Linggle)

| Items | Mean | *SD* |
|---|---|---|
| 1. I can use "Linggle" to find synonyms and vary expressions. | 3.2 | 1.26 |
| 2. The training time was enough to learn "Linggle." | 3.8 | 0.94 |
| 3. I will try to use "Linggle" to vary my expressions in the future. | 3.1 | 1.22 |
| 4. Overall, I learned a lot from today's training. | 3.6 | 0.96 |

The results also showed that more students seemed to favor MW and OLD, while some were uncertain about the potential of using Linggle for finding synonyms. See Tables 8 to 10. (Note for the item response scale: 1: strongly disagree; 2: disagree; 3: somewhat disagree; 4: agree; 5: strongly agree.). In spite of that, most students agreed that the tool training sessions were helpful and the training time was sufficient for them to learn how to use these tools for finding synonyms. Many reported that they would try to use these tools to vary their expressions in the future.

To gain insights into students' perceptions of the three tools, their responses to the open-ended survey items were further analyzed. The students mentioned how they could benefit from using these tools for synonym searches because of the distinct features and functions provided by these tools. They also indicated drawbacks which hindered them from effective consultation. Despite that, most students considered it beneficial to learn to use these tools for finding synonyms so that they could have options when they needed to locate more suitable search results. See Table 11.

*Table 11*. Perceived benefits and drawbacks of the three tools

| Tool | Benefits | Drawbacks |
|---|---|---|
| MW | Microsoft WORD thesaurus function is the most useful tool for me because it can be searched on the Word page directly, so I don't have to switch to other web pages. And most of the vocabulary I looked for so far was useful and correct. (S4)<br>It's the easiest one to operate. (S8)<br>Microsoft WORD synonyms is built in WORD so I don't have to open a new page. (S15) | Too many options; don't know how to choose the right synonyms. (S1)<br>Because Microsoft thesaurus can only provide the words but no examples, we need to think about whether we can use it to replace the original one or not. (S5) |
| OLD | The synonyms are more accurate. (S16)<br>Oxford shows not only the synonyms but also the definition or explanation, and more importantly, it provides us with some examples so that we can clearly understand whether that's the word we're looking for. (S5)<br>OLD provides different synonyms of different meanings. (S15) | Searching on OLD was a little bit distracting today. The server might not be able to serve so many people at one time. (S3) |
| Linggle | The steps are simple. (S2)<br>Linggle provides some examples. (S12)<br>It shows the number of usage of the word, which makes the word more reliable. (S6) | I couldn't find particular words on Linggle. (S3) |

In response to the third research question, the majority agreed that tool consultation combined with paraphrasing strategies could enhance their paraphrasing awareness and ability. They noted that with guided tool training, they were able to use the three tools to find synonyms for word/phrase replacement. More crucially, the paraphrasing strategies helped them to move beyond lexical substitutions, allowing them to re-order or restate with syntactic changes. Irrespective of students' frequency of consultation and preferences, it can be concluded that by offering students more than one tool, they can be motivated to search for and compare synonyms in order to vary expressions. This finding can be added to L2/EFL paraphrasing studies regarding combining paraphrasing strategy instruction with the facilitation of e-tools like the ones provided in this research. Moreover, the result can contribute to the growing DDL literature concerning the potential of teaching both concordancers and non-concordancers to EFL learners in order to meet their needs and preferences so as to encourage their consultation (Bailey & Withers, 2018; Boulton, 2009; Han & Shin, 2017; Yoon, 2016).

## 5. Conclusion

This study examined the effects of integrating tool consultation with paraphrasing strategy instruction for enhancing EFL students' paraphrasing performance. Three tools were provided: (1) *Microsoft Word thesaurus,* (2) *Oxford Living Dictionaries (Synonyms),* and (3) *Linggle (Synonyms)*. The results showed that the students made significant progress in the posttest when tool consultation was integrated with the paraphrasing strategy instruction. The majority of the lexical and phrasal paraphrases was suitable while only a few were inaccurate. Although students regarded *Microsoft Word* thesaurus and *Oxford Living Dictionaries* more useful for finding synonyms than *Linggle,* they held positive attitudes toward using the three tools for practicing paraphrasing. As was presented, the case study student, Cathy, and several students in this study demonstrated the ability to consult the three tools for finding synonyms, apply paraphrasing strategies, and make lexico-grammatical changes to restructure and restate the original sentences. These findings can be added to the growing body of literature on data-driven learning with the focus on a structured, step-by-step tool consultation training as a scaffolded approach, which encourages students to make use of the tools at their disposal while paraphrasing. Moreover, considering the fact that DDL may not be suitable for all learners, this study showed the possibility of including both concordancers and non-concordancers to give students options in terms of tool selection rather than forcing them to experience DDL. For learners with lower-level English proficiency, introducing more than one tool for finding synonyms is important since it can develop their ability to evaluate the appropriateness of word usage and justify the word context. With guided tool consultation plus strategy instruction, students can learn to make both lexical and syntactical changes. The results of this study can contribute to the literature by highlighting tool consultation combined with paraphrasing strategies as learning processes to raise students' awareness of plagiarism and to avoid patchwriting.

## 6. Limitations and implications

This study demonstrated how strategy-based tool consultation could help students improve paraphrasing. However, there are some limitations. First, based on a single-group design, this study did not include a control group. The findings might have been different if a control group could have been included. Second, this study only examined the effects of the three e-tools for paraphrasing. It did not look into the process regarding how students evaluate results and identify the most suitable synonyms for replacement. Differentiating nuances such as how synonymous one word is with another may be challenging. Therefore, future studies can explore learners' tool usage, including evaluating results and identifying suitable synonyms through using the think-aloud approach. Finally, the duration of this study was just one semester. Since learners' preferences may vary, it is suggested that a longer period of time should be allowed in future research.

Several pedagogical implications can be drawn from this study. First, the three tools introduced in this study seemed to benefit the students differently. Teachers can choose the tools according to students' learning needs and proficiency levels and offer training and opportunities for discussion regarding the difficulties they encounter. Second, to facilitate the EFL students' paraphrasing, we developed step-by-step strategy instruction with tool consultation for paraphrasing. Writing teachers can engage students in this practice to raise students' lexical and semantic awareness, monitor their paraphrasing process, and improve their paraphrasing quality. Noteworthy was that searching for synonyms constantly may influence the paraphrasing process. Students may need to focus on higher-order thinking skills (e.g., structuring main ideas in their own words) and reduce their frequency of searching to avoid interruption. Third, some technical terms may not be found with the three e-tools. In this study, screen recordings revealed that some participants searched for technical terms such as *aerobic* but failed to find suitable synonyms with the use of the three tools. It is suggested that during training, teachers can inform students that technical terms like this may not be found and often do not need to be replaced. Last, before the tools are introduced, they should be tested to avoid any unexpected situations. In this study, during the tool training, OLD could not be accessed in the lab and the students could only use their cell phones to consult this dictionary for finding synonyms. The problem was also noted by two students when they used OLD at home. This is worth noting so that more effective instructional guidance can be offered in the classroom settings.

## Acknowledgment

I would like to thank the reviewers for their detailed feedback, which improved the quality of this paper. This study was funded by the MOST 109-2410-H-194-098.

## References

Bailey, C., & Withers, J. (2018). What can screen capture reveal about students' use of software tools when undertaking a paraphrasing task? *Journal of Academic Writing, 8*(2), 176-190.

Boisson, J., Kao, T. H., Wu, J. C., Yen, T. H., & Chang, J. S. (2013). Linggle: A Web-scale linguistic search engine for words in context. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics: System Demonstrations* (pp. 139-144). Association for Computational Linguistics.

Boulton, A. (2009). Corpora for all? Learning styles and data-driven learning. In M. Mahlberg, V. Gonzalez-Díaz, & C. Smith (Eds.), *Proceedings of 5th corpus linguistics conference*. http://ucrel.lancs.ac.uk/publications/cl2009/150_FullPaper.doc

Boulton, A. (2010). Data-driven learning: Taking the computer out of the equation. *Language Learning, 60*(3), 534-572.

Bridle, M. (2019). Learner use of a corpus as a reference tool in error correction: Factors influencing consultation and success. *Journal of English for Academic Purposes, 37*, 52-69.

Chen, M. H., Huang, S. T., Chang, J. S., & Liou, H. C. (2015). Developing a corpus-based paraphrase tool to improve EFL learners' writing skills. *Computer Assisted Language Learning*, *28*(1), 22-40.

Cheng, Y.-H. (2021). EFL college students' concordancing for error correction. *English Teaching & Learning*, *45*(4), 431-460.

Crosthwaite, P., Wong, L. L., & Cheung, J. (2019). Characterising postgraduate students' corpus query and usage patterns for disciplinary data-driven learning. *ReCALL, 31*(3), 255–275.

Escudero, I., Fuertes, N., & López, L. (2019). Paraphrasing strategy in EFL Ecuadorian B1 students and implications on reading comprehension. *English Language Teaching*, *12*(1), 56-66.

Fitria, T. N. (2021). QuillBot as an online tool: Students' alternative in paraphrasing and rewriting of English writing. *Englisia: Journal of Language, Education, and Humanities*, *9*(1), 183-196.

Fitzpatrick, M. (2011). *Engaging writing 2: Essential skills for academic writing*. Pearson Longman.

Folse, K. S., & Pugh, T. (2015). *Great writing 5: From great essays to research* (3rd ed.). Cengage.

Han, S., & Shin, J. A. (2017). Teaching Google search techniques in an L2 academic writing context. *Language Learning & Technology*, *21*(3), 172-194.

Hirvela, A., & Du, Q. (2013). "Why am I paraphrasing?": Undergraduate ESL writers' engagement with source-based academic writing and reading. *Journal of English for Academic Purposes*, *12*(2), 87-98.

Keck, C. (2006). The Use of paraphrase in summary writing: A Comparison of L1 and L2 writers. *Journal of Second Language Writing*, *15*(4), 261-278.

Keck, C. (2014). Copying, paraphrasing, and academic writing development: A Re-examination of L1 and L2 summarization practices. *Journal of Second Language Writing*, *25*, 4-22.

Kostka, I., & Maliborska, V. (2016). Using Turnitin to provide feedback on L2 writers' texts. *TESL-EJ*, *20*(2), 1-22.

Lai, S. L., & Chang, J. S. (2020). Toward a pattern-based referencing tool: Learner interactions and perceptions. *ReCALL*, *32*(3), 272-290.

Lai, S. L., & Chen, H. J. H. (2015). Dictionaries vs concordancers: Actual practice of the two different tools in EFL writing. *Computer Assisted Language Learning*, *28*(4), 341-363.

Lee, P., & Lin, H. (2019). The Effect of the inductive and deductive data-driven learning (DDL) on vocabulary acquisition and retention. *System*, *81*, 14-25.

Li, S. (2017). Using corpora to develop learners' collocational competence. *Language Learning & Technology, 21*(3), 153–171.

Li, Y., & Casanave, C. P. (2012). Two first-year students' strategies for writing from sources: Patchwriting or plagiarism? *Journal of Second Language Writing*, *21*, 165–180.

Liao, M. T., & Tseng, C. Y. (2010). Students' behaviors and views of paraphrasing and inappropriate textual borrowing in an EFL academic setting. *Journal of Pan-Pacific Association of Applied Linguistics*, *14*(2), 187-211.

Liou, H. C. (2016). Exploring source use by different EFL students: Text quality, lexical profiles, and citation features. *Asian Journal of English Language Teaching*, *26*, 55-84.

Liou, H. C. (2019). Learner concordancing for EFL college writing accuracy. *English Teaching & Learning*, *43*(2), 165-188.

Liu, G. Z., Lin, V., Kou, X., & Wang, H. Y. (2016). Best practices in L2 English source use pedagogy: A Thematic review and synthesis of empirical studies. *Educational Research Review*, *19*, 36-57.

Liu, G. Z., Lu, H. C., Lin, V., & Hsu, W. C. (2018). Cultivating undergraduates' plagiarism avoidance knowledge and skills with an online tutorial system. *Journal of Computer Assisted Learning*, *34*(2), 150-161.

Loh, Y. L. (2013). Errors in paraphrasing and strategies in overcoming them. *Journal of Creative Practices in Language Learning and Teaching*, *1*(1), 4-17.

Marr, J. W. (2019). Making the mechanics of paraphrasing more explicit through Grammatical Metaphor. *Journal of English for Academic Purposes*, *42*, 1-7.

Mizumoto, A., & Chujo, K. (2016). Who is data-driven learning for? Challenging the monolithic view of its relationship with learning styles. *System*, *61*, 55-64.

Mueller, C. M., & Jacobsen, N. D. (2015). A Comparison of the effectiveness of EFL students' use of dictionaries and an online corpus for the enhancement of revision skills. *ReCALL*, *28*(1), 3–21.

Na, C. D., Nhat, C. M., & Nguyen, X. (2017). Paraphrasing in academic writing: A Case study of Vietnamese learners of English. *Language Education in Asia, 8*(1), 9-24.

Pecorari, D., & Shaw, P. (2012). Types of student intertextuality and faculty attitudes. *Journal of Second Language Writing*, *21*(2), 149-164.

Prentice, F. M., & Kinden, C. E. (2018). Paraphrasing tools, language translation tools and plagiarism: an exploratory study. *International Journal for Educational Integrity*, *14*, 1-16.

Rogerson, A. M., & McCarthy, G. (2017). Using Internet based paraphrasing tools: Original work, patchwriting or facilitated plagiarism? *International Journal for Educational Integrity*, *13*, 1-15.

Shi, L. (2012). Rewriting and paraphrasing source texts in second language writing. *Journal of Second Language Writing*, *21*(2), 134-148.

Sun, Y. C. (2012). Does text readability matter? A Study of paraphrasing and plagiarism in English as a Foreign Language writing context. *Asia-Pacific Education Researcher (De La Salle University Manila)*, *21*(2), 296-306.

Yamanishi, H., Ono, M., & Hijikata, Y. (2019). Developing a scoring rubric for L2 summary writing: A Hybrid approach combining analytic and holistic assessment. *Language Testing in Asia*, *9*(1), 1-22.

Yoon, C. (2016). Concordancers and dictionaries as problem-solving tools for ESL academic writing. *Language Learning & Technology*, *20*(1), 209-229.

Zhang, C. (2013). Effect of instruction on ESL students' synthesis writing. *Journal of Second Language Writing*, *22*(1), 51-67.

Zhu, P. (2015). N-Grams based linguistic search engine. *International Journal of Computational Linguistics Research*, *6*(1), 1-7.

# Appendix

**Rubric for Scoring Summary Writing**

| Dimension | Score & Level | | Criteria |
|---|---|---|---|
| 1 Content | 4 | very good | Can grasp all of the main ideas. Can develop the main point substantially by occasionally using secondary information. |
| | 3 | good | Can grasp most of the main ideas. Includes somewhat incorrect information or information beyond the original text, but it does not substantially deviate from the main point. |
| | 2 | fair | Can grasp only limited main ideas. Cannot demonstrate an adequate development of the main point. Noticeably includes incorrect information beyond the original text. |
| | 1 | poor | Cannot identify main ideas. Cannot grasp main idea correctly. |
| 2 Paraphrase (Quantity) | 4 | very good | Can paraphrase 80% or more of the expressions included in the summary in one's own words. |
| | 3 | good | Can paraphrase from 50 % to less than 80% of the expressions included in the summary in one's own words. |
| | 2 | fair | Can paraphrase only from 25% to less than 50% of the expressions included in the summary in one's own words. |
| | 1 | poor | Can paraphrase only less than 25% of the expressions included in the summary in one's own words. |
| 3 Paraphrase (Quality) | 4 | very good | Can actively attempt to paraphrase. Can demonstrate effective paraphrases where both sentence construction and vocabulary choice are different from the original text. |
| | 3 | good | Can actively attempt to paraphrase. Can paraphrase using vocabulary different from the original text. Seldom changes sentence construction from the original text. |
| | 2 | fair | Includes few expressions consisting of more than 4 words in a row copied from the original text. Can only demonstrate paraphrases using vocabulary from the original text. Deletes expressions partially or changes word order. |
| | 1 | poor | Includes a number of expressions consisting of more than 4 words in a row copied from the original text. Cannot demonstrate effective paraphrases. |
| 4 Language Use | 4 | very good | Can demonstrate a sophisticated range of vocabulary with effective words/idiom choice and usage. Can demonstrate effective and complex sentence construction with few grammatical errors. |
| | 3 | good | Can demonstrate an adequate range of vocabulary with good words/idiom choice and usage. Can demonstrate simple but effective sentence construction. Includes minor and occasional errors. |
| | 2 | fair | Can demonstrate only a limited range of vocabulary, words/idiom choice and usage. Can demonstrate simple sentence construction. Meaning is obscure due to frequent major errors. |
| | 1 | poor | Can demonstrate little knowledge of vocabulary, idioms, and word form. Can demonstrate little knowledge of sentence construction rules and English writing conversations. Meaning is obscure due to a number of minor and major errors. |
| 5 Overall Quality | 4 | very good | As a response to this task, the overall quality of this summary is…. |
| | 3 | good | |
| | 2 | fair | |
| | 1 | poor | |

*Note.* Adapted from Yamanishi et al. (2019).

Huang, A. Y. Q., Chang, J. W., Yang, A. C. M., Ogata, H., Li, S. T., Yen, R. X., & Yang, S. J. H. (2023). Personalized Intervention based on the Early Prediction of At-risk Students to Improve Their Learning Performance. *Educational Technology & Society*, *26*(4), 69-89. https://doi.org/10.30191/ETS.202310_26(4).0005

# Personalized Intervention based on the Early Prediction of At-risk Students to Improve Their Learning Performance

**Anna Y. Q. Huang[1], Jei Wei Chang[1], Albert C. M. Yang[2,4], Hiroaki Ogata[3], Shun Ting Li[1], Ruo Xuan Yen[1] and Stephen J. H. Yang[1*]**

[1]Computer Science & Information Engineering, National Central University, Taoyuan City, Taiwan // [2]Graduate School of Informatics, Kyoto University, Japan // [3]Academic Center for Computing and Media Studies, Kyoto University, Japan // [4]Computer Science and Engineering, National Chung-Hsing University, Taichung City, Taiwan // anna.yuqing@gmail.com // a0978122159@gmail.com // albertyang1994@gmail.com // hiroaki.ogata@gmail.com // ariel055132@g.ncu.edu.tw // apple310565@gmail.com // stephen.yang.ac@gmail.com

[*]Corresponding author

(Submitted June 24, 2022; Revised March 6, 2023; Accepted March 24, 2023)

**ABSTRACT:** To improve students’ learning performance through review learning activities, we developed a personalized intervention tutoring approach that leverages learning analysis based on artificial intelligence. The proposed intervention first uses text-processing artificial intelligence technologies, namely bidirectional encoder representations from transformers and generative pretrained transformer-2, to automatically generate Python programming remedial materials; subsequently, learning performance prediction models constructed using various machine learning methods are used to determine students’ learning type, enabling the automatic generation of personalized remedial materials. The participants in this study were 78 students from a university in northern Taiwan enrolled in an 8-week Python course. Students in the experimental ($n$ = 36) and control ($n$ = 42) groups engaged in the same programming learning activities during the first 5 weeks of the course, and they completed a pretest of programming knowledge in Week 6. For the review activity in Week 7, the 36 students in the experimental group received personalized intervention, whereas those in the control group received traditional class tutoring. We examined the effect of the self-regulated learning and personalized intervention on the learning performance of students. Compared with the traditional class tutoring, the personalized intervention review activity not only helped students obtain higher learning performance but also prompted greater improvements in the following learning strategies: rehearsal, critical thinking, metacognitive self-regulation, effort regulation, and peer learning. We also observed that students’ rehearsal and help-seeking learning strategies indirectly affected learning performance through students’ note-taking in the provided e-book.

**Keywords:** Personalized intervention, Self-regulated learning, Machine learning, Artificial intelligence

## 1. Introduction

With its roots in learning analysis, precision education aims to improve students’ learning performance through the four steps of diagnosis, prediction, treatment, and prevention, drawing inspiration from precision medicine (Lu et al., 2018). Personalized intervention involves formulating specific measures according to the needs of each student, guiding students to overcome learning difficulties, helping students confront high risks, and supporting students in improving their learning performance (Zhang et al., 2020). Thus, the objective of both precision education and personalized intervention is designing unique interventions for different students.

According to the US Department of Education (2010), personalized learning refers to teachers customizing learning objectives, teaching methods, and teaching content (and the sequence in which it is presented) according to the needs of learners and subsequently guiding students to engage in meaningful learning activities. Personalized learning aims to improve students’ learning by meeting their diverse needs, and research has indicated that students learn more effectively when teaching meets their needs (Benedict, 2010; Lin et al., 2016). Thus, personalized learning can be regarded as an alternative to traditional models because it focuses on providing guidance and addressing knowledge gaps on the basis of students’ current level of understanding (Johnson & Samora, 2016); notably, teachers have been increasingly adopting this approach in remedial tutoring for students (Foshee et al., 2016; Hsieh et al., 2013; Lin et al., 2016). Leveraging rich and diverse learning data, artificial intelligence and machine learning techniques can be used to analyze and predict student learning performance or to reveal their learning patterns at an early stage (Lu et al., 2018; Marbouti et al., 2016). Personalized learning based on learning analytics does not merely involve the assessment of learning and performance, but it can also improve learner engagement in the learning process (Bernacki et al., 2021). Siemens and Gašević (2012) defined learning analytics as measuring, collecting, analyzing, and reporting data about learners and their contexts to understand and optimize learning and the contexts in which it occurs. Personalized intervention applied in after-class tutoring has developed into an effective approach to improving students' learning effectiveness (Zhang et al., 2020).

In personalized intervention, students' learning status in terms of target concepts must first be predicted or identified; subsequently, they are provided with personalized learning adjustment suggestions. The personalized intervention proposed in this study first determines the students' mastery of concepts and then provides students with relevant remedial materials according to their learning status to help them review what they have learned, thereby improving students' conceptual proficiency. We used natural language processing (NLP) technologies, namely bidirectional encoder representations from transformers (BERT) (Devlin et al., 2018) and generative pretrained transformer (GPT-2) (Radford et al., 2018), for the generation of personalized intervention remedial learning content according to the status of students. Because self-regulated learning (SRL) can be used to explore students' cognition during the programming process (Zimmerman, 1989; Zimmerman, 1990), we used the learning strategy subscale of the motivated strategies for learning questionnaire (MSLQ) (Pintrich, 1991; Pintrich et al., 1993) to group students; subsequently, we explored the impact of the personalized intervention on students with different learning strategies. The following research questions guided this study:

- RQ1: For different learning concepts, what are the key learning features that affect students' mastery of concepts?
- RQ2: Can personalized intervention tutoring improve students' learning performance to a greater extent than traditional classroom tutoring?
- RQ3: Can learning strategies and online learning features effect on learning performance?
- RQ4: What is the impact of personalized intervention tutoring on the learning performance of students with different learning strategy abilities?

## 2. Literature review

### 2.1. SRL in a programming course

In SRL, learners examine their learning behavior and make adjustments to achieve learning goals and tasks. According to Zimmerman (1990), the first component of SRL is the student's metacognitive strategies for planning, self-monitoring, and controlling learning at various stages of the learning process. The second component of SRL is students' motivation and emotional processes for engaging in learning tasks (Pintrich, 1999; Zimmerman, 1989); for this component, researchers can use variables such as self-efficacy, task value, intrinsic goal orientation, and test anxiety as assessment dimensions. The third component of SRL is student behavioral processes, such as how students create and structure their learning environment (Zimmerman, 1989). The preceding description highlights that SRL is an active and constructive process by which students set goals for their learning and attempt to monitor, regulate, and control their cognition, motivation, and behavior in the learning process (Pintrich, 2000).

Problem-solving is a key skill required for the 21st century (Lai & Hwang, 2014), and computer programming has emerged as a popular subject for developing such skills. Because of the prominence of computer technology in modern society, the programming process is increasingly being used in educational settings to cultivate students' problem-solving ability. In the programming process, the problem is the main focus and the goal is to solve the problem through computer programming. Programming for problem-solving has become an essential ability for learners to construct the knowledge needed to perform new tasks; in this learning process related to new tasks, we can consider students' self-adaptive learning to observe their cognitive construction processes. In SRL, student cognition is the main element, and the relationship between students' metacognition, motivation, and behavioral participation as they complete learning tasks can be explored. Thus, in the learning process of programming for problem-solving, we can use SRL and adopt students' cognition as the main observation; we can then further explore each student's learning status in the programming process. In recent years, some researchers have begun to guide students to learn programming languages through SRL and investigate the correlation between students' SRL ability and learning performance (Cheng, 2021; Song et al., 2021).

### 2.2. Personalized intervention for at-risk students

In the field of teaching, at-risk students can be helped in a timely manner through early identification. Artificial intelligence technology and machine learning are increasingly being used to construct models for identifying at-risk students or learning patterns in educational settings. These models have mostly been used to predict student learning trends, student performance levels (Villagrá-Arnedo et al., 2017), whether students will pass or fail (Huang et al., 2020), and their academic scores (Lu et al., 2018). Most predictive models adopt machine learning methods, such as decision trees, linear regression, and support vector machines. These models can be used to identify at-risk students and determine when human intervention is needed as well as to provide assistance to such students. Although models can be used to identify at-risk students, the tutoring of these students to help improve their academic performance is still handled by teachers. Research on automatic generation of personalized learning suggestions for students with different learning trends, with the aim of improving the subsequent learning performance of at-risk students is limited. Therefore, the development of interventions for at-risk students has become a popular topic.

Personalized intervention refers to the provision of unique interventions according to the distinct learning characteristics and learning states of students (Zhang et al., 2020). In early education research, most personalized interventions were based on teachers' observations of students. Because of the effort involved, teachers cannot implement personalized interventions for all students in a given class. Fortunately, learning analysis based on big data analysis technology has grown rapidly, thereby facilitating the development of personalized interventions. For example, on the basis of learners' learning styles and cognitive abilities, Yi et al. (2017) implemented personalized interventions in an online learning environment; the intervention involved sending notification messages and emails to all students in the classroom. Most current personalized interventions send personalized learning messages according to students' personal characteristics (e.g., learning styles); however, personalized interventions can also be adopted for remedial materials aimed at addressing learning difficulties. Therefore, we developed a system that provides personalized remedial learning material content according to students' learning proficiency.

### 2.3. Automatic generation of remedial materials in personalized intervention

After learning, students may still have gaps in their understanding of certain concepts. Remedial materials are designed to help students fill those gaps in a timely manner (Bauman & Tuzhilin, 2018). Big data–based learning analysis is increasingly being adopted to provide students with personalized remedial materials to help them master subject content (Bauman & Tuzhilin, 2018; Bethard et al., 2012). Although the use of personalized remedial materials is receiving growing research attention, related research has continued to rely on teachers to generate such materials in advance to address students' learning difficulties. The automatic generation of remedial teaching materials for personalized intervention would thus be of considerable use to educators.

NLP toolkits for text processing have become a popular research topic. Novel NLP tools introduced recently include the Natural Language Toolkit (NLTK), TextBlob, and FLAIR. In NTLK, before the text is applied for different tasks, text processing, such as tokenization, tagging, parsing, is performed (Bird, 2006). The main goal of TextBlob is to calculate the polarity and subjectivity of text. FLAIR allows the application of different models, such as named entity recognition and part-of-speech tagging, to user tasks (Akbik et al., 2019). However, the aforementioned three toolkits are used mostly for sentiment analysis and text classification; they are not suitable for applications involving automatic text content generation.

Models pretrained on large corpora can learn common language representations from large amounts of unlabeled textual data; these language representations can then be used for applications in other natural language tasks, thereby avoiding training new models from scratch (Lu et al., 2021; Yang, 2021). Users only need to fine-tune the pretrained NLP model to be able to complete various downstream NLP tasks, with no need for users to design new neural network architectures for different tasks, thus substantially improving the training efficiency (Radford et al., 2018). The BERT pretrained model allows for the intuitive building of the model pipeline and for modeling many downstream tasks. Thus, researchers only need to input specific inputs and outputs into BERT; the model then fine-tunes all parameters during the training process (Devlin et al., 2018). This mechanism enables researchers to represent the steps of BERT in a more interpretable and localizable manner (Tenney et al., 2019). Therefore, we used two well-known pretrained models—BERT (Devlin et al., 2018) and GPT (Radford et al., 2018)—to automatically generate remedial materials.

## 3. Methodology

### 3.1. Participants and instruments

We conducted an 8-week experiment beginning in March 2022 to investigate whether the proposed personalized intervention tutoring could improve students' reviewing performance as they learned Python. The study participants are 78 students in two classes at a university in northern Taiwan; the two classes will be assigned as an experimental class and a control class. There were 36 students in the class of the experimental group, including 16 males and 20 females; the class of the control group consisted of 42 students, consisting of 28 males and 14 females. The only difference between the two groups was that, one week before the test, the students in the experimental group reviewed Python knowledge through personalized intervention tutoring, whereas those in the control group received traditional classroom tutoring.

The MSLQ is commonly used to measure students' self-regulation ability; it comprises two subscales—learning motivation and learning strategy—with items being rated on a five-point Likert scale (Pintrich, 1991, Pintrich et al., 1993). Thus, the instructor measured students' self-regulation ability by using the learning strategy subscale in the MSLQ. The learning strategy subscale of the MSLQ contains a total of 31 items on rehearsal (4 items), elaboration (6 items), organization (4 items), critical thinking (5 items), metacognitive self-regulation (12 items), time and study environment (8 items), effort regulation (4 items), peer learning (3 items), and help seeking (4 items). For the experimental group, the Cronbach reliability coefficients for our data in the pretest and posttest were .909 and .912, respectively. For the control group, these values were .884 and .948, respectively.

To reveal students' Python programming background knowledge, we conducted a programming knowledge test during the first week of the course. Only 41 and 29 students in the control and experimental groups, respectively, completed this test. No significant difference was observed between the experimental and control groups (Table 1); that is, the students in the two groups had substantial and comparable background knowledge of Python programming at the beginning of the course.

*Table 1.* Python programming background knowledge in the control and experimental groups

| Group | *N* | Mean | *SD* | *t*-value |
|---|---|---|---|---|
| $G_C$ | 41 | 42.78 | 17.95 | .70 |
| $G_E$ | 29 | 39.97 | 14.67 | |

*Note.* Independent sample *t* tests; $^{*}p < .05$.

In this study, ethics review approval (NTU-REC No.: 202005ES032) was granted by the National Taiwan University Research Ethics Committee. In addition, all participants in this study were informed that their learning event data would be collected, and all participants signed an informed consent form.

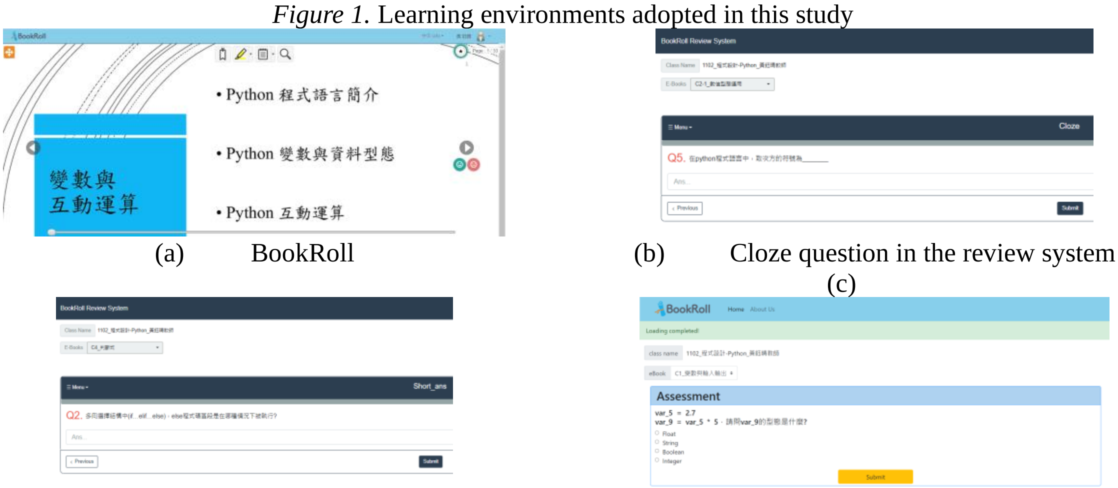
### 3.2. Experimental design of learning activities

We constructed a Python integrated learning environment, comprising an e-book reading system called BookRoll, a review system, and an assessment system, specifically for students in non information fields. BookRoll is an online e-book learning platform developed by the School of Social Information at Kyoto University (Japan). Because after-class review has a considerable impact on academic performance, the review system established in this study provides students with questions through cloze and short-answer questions. Finally, the assessment system allows students to tests their knowledge of key concepts after class through the use of multiple-choice questions; in this manner, students obtain an overview of their proficiency in key concepts.

In the Python programming course, the principle concepts are program output and input ($C_1$), strings ($C_2$), lists ($C_3$), and selection logic ($C_4$). The learning activities of the experimental and control groups are presented in Figure 2. In Week 1, students completed a test on Python programming background knowledge as well as a pretest of SRL ability measured using the MSLQ learning strategy subscale; during this week, students also received an introduction on how to operate the various learning environments (see Figure 1(a) to 1(d)) adopted in this study. Students in both groups engaged in two types of learning activities, namely programming learning activities and programming review activities, with only the review activities differing between the groups: The control group received traditional classroom tutoring review activities, whereas the experimental group received personalized intervention review activities.

From Weeks 2 to 5, the students in the experimental and control groups learned the basic concepts of Python programming ($C_1$–$C_4$) through three programming learning activities: programming instruction, programming knowledge review, and programming knowledge self-assessment. For the programming instruction activity, the teacher first uploads learning materials to BookRoll (see Figure 1(a)) and subsequently explains Python programming concepts in the classroom. In the programming knowledge review, students can review the key content of the learning materials through the two types of questions in the review system: cloze (see Figure 1(b)) and short answer (see Figure 1(c)). In the programming knowledge self-assessment, students use the assessment system (see Figure 1(d)) to confirm their mastery of key concepts through multiple-choice questions.

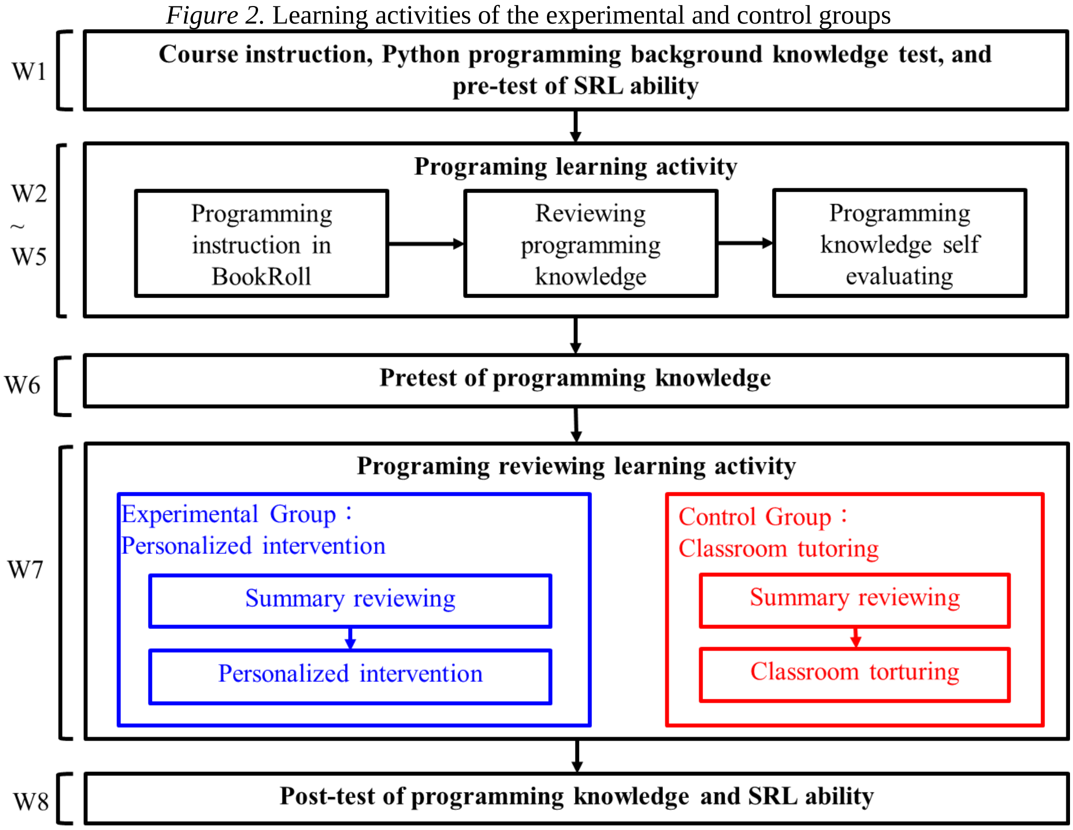
*Figure 1*. Learning environments adopted in this study




(a) BookRoll (b) Cloze question in the review system (c)

(c) Example short-answer question in the review system (d) Assessment system

*Figure 2*. Learning activities of the experimental and control groups

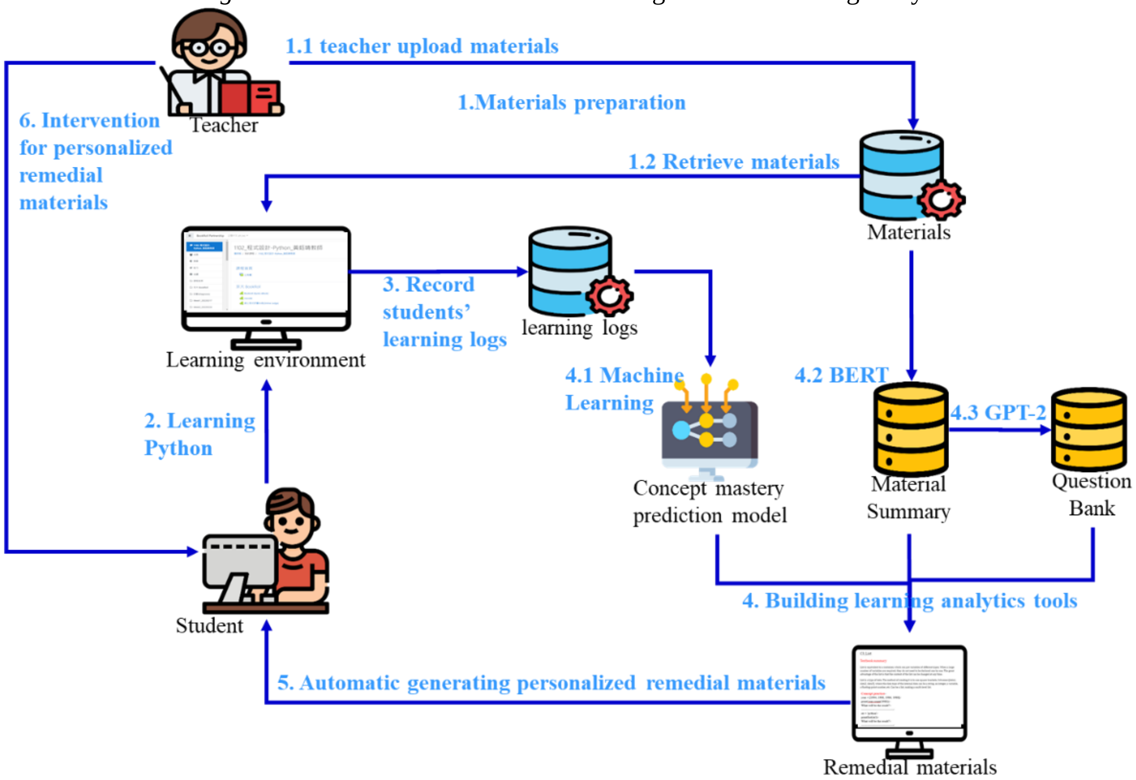



To help students prepare for the programming knowledge exam, we conducted various programming review activities for both groups in Week 7. First, the students in the experimental and control groups completed the same summary review activity, which focused on the key content of the four concepts ($C_1$ to $C_4$) learned in the course. Subsequently, the students in the experimental group received the personalized intervention (see Figure 2), as described in Section 3.3, whereas those in the control group received traditional classroom tutoring where students could clarify their doubts by asking the teacher questions and receiving related guidance. To measure the effect of the different review activities, a pretest of programming knowledge was conducted in Week 6, and then a posttest of programming knowledge and SRL ability was conducted in Week 8. The pretest and posttest questions are presented in Appendix A.

### 3.3. Personalized intervention process

To provide teachers with reference information regarding students' learning behavior during the tutoring intervention process, we designed a personalized intervention tutoring process that involves 6 steps based on learning analysis (see Figure 3). Step 1 involves material preparation. In Step 1.1, teachers upload the learning content to the materials database before class. In Step 1.2, the learning content is sent from the materials database to the online learning environments such as BookRoll. In Step 2, students learn Python programming in BookRoll under the guidance of a teacher. Step 3 involves students using the online learning environment, with their learning events collected and recorded in the learning log database. Step 4 involves the construction of learning analysis tools; in Step 4.1, a concept mastery prediction model is constructed using machine learning methods (see Section 3.5); in Step 4.2, a summary of materials is automatically generated using BERT and stored in the materials summary database (see Section 3.4.1); and in Step 4.3, questions are automatically generated using GPT-2 and stored in the question bank (see Section 3.4.2). On the basis of the learning analysis tool established in Step 4, in Step 5, personalized learning materials are automatically generated and sent to each student. Subsequently, in Step 6, teachers conduct personalized tutoring interventions for students according to the personalized materials automatically generated in Step 5.

*Figure 3*. Personalized intervention tutoring based on learning analytics




### 3.4. Automatic generation process of personalized remedial materials

In Step 5 of the personalized intervention process (see Figure 3), personalized remedial content is generated for each student; the steps in this content generation process are illustrated in Figure 4. The personalized remedial materials are obtained using two components: learning diagnosis and remedial content generation. Learning diagnosis involves three analysis systems: concept mastery prediction (see Section 3.5), self-evaluation results in the assessment system (see Section 3.2), and the extracted features (presented in Appendix B). The primary goal of the learning diagnosis component is to place each student into one of the following performance categories: proficient, practical improvement, and nonproficient. In the remedial content generation component, remedial content is automatically generated through two retrievers, one related to questions and the other to material summaries. The question retriever is designed to retrieve the practice questions automatically generated by the GPT-2 from the question bank; the summary retriever is designed to retrieve material summaries automatically extracted by BERT. Figure 5 depicts a screenshot of the content of the automatically generated personalized remediation material.

*Figure 4*. Steps in the automatic generation of personalized remedial materials

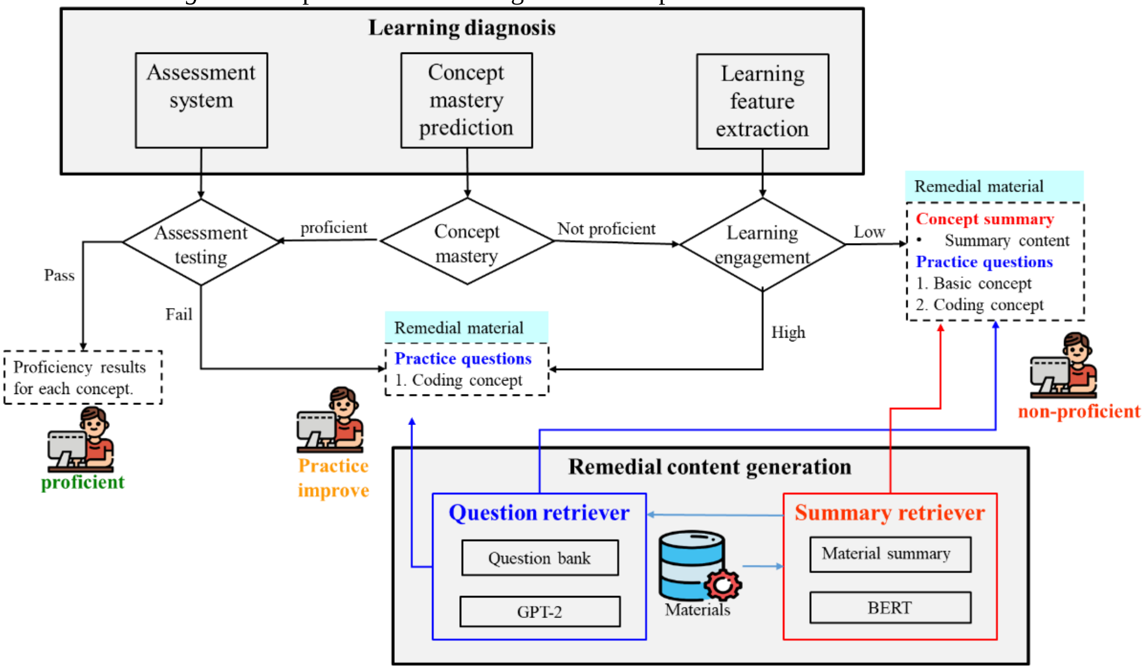
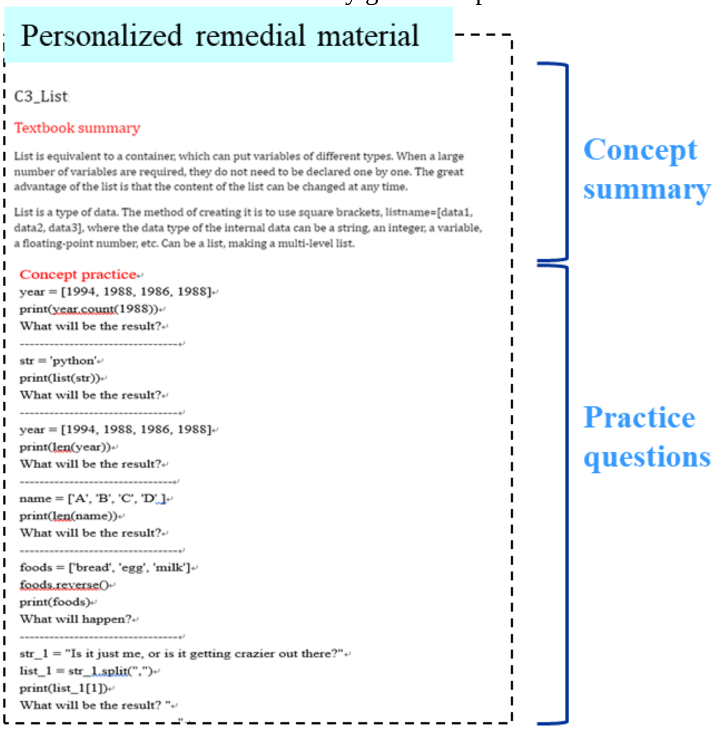



*Figure 5*. Screenshot of the automatically generated personalized remediation material



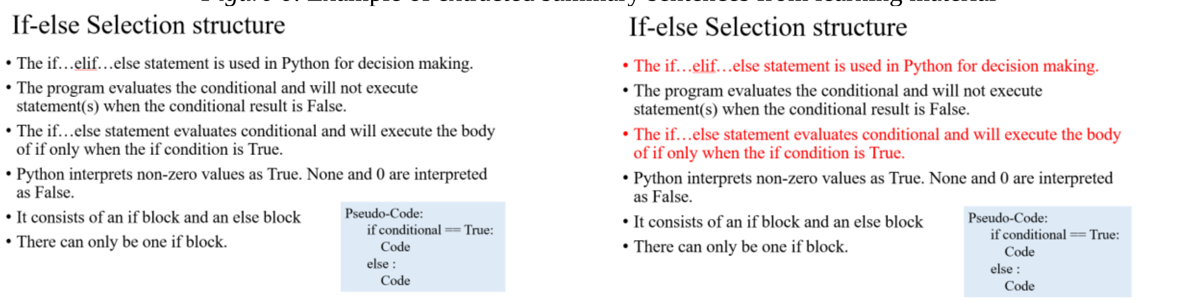

Proficient students are those students who are predicted to be proficient in learning concepts by the concept mastery prediction model and who receive a passing result in the assessment system. Because of their proficiency in learning concepts, these students do not need to receive additional supplementary materials, but this study will give proficiency results for each concept, thereby encouraging students to continue to study hard. Nonproficient students are those predicted to be nonproficient by the concept mastery prediction model and who have low values for extracted features. These low values indicate that these students lack active online learning behaviors. Students' poor online learning behaviors may lead to a lack of understanding of the learning content, which in turn can lead to them being deemed as nonproficient by the concept mastery prediction model. To help these students rapidly review key material before the test, we provide them with summary content through BERT, guide them to organize and acquire key concepts, and provide them with practice questions including basic concept questions and coding concept questions.

Finally, students are placed in the practice improvement category in two cases: (1) they are predicted to be proficient by the concept mastery model, but they receive a failing grade in the assessment system and (2) they are predicted to be nonproficient by the concept mastery model, but they have high values for the extracted features, indicating active participation in online learning. In the second case, although the students are active in online learning, they are still predicted to be nonproficient in key concepts. Such students require additional practice to achieve proficiency, hence the name of this group. Students in the practice improvement category lack the ability to apply what they have learned to programming; therefore, they receive coding concept questions, thereby helping them develop their programming skills as they solve problems.

### *3.4.1. Material summary automatic generation process*

The main goal of summarization is to extract the main idea of a document, generally combining relevant or important information into a concise structure (Allahyari et al., 2017). Summarization help students to not only rapidly obtain key content from learning material but also improve their review effectiveness. The BERT extractive summarizer model (see https://github.com/dmmiller612/bert-extractive-summarizer) first embeds the sentences in the input text into BERT's sentence embedding vector, then uses k-means to group all sentences in the text, and finally extracts the sentences closest to the cluster centroids as summaries (Miller, 2019). The BERT extractive summarizer model allows the user to specify the number of summary sentences to generate. For the summary extraction of Python learning materials, we first sorted the textbook content from the e-book and then used the BERT extractive summarizer model to extract the summary sentences and store them in the material summary database. Figure 6(a) and 6(b) are list examples of summary sentences extracted from the learning material.

*Figure 6*. Example of extracted summary sentences from learning material




(a)Screenshot of learning content from e-book. (b)The extracted summary sentences are marked in red font.

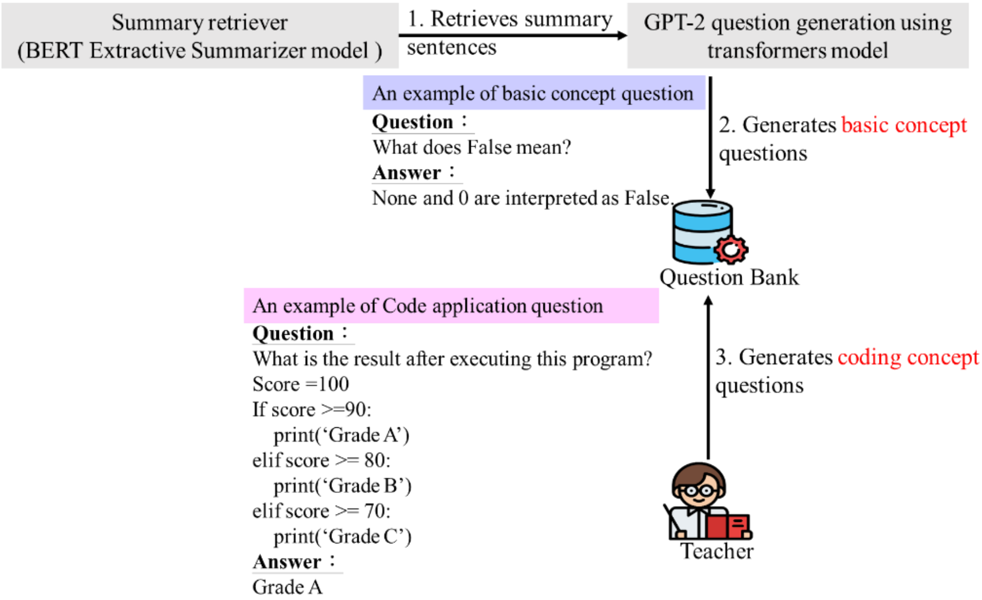
### *3.4.2. Question automated generation process*

Proposed by OpenAI, the GPT model is a language model that predicts the next word in an incomplete sentence; after the predicted new words are added, the next word is predicted again until a complete sentence is produced. The GPT-2 language model (Radford et al., 2018) focuses on the completion of some tasks such as answering questions and generating text output. The GTP can thus serve as a sentence generator; for this reason, it has been widely used in dialogue systems, medical text simplification, and many other applications (Ghojogh & Ghodsi, 2020).

The GPT-2 model of question generation using transformers (see https://github.com/patil-suraj/question_generation#question-generation-using-transformers) aims to generate questions with pretrained transformers through simplified data preprocessing. It can generate different types of questions and answers by using three models: the single-task question generation model, multitask question generation model, and end-to-end question generation model. The goal for the model in this study was to generate questions and answers together, with the answers appearing in the text. Therefore, we adopted the single-task question generation model.

Programming knowledge includes basic conceptual knowledge and coding conceptual knowledge. Basic conceptual knowledge can be extracted from the material content of textbooks, while coding conceptual knowledge focuses on explaining programming syntax through program code. In order to improve students' review efficiency through program practice questions, this study generates basic concept questions and coding concept questions for each learning concept. Figure 7 shows two examples of generating basic concept and coded concept questions. We applied GPT-2 model of question generation using transformers, which uses a single-task question generation model to automatically generate basic conceptual questions and answers from material summaries automatically generated by the summary retriever, and handed over to the instructor to confirm and modify the GPT-2 automatic generated basic concept questions and answers. For coding concept questions, the instructor will design coding application practice questions. Finally, both basic concept questions and coded concept questions will be stored in the question bank.

*Figure 7*. Examples of generated questions for basic concept questions and coding application questions



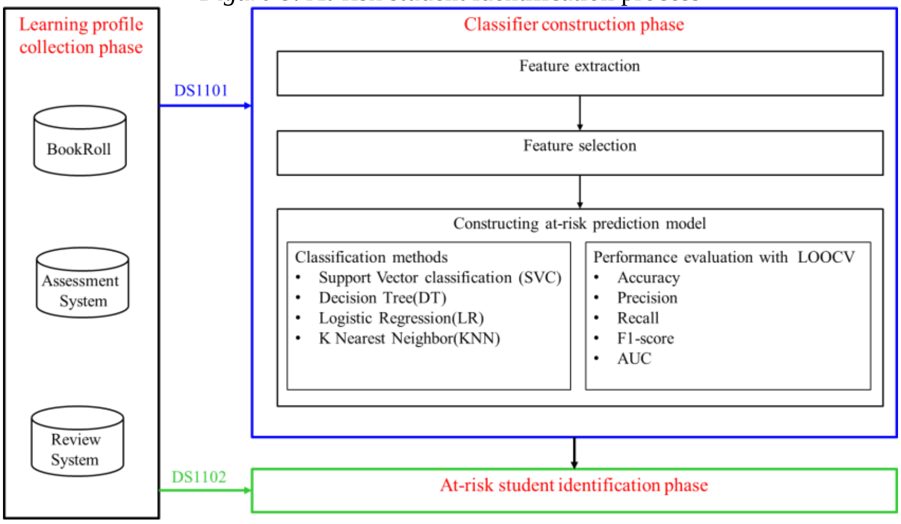

### 3.5. Construction process of at-risk student prediction model

To identify at-risk students in the targeted programming course, we proposed an at-risk student identification process (Figure 8), which comprises the learning profile collection, classifier construction, and at-risk student identification phases. In the learning portfolio collection phase, the learning events of students in the integrated learning environments (i.e., BookRoll, assessment system, and review system) were collected. In the classifier construction phase, with reference to the training samples, a classifier was constructed using the following steps: feature extraction, feature selection, and classifier construction. Finally, in the at-risk student identification phase, validation samples were identified through the constructed classifier.

In the learning portfolio collection phase, we collected students' learning logs from BookRoll, review system, and assessment system. In the at-risk student identification phase, the classification model is first constructed based on the training data set of DS1101, and then it is possible to predict whether the students in the data set DS1102 will pass or fail. In the classifier construction phase, we generated an at-risk student prediction model through three steps: feature extraction, feature selection, and construction. In the feature extraction step, relevant features were extracted from log data. Detailed instructions for each extracted feature are described in Appendix B. In the feature selection step, we use three selection methods—minimal redundancy maximal relevance (mRMR), chi-square test (Chi2), and relief algorithm—to identify the most relevant and powerful features from among the extracted features (Chandrashekar & Sahin, 2014). The mRMR method is used to identify the set of features in the original feature set that exhibits a high correlation with the output and a low correlation between the features themselves (Li et al., 2012); the chi-square test estimates whether a class label is independent of a feature (Jin et al., 2006); and the relief algorithm calculates a statistic for each feature that can be used to estimate feature quality or relevance to the output (Kira & Rendell, 1992).

In the construction phase of the at-risk student prediction model, we constructed prediction models by using support vector machine, decision tree, logistic regression, and k-nearest neighbor. Support vector machine is a supervised learning model that finds a separable hyperplane for samples by mapping samples to a high-dimensional space; it then predicts the class of new data samples. Logistic regression is used to construct a classification model by finding a regression line based on the probability of occurrence of sample classes (Kutner et al., 2005). K-nearest neighbor is a nonparametric classification method that establishes a classification model through the k nearest data points for a data sample (Cover & Hart, 1967). Finally, with decision trees, the focus is on building a classification model to find a tree structure that represents all known training data samples; the aim is to reveal hidden rules that identify categories based on feature values (Quinlan, 1983). To evaluate the prediction performance, this study used five metrics of accuracy, precision, recall, F1 measure, and area under the curve (AUC) (Ferri et al., 2009; Fawcett, 2004).

*Figure 8*. At-risk student identification process




## 4. Results and discussion

### 4.1. The critical online learning features for each programming learning concept

For the early identification of at-risk students, we used four feature selection methods to improve the prediction performance of the key learning features of each learning concept. The prediction performance of the four models is presented in Appendix C. Table 2 presents summaries of the prediction performance of classifications with the four feature selection methods for concepts $C_1$ to $C_4$. The results indicate that the classification using mRMR feature selection generally yielded the highest AUC values, ranging from 0.90 to .94. Therefore, we adopted mRMR as our feature selection method.

*Table 2*. Summaries of prediction performance of classifications with feature selection methods for concepts $C_1$ to $C_4$

| Feature selection | Accuracy | Precision | Recall | F1 | AUC |
|---|---|---|---|---|---|
| mRMR | .89~.94 | .94~.95 | .89~.90 | .88~.90 | .90~.94 |
| Chi-Square | .78~.89 | .84~.94 | .78~.89 | .76~.90 | .69~.94 |
| Relief | .80~.89 | .91~.93 | .80~.89 | .84~.89 | .83~.92 |

For exploring the key learning features that affect students' mastery of each learning concept through the features selected by the mRMR method (RQ 1). Table 3 indicates the key learning features selected by the mRMR for concepts $C_1$–$C_4$. For concepts $C_1$ and $C_2$, the key features for determining whether students have mastered the concepts were mostly related to the features extracted from BookRoll and the review system, namely review time ($f_4$), preview time ($f_5$), and correct answer rate in assessment ($f_7$). This is because concepts $C_1$ and $C_2$ involve relatively basic knowledge in programming; students can understand these two concepts by reading and reviewing the content of the e-book. For the content of concepts $C_3$ and $C_4$, the importance of the features associated with the assessment system ($f_6$, $f_7$) and review systems ($f_8$, $f_9$) increased accordingly. The conceptual content of concept $C_3$ (List) centers on the definitions of composite data types. In addition to understanding the definition of composite data type List in the textbook, students must also learn the operation and application of List. The conceptual content of concept $C_4$ (choice structure) centers on program logic. Students must familiarize themselves with various types of selection structures and methods through practical exercises; examples include the if...else structure, if–elif–else structure, and nested if structure. Therefore, because of the complexity of concepts $C_3$ and $C_4$, students cannot master them merely by reading or reviewing the content of the textbook. They would need to also further enhance their conceptual proficiency by practicing the questions relating to application or program tracking. That is, for students to become proficient in concepts $C_3$ and $C_4$, they must refer to the programming applications or coding track questions covered in the assessment system.

*Table 3*. Key learning features selected using the mRMR method for each learning concept

| Concept | At-risk prediction model | Key features |
|---|---|---|
| $C_1$ | SVM with mRMR | $f_4, f_9, f_5, f_7$ |
| $C_2$ | SVM with mRMR | $f_4, f_1, f_3, f_7$ |
| $C_3$ | DT with mRMR | $f_7, f_9, f_8, f_1, f_6$ |
| $C_4$ | DT with mRMR | $f_7, f_6, f_9, f_1$ |

**4.2. Effect of personalized intervention tutoring approach on students' learning performance**

The independent sample *t*-test results for the pretest and posttest are listed in Table 4. We observed no significant difference in the pretest results between the experimental and control groups ($t = -.12$, $p > .05$). This result suggests that students in the two groups achieved the same level of programming knowledge after engaging in Python programming learning activities for 4 weeks. The posttest scores of the experimental and control groups were 89.63 and 83.29 ($t = -2.44$, $p < .05$), respectively, representing a significant difference. The personalized intervention tutoring approach aims to recommend programming learning materials based on the predicted results of students' learning performance through machine learning methods. This result is consistent with previous research findings that learning resources recommended by machine learning can guide students to achieve higher learning performances (El-Bishouty et al., 2018). Accordingly, in response to RQ2, we found that compared with traditional classroom tutoring review, personalized intervention review can more effectively improve students' learning performance.

*Table 4.* Independent *t*-test results of pretest and posttest between control and experimental groups

| Group | *N* | Pre-test | | Post-test | |
|---|---|---|---|---|---|
| | | Mean/*SD* | *t*-value | Mean/*SD* | *t*-value |
| $G_C$ | 42 | 74.62/14.33 | -.12 | 83.21/15.65 | -2.44* |
| $G_E$ | 36 | 75.02/13.71 | | 89.72/6.86 | |

*Note.* $^*p < .05$.

After the programming review learning activity, students in the experimental group were asked to respond to four feedback questions to elucidate their views on the personalized intervention tutoring they received. The four feedback questions (see Appendix D) were answered using a 5-point Likert scale, ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Table 5 presents a summary of their responses. Of the 36 students in the experimental group, 30 provided feedback. The first question focused on how helpful students found the review activities conducted by the teacher. The average score for this question was 4.17, indicating that most students had positive perceptions of the classroom review activities and regarded them as effective. Notably, 16 of the 30 students who answered the question were classified as proficient. For students who are familiar with each concept, although no remedial materials will be given in this study, the familiarity evaluation results of each concept will still be provided to encourage students to continue to study hard. Since only 14 students in this study were not conceptually proficient, only these 14 students answered questions 2-4 and were used to discuss the feedback received on the remedial materials.

*Table 5.* Responses to the four feedback questions

| Question | Number of responds | Mean/ *SD* | Number of selected students for each point | | | | |
|---|---|---|---|---|---|---|---|
| | | | 1(SD) | 2(D) | 3(N) | 4(A) | 5(SA) |
| Q1 | 30 | 4.17/.79 | 0 | 1 | 4 | 14 | 11 |
| Q2 | 14 | 4.29/.83 | 0 | 1 | 0 | 7 | 6 |
| Q3 | 14 | 4.36/.84 | 0 | 1 | 0 | 6 | 7 |
| Q4 | 14 | 4.36/.84 | 0 | 1 | 0 | 6 | 7 |

*Note.* SD: strongly disagree; D: disagree; N: neutral; A: agree; SA: strongly agree.

Thirteen students indicated that the personalized intervention remedial materials not only helped them to review the conceptual content independently but also helped them to review the conceptual content that they were not familiar with (Table 5). This means that the students generally had a positive perception of the materials and were satisfied with personalized intervention tutoring approach. For Questions 2–4, only one student provided a rating of 2 points. The teacher interviewed this student to inquire into why he disagreed that the individualized intervention tutoring approach was effective. The student indicated that although the teacher's review activities were explained clearly, for computer programming, more time should be dedicated to practicing programming skills on the computer; moreover, he regarded conceptual learning as having low importance, which explains his disagreement with the relevant item on the questionnaire. This student argued that an individualized intervention tutoring approach that provides both conceptual review material regarding programming knowledge and personalized programming exercises would be of greater benefit to students' programming skills.

### 4.3. Impact of learning strategies and online learning features on learning performance

For the experimental and control groups, only 35 and 32 students, respectively, completed both the pretest and posttest for student's SRL ability. We used the SRL ability pretest result as a covariate and used analysis of covariance (ANCOVA) to identify significant differences between the two groups in the posttest of SRL ability; the ANCOVA results are presented in Table 6. For the rehearsal ($F$ = 8.15, $p$ < .05), critical thinking ($F$ = 11.93, $p$ < .05), metacognitive self-regulation ($F$ = 12.24, $p$ < .05), effort regulation ($F$ = 10.42, $p$ < .05), and peer learning ($F$ = 4.15, $p$ < .05) dimensions, the students in the experimental group scored significantly higher than those in the control group. These results indicate that, after the review activities, students who engaged in personalized intervention activities had significantly higher abilities in the aforementioned five learning strategies than those who engaged in traditional classroom tutoring activities. That is, personalized tutoring intervention was more effective than traditional classroom tutoring activities in improving students' abilities in the aforementioned five learning strategies.

*Table 6*. Analysis of covariance results of posttest SRL ability for control and experimental groups.

| MSLQ | Group | $N$ | Mean/$SD$ of pre-test | Post-test Mean/$SD$ | Post-test Adjusted Mean | Post-test Std. Error | $F$ |
|---|---|---|---|---|---|---|---|
| Rehearsal | $G_C$ | 32 | 3.80/.42 | 4.10/.58 | 3.61 | .094 | 8.153$^{**}$ |
| | $G_E$ | 35 | 3.93/.34 | 4.14/.44 | 3.99 | .090 | |
| Elaboration | $G_C$ | 32 | 3.68/.44 | 4.12/.57 | 3.74 | .085 | 2.912 |
| | $G_E$ | 35 | 3.76/.50 | 4.09/.49 | 3.94 | .081 | |
| Organization | $G_C$ | 32 | 3.57/.36 | 3.76/.40 | 3.71 | .095 | 1.057 |
| | $G_E$ | 35 | 3.45/.30 | 3.71/.30 | 3.85 | .091 | |
| Critical thinking | $G_C$ | 32 | 3.56/.47 | 3.72/.55 | 3.06 | .081 | 11.927$^{***}$ |
| | $G_E$ | 35 | 3.35/.78 | 3.74/.76 | 4.00 | .077 | |
| Metacognitive self-regulation | $G_C$ | 32 | 3.62/.48 | 3.67/.43 | 3.45 | .051 | 12.243$^{***}$ |
| | $G_E$ | 35 | 3.53/.51 | 4.05/.68 | 3.70 | .048 | |
| Time and study environment | $G_C$ | 32 | 3.72/.45 | 3.95/.51 | 3.55 | .047 | 0.571 |
| | $G_E$ | 35 | 3.60/.47 | 3.94/.64 | 3.60 | .045 | |
| Effort regulation | $G_C$ | 32 | 3.58/.33 | 4.02/.52 | 3.57 | .084 | 10.423$^{**}$ |
| | $G_E$ | 35 | 3.42/.29 | 3.72/.35 | 3.93 | .080 | |
| Peer learning | $G_C$ | 32 | 3.49/.23 | 3.65/.33 | 3.71 | .136 | 4.149$^{*}$ |
| | $G_E$ | 35 | 3.54/.42 | 3.94/.56 | 3.32 | .131 | |
| Help seeking | $G_C$ | 32 | 3.63/.75 | 3.32/.92 | 3.37 | .077 | 3.705 |
| | $G_E$ | 35 | 3.37/.39 | 3.55/.51 | 3.58 | .073 | |

*Note.* $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

In terms of the effect that students' SRL ability has on their academic performance, Song, Hong, and Oh (2021) found that although students' learning strategy ability was not significantly correlated with their academic performance, the online learning features in programming courses, such as the number of chosen program tasks and overall code-run trials, were significantly correlated with students' learning abilities. Therefore, we examined the correlation between online learning features (listed in Table 6) and learning strategy ability as well as the correlation between online learning features and learning performance; the results are presented in Table 7. Learning Features $f_3$ and $f_8$ were significantly correlated with the pretest result for learning performance ($r$ = .37, $p$ < .05 for $f_3$; $r$ = .48, $p$ < .01 for $f_8$). Moreover, although $f_8$ was not related to any of the learning strategies, feature $f_3$ was significantly related to $LS_1$ (rehearsal), $LS_3$ (organization), $LS_5$ (metacognitive self-regulation), $LS_6$ (time and study environment), $LS_7$ (effort regulation), $LS_8$ (peer learning), and $LS_9$ (help seeking).

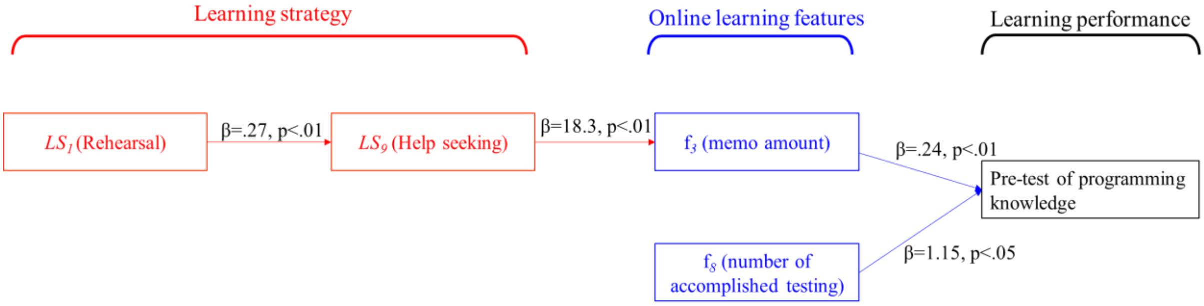
On the basis of the correlation results between online learning features and learning strategies (Table 7), we used multiple regression analysis to explore the correlations among online learning features, learning strategies, and learning performance (shows in Appendix E). Our results indicated that features $f_3$ ($\beta$ = .24, $p$ < .01) and $f_8$ ($\beta$ = 1.15, $p$ < .05) were significant predictors of pretest learning performance, $LS_9$ (help seeking) ($\beta$ = 18.30, $p$ < .01) was a significant predictor of feature $f_3$, and $LS_1$ (rehearsal) ($\beta$ = .27, $p$ < .05) was a significant predictor of $LS_9$ (help seeking). In terms of learning strategies, these results suggest that rehearsal first affects help seeking, which then affects the amount of notes students add, eventually affecting their learning performance. Figure 9 presents the conceptual diagram of the present regression analysis results for predictors $f_3$, $f_8$, $LS_1$, and $LS_9$ for pretest learning performance. The results in Table 6 indicate that students in the experimental group had significantly higher values than those in the control group for the rehearsal learning strategy. In addition, the results in Table 7 reveal that the rehearsal learning strategy was a predictor of help seeking, which in turn was a predictor of memo amount ($f_3$); finally, the amount of notes students wrote ($f_3$) and number of tests they completed ($f_8$) were predictors of learning performance. In sum to reply RQ3, the personalized intervention tutoring approach improved students' ability in terms of help seeking by promoting their rehearsal learning strategy, which in turn prompted students to write more notes on BookRoll, ultimately leading to improved learning performance.

*Table 7*. Pearson correlations among the extracted features, learning strategies, and learning performance.

| | $f_1$ | $f_2$ | $f_3$ | $f_4$ | $f_5$ | $f_7$ | $f_8$ | $f_9$ |
|---|---|---|---|---|---|---|---|---|
| $LS_1$ (Rehearsal) | 0.008 | 0.166 | $.409^{*}$ | 0.142 | $.447^{**}$ | 0.164 | 0.054 | -0.044 |
| $LS_2$ (Elaboration) | 0.002 | 0.195 | 0.301 | 0.11 | 0.251 | 0.219 | -0.05 | 0.144 |
| $LS_3$ (Organization) | 0.041 | $.451^{**}$ | $.428^{**}$ | 0.122 | 0.212 | 0.081 | 0.158 | 0.005 |
| $LS_4$ (Critical thinking) | -0.033 | 0.075 | 0.169 | 0.075 | 0.047 | 0.001 | -0.207 | -0.034 |
| $LS_5$ (Metacognitive self-regulation) | 0.14 | 0.263 | $.480^{**}$ | 0.287 | 0.246 | 0.134 | 0.157 | 0.105 |
| $LS_6$ (Time and study environment) | 0.318 | 0.185 | $.505^{**}$ | $.329^{*}$ | 0.309 | 0.304 | 0.297 | 0.191 |
| $LS_7$ (Effort regulation) | 0.05 | 0.054 | $.455^{**}$ | $.433^{**}$ | -0.018 | 0.148 | 0.08 | 0.135 |
| $LS_8$ (Peer learning) | 0.324 | 0.218 | $.338^{*}$ | 0.208 | $.445^{**}$ | 0.308 | 0.285 | $.441^{**}$ |
| $LS_9$ (Help seeking) | 0.249 | 0.09 | $.495^{**}$ | 0.226 | 0.229 | 0.036 | 0.184 | 0.021 |
| Pre-test | 0.28 | 0.3 | $.37^{*}$ | 0.26 | 0.17 | -0.03 | $.48^{**}$ | 0.15 |
| Post-test | -0.11 | 0.08 | -0.04 | -0.13 | 0.03 | 0.09 | 0.08 | 0.21 |

*Note.* $^{*}p$ < .05; $^{**}p$ < .01.

*Figure 9*. Conceptual diagram of the regression analysis results




## 4.4. Effect of personalized intervention on learning performance for each learning strategy

To respond to RQ4, we used two-way analysis of variance (ANOVA) to explore the impact of the interaction of review activity and learning strategies on learning performance. Because review activity is a categorical variable pointing to either personalized intervention or traditional classroom tutoring, we adopted k-means clustering to divide students' learning strategy ability into categorical variables at three levels: high, medium, and low. Among the nine learning strategies, four had significant effects on the interaction between the posttest results for learning strategies and learning performance, namely review activity × rehearsal ($F$ = 6.94, $p$ = .002), review activity × elaboration ($F$ = 4.21, $p$ = .019), review activity × organization ($F$ = 6.58, $p$ = .002), and review activity × critical thinking ($F$ = 9.53, $p$ = .000). The present two-way ANCOVA results are presented in Table 8, and the descriptive statistics for the high, medium, and low groups for rehearsal, elaboration, organization, and critical thinking are presented in Appendix F. Our results indicated that the review activity approach, four learning strategies (rehearsal, elaboration, organization, and critical thinking), and the interaction between the learning strategies and review activity had significant effects on students' learning performance.

*Table 8.* Two-way ANOVA results for review activity and four learning strategies—rehearsal, elaboration, organization, and critical thinking

| Variables | *SS* | *df* | *MS* | *F* | Significant |
|---|---|---|---|---|---|
| Review activity | 2119.55 | 1 | 2119.55 | 18.53 | .000 |
| Rehearsal | 2608.87 | 2 | 1304.44 | 11.40 | .000 |
| Review activity * Rehearsal | 1588.77 | 2 | 794.38 | 6.94 | .002 |
| Review activity | 1847.91 | 1 | 1847.91 | 14.30 | .000 |
| Elaboration | 1498.82 | 2 | 749.41 | 5.80 | .005 |
| Review activity * Elaboration | 109.13 | 2 | 545.06 | 4.21 | .019 |
| Review activity | 2071.90 | 1 | 2071.90 | 16.90 | .000 |
| Organization | 2216.64 | 2 | 1108.32 | 9.04 | .000 |
| Review activity * Organization | 1613.81 | 2 | 806.91 | 6.58 | .002 |
| Review activity | 212.80 | 1 | 212.80 | 18.52 | .000 |
| Critical Thinking | 1927.18 | 2 | 963.59 | 8.41 | .001 |
| Review activity * Critical Thinking | 2183.45 | 2 | 1091.72 | 9.53 | .000 |

*Table 9*. Simple main-effect analysis of learning performance in terms of the four learning strategies

| Variables | *SS* | *df* | *MS* | *F*-value | Post-Hoc |
|---|---|---|---|---|---|
| Review activity | | | | | |
| ($G_C$)Review with traditional class tutoring | 3351.44 | 2 | 1675.72 | $9.77^{**}$ | $Re_L = Re_M > Re_H$ |
| ($G_E$)Review with personalized intervention | 10.09 | 2 | 5.05 | 1.07 | |
| Learning strategy: Rehearsal | | | | | |
| $Re_L$ | 31.07 | 1 | 31.07 | .92 | |
| $Re_M$ | 273.97 | 1 | 273.97 | .10 | |
| $Re_H$ | 2747.76 | 1 | 2747.76 | $9.65^{**}$ | $G_E > G_C$ |
| Review activity | | | | | |
| ($G_C$)Review with traditional class tutoring | 233.11 | 2 | 1165.06 | $5.89^{**}$ | $El_L = El_M > El_H$ |
| ($G_E$)Review with personalized intervention | 51.21 | 2 | 25.60 | .53 | |
| Learning strategy: Elaboration | | | | | |
| $El_L$ | 9.20 | 1 | 9.20 | .28 | |
| $El_M$ | 173.23 | 1 | 173.22 | 1.66 | |
| $El_H$ | 216.00 | 1 | 216.00 | $7.39^{*}$ | $G_E > G_C$ |
| Review activity | | | | | |
| ($G_C$)Review with traditional class tutoring | 2734.82 | 2 | 1367.41 | $7.30^{**}$ | $Or_L = Or_M > El_H$ |
| ($G_E$)Review with personalized intervention | 127.22 | 2 | 63.61 | 1.36 | |
| Learning strategy: Organization | | | | | |
| $Or_L$ | 104.17 | 1 | 104.17 | 2.688 | |
| $Or_M$ | 216.75 | 1 | 216.75 | 2.207 | |
| $Or_H$ | 2700 | 1 | 2700 | $1.15^{**}$ | $G_E > G_C$ |
| Review activity | | | | | |
| ($G_C$)Review with traditional class tutoring | 3422.29 | 2 | 1711.15 | $1.08^{***}$ | $CT_L = CT_M > CT_H$ |
| ($G_E$)Review with personalized intervention | 19.95 | 2 | 9.98 | .20 | |
| Learning strategy: Critical Thinking | | | | | |
| $CT_L$ | 105.94 | 1 | 105.94 | .98 | |
| $CT_M$ | 5.98 | 1 | 5.98 | .13 | |
| $CT_H$ | 3217.47 | 1 | 3217.47 | $12.36^{**}$ | $G_E > G_C$ |

*Note.* $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

On the basis of the results in Table 8, we conducted a simple main effects analysis to examine the effect of review activities on the learning performance of students with different SRL ability levels; the results are listed in Table 9. In the traditional classroom tutoring, students with different ability levels in rehearsal, elaboration, organization, and critical thinking exhibited significantly different learning performances. In the personalized intervention, no significant differences were observed in students' learning performance at different ability levels for the four aforementioned learning strategies. The post hoc results of traditional classroom tutoring revealed that students with medium and low ability levels in rehearsal ($F = 9.77$, $p < .01$, $Re_L = Re_M > Re_H$), elaboration ($F = 5.89$, $p < .01$, $El_L = El_M > El_H$), organization ($F = 7.30$, $p < .01$, $Or_L = Or_M > El_H$), and critical thinking ($F = 1.08$, $p < .001$, $CT_L = CT_M > CT_H$) had significantly higher learning performance than students with high ability levels in these four learning strategies. This result suggests that students with strong learning strategy abilities may need more review content to maximize their learning performance after the review activity. This result also confirms that personalized intervention activities can help students with high learning strategy ability maximize their learning performance through the provision of additional review content; this explains why the experimental group students at all three learning strategy levels did not differ significantly in learning performance.

We subsequently examined the effect of different learning strategy levels on the learning performance prompted by the two review activities; for high-level rehearsal ($F$ = 9.62, $p$ < .01, $G_E > G_C$), elaboration ($F$ = 7.39, $p$ < .05, $G_E > G_C$), organization ($F$ = 1.15, $p$ < .01, $G_E > G_C$), and critical thinking ($F$ = 12.36, $p$ < .01, $G_E > G_C$), we observed significant differences in learning performance between the experimental and control groups (see Table 9). However, no significant difference was noted in students' learning performance between the two groups for low- and medium-level learning strategies. These results suggest that students with strong learning strategy abilities can achieve significantly higher learning performance through personalized intervention activities than through traditional class tutoring activities.

## 5. Conclusions

Artificial intelligence and machine learning technology has stimulated the development of personalized interventions in remedial coaching. We proposed a personalized intervention approach based on artificial intelligence technology for use in a computer programming course. Our results indicate that after the review activity, students who received personalized intervention had significantly higher learning performance than those who merely received class tutoring. This result confirms the effectiveness of the personalized intervention approach in helping students to review content.

Because SRL can be used to explore students' cognition during problem-solving as part of programming learning, we also considered how students' SRL abilities would influence their learning performance under the proposed personalized intervention approach. Our results reveal that the proposed personalized intervention prompted improvements in the following learning strategies: rehearsal, critical thinking, metacognitive self-regulation, effort regulation, and peer learning. We found that, although students' ability in each learning strategy was not directly related to learning performance, the learning strategies of rehearsal and help seeking indirectly affected learning performance through Learning Feature $f_3$ (memo amount).

In terms of the interaction effect between individualized intervention and learning strategies on learning performance, we observed significant results for rehearsal, refinement, organization, and critical thinking. Students with strong abilities in these four learning strategies achieved higher learning performance in the personalized intervention approach than in the class tutoring approach. That is, for such students, the additional review content provided by the personalized intervention effectively supported their reviewing, thereby improving their learning performance.

### 5.1. Limitation

The participants in this study were mainly students from noninformation fields; thus, the curriculum of this course would not be suitable for students majoring in computer science. Moreover, the proposed personalized intervention review activity focuses on programming-related knowledge, with the assessment of learning performance also focusing on programming concepts; that is, students' actual program coding ability was not tested. Therefore, the proposed intervention may require some modifications for effectively improving students' coding ability. The participants in this manuscript are only the number of students in two classes (the students in the experimental group and the control group were 36 and 42, respectively), and subsequent large-scale empirical studies are need to be verified.

## Acknowledgment

This work was supported by grants from the Taiwan Ministry of Science and Technology, including MOST-111-2410-H-008-010-MY3 (its Research Ethics Committee case number is 202205ES065), MOST-109-2511-H-008-007-MY3 (its Research Ethics Committee case number is 202005ES032), MOST-108-2511-H-008-009-MY3 (its Research Ethics Committee case number is 201812ES028), and the Taiwan Ministry of Education.

## References


Akbik, A., Bergmann, T., Blythe, D., Rasul, K., Schweter, S., & Vollgraf, R. (2019). FLAIR: An Easy-to-use framework for state-of-the-art NLP. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics* (pp. 54-59).

Allahyari, M., Pouriyeh, S., Assefi, M., Safaei, S., Trippe, E. D., Gutierrez, J. B., & Kochut, K. (2017). *Text summarization techniques: A Brief survey*. https://doi.org/10.48550/arXiv.1707.02268

Bauman, K., & Tuzhilin, A. (2018). Recommending remedial learning materials to students by filling their knowledge gaps. *MIS Quarterly, 42*(1), 313-332.

Benedict, N. (2010). Virtual patients and problem-based learning in advanced therapeutics. *American Journal of Pharmaceutical Education, 74*(8), 1-5.

Bernacki, M. L., Greene, M. J., & Lobczowski, N. G. (2021). A Systematic review of research on personalized learning: Personalized by whom, to what, how, and for what purpose (s)? *Educational Psychology Review, 33*(4), 1675-1715.

Bethard, S., Okoye, I., Hang, H., Sultan, A., Martin, J. H., & Sumner, T. (2012). Identifying science concepts and student misconceptions in an interactive essay writing tutor. In *Proceedings of the 7th Workshop on the Innovative Use of NLP for Building Educational Applications* (pp. 12-21). Association for Computational Linguistics.

Bird, S. (2006). NLTK: The Natural language toolkit. In *Proceedings of the COLING/ACL 2006 Interactive Presentation Sessions* (pp. 69-72). Association for Computational Linguistics.

Chandrashekar, G., & Sahin, F. (2014). A Survey on feature selection methods. *Computers & Electrical Engineering, 40*(1), 16-28.

Cheng, G. (2021). The Effects of self-regulated learning strategies on computer programming achievement in teacher education. *Frontiers in Education Technology, 4*(4), 37-58.

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE transactions on information theory, 13*(1), 21-27.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). *Bert: Pre-training of deep bidirectional transformers for language understanding.* https://doi.org/10.48550/arXiv.1810.04805

El-Bishouty, M., Aldraiweesh, A., Alturki, U., Tortorella, R., Yang, J., Chang, T., Sabine, G., & Kinshuk. (2018). Use of Felder and Silverman learning style model for online course design. *Educational Technology Research and Development, 67*, 161–177.

Fawcett, T. (2004). ROC graphs: Notes and practical considerations for researchers. *Machine learning, 31*(1), 1-38.

Ferri, C., Hernández-Orallo, J., & Modroiu, R. (2009). An Experimental comparison of performance measures for classification. *Pattern recognition letters, 30*(1), 27-38.

Foshee, C. M., Elliott, S. N., & Atkinson, R. K. (2016). Technology-enhanced learning in college mathematics remediation. *British Journal of Educational Technology, 47*(5), 893-905.

Ghojogh, B., & Ghodsi, A. (2020). Attention mechanism, transformers, BERT, and GPT: Tutorial and survey. https://doi.org/10.31219/osf.io/m6gcn

Hsieh, T.-C., Lee, M.-C., & Su, C.-Y. (2013). Designing and implementing a personalized remedial learning system for enhancing the programming learning. *Journal of Educational Technology & Society, 16*(4), 32-46.

Huang, A. Y., Lu, O. H., Huang, J. C., Yin, C. J., & Yang, S. J. (2020). Predicting students' academic performance by using educational big data and learning analytics: Evaluation of classification methods and learning logs. *Interactive Learning Environments, 28*(2), 206-230.

Jin, X., Xu, A., Bie, R., & Guo, P. (2006). Machine learning techniques and chi-square feature selection for cancer classification using Sage gene expression profiles. In J. Li, Q. Yang, & A. H. Tan (Eds.), *Data Mining for Biomedical Applications. BioDM 2006. Lecture Notes in Computer Science (Vol. 3916)*. Springer. https://doi.org/10.1007/11691730_11

Johnson, D., & Samora, D. (2016). The Potential transformation of higher education through computer-based adaptive learning systems. *Global Education Journal, 2016*(1), 1-17.

Kira, K., & Rendell, L. A. (1992). A Practical approach to feature selection. In *Machine learning proceedings 1992* (pp. 249-256). Elsevier.

Kohavi, R. (1995). A Study of cross-validation and bootstrap for accuracy estimation and model selection. *IJCAI95: Proceeding of International Joint Conference on Articial Intelligence, 14*(2), 1137-1145.

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models*. McGraw-Hill New York.

Lai, C. L., & Hwang, G. J. (2014). Effects of mobile learning time on students' conception of collaboration, communication, complex problem–solving, meta–cognitive awareness and creativity. *International Journal of Mobile Learning and Organization, 8*(3-4), 276-291.

Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R. P., Tang, J., & Liu, H. (2017). Feature selection: A Data perspective. *ACM computing surveys (CSUR), 50*(6), 1-45.

Li, B. Q., Hu, L. L., Chen, L., Feng, K. Y., Cai, Y. D., & Chou, K. C. (2012). Prediction of protein domain with mRMR feature selection and analysis. *PloS one, 7*(6), e39308. https://doi.org/10.1371/journal.pone.0039308

Lin, C. C., Guo, K. H., & Lin, Y. C. (2016). A Simple and effective remedial learning system with a fuzzy expert system. *Journal of Computer Assisted Learning, 32*(6), 647-662.

Lu, O. H., Huang, A. Y., Huang, J. C., Lin, A. J., Ogata, H., & Yang, S. J. (2018). Applying learning analytics for the early prediction of Students' academic performance in blended learning. *Educational Technology & Society, 21*(2), 220-232.

Lu, O. H., Huang, A. Y., Tsai, D. C., & Yang, S. J. (2021). Expert-authored and machine-generated short-answer questions for assessing students learning performance. *Educational Technology & Society, 24*(3), 159-173.

Marbouti, F., Diefes-Dux, H. A., & Madhavan, K. (2016). Models for early prediction of at-risk students in a course using standards-based grading. *Computers & Education, 103,* 1-15.

Miller, D. (2019). *Leveraging BERT for extractive text summarization on lectures*. https://doi.org/10.48550/arXiv.1906.04165

Pintrich, P. R. (1991). *A Manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ).* The Regents of The University of Michigan.

Pintrich, P. R. (1999). The Role of motivation in promoting and sustaining self-regulated learning. *International journal of educational research, 31*(6), 459-47.

Pintrich, P. R. (2000). The Role of goal orientation in self-regulated learning. In *Handbook of self-regulation* (pp. 451-502). Elsevier.

Pintrich, P. R., Smith, D. A., Garcia, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and psychological measurement, 53*(3), 801-813.

Quinlan, J. R. (1983). Learning efficient classification procedures and their application to chess end games. In *Machine learning* (pp. 463-482). Springer.

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). *Improving language understanding by generative pre-training*. https://www.cs.ubc.ca/~amuham01/LING530/papers/radford2018improving.pdf

Siemens, G., & Gašević, D. (2012). Special issue on learning and knowledge analytics. *Educational Technology & Society, 15*(3), 1–2.

Song, D., Hong, H., & Oh, E. Y. (2021). Applying computational analysis of novice learners' computer programming patterns to reveal self-regulated learning, computational thinking, and learning performance. *Computers in Human Behavior, 120*, 106746. https://doi.org/10.1016/j.chb.2021.106746

U.S. Department of Education. (2010). *Transforming American education: Learning powered by technology*. Office of Educational Technology.

Villagrá-Arnedo, C. J., Gallego-Durán, F. J., Llorens-Largo, F., Compañ-Rosique, P., Satorre-Cuerda, R., & Molina-Carmona, R. (2017). Improving the expressiveness of black-box models for predicting student performance. *Computers in Human Behavior, 72*, 621–631.

Wolff, A., Zdrahal, Z., Herrmannova, D., Kuzilek, J., & Hlosta, M. (2014). Developing predictive models for early detection of at-risk students on distance learning modules. In *Machine Learning and Learning Analytics Workshop at the 4th International Conference on Learning Analytics and Knowledge (LAK14)*.

Wong, T. T. (2015). Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition, 48*(9), 2839-2846.

Yang, S. J. H. (2021). Guest editorial: Precision education - A New challenge for AI in education. E*ducational Technology & Society, 24*(1), 105–108. https://doi.org/10.30191/ETS.202101_24(1).0008

Yi, B., Zhang, D., Wang, Y., Liu, H., Zhang, Z., Shu, J., & Lv, Y. (2017). Research on personalized learning model under informatization environment. In *2017 International Symposium on Educational Technology (ISET)* (pp. 48–52). IEEE. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8005386

Zhang, J. H., Zou, L. C., Miao, J. J., Zhang, Y. X., Hwang, G. J., & Zhu, Y. (2020). An Individualized intervention approach to improving university students' learning performance and interactive behaviors in a blended learning environment. *Interactive Learning Environments, 28*(2), 231–245.

Zimmerman, B. J. (1989). A Social cognitive view of self-regulated academic learning. *Journal of educational psychology, 81*(3), 329–339. https://doi.org/10.1037/0022-0663.81.3.329

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An Overview. *Educational psychologist, 25*(1), 3-17.

# Appendix A

Pretest and posttest questions for programming knowledge

| Concept: C1 | |
|---|---|
| Pre-test | Post-test |
| **Question C1_1:**<br>Which of the following variables is named incorrectly?<br>Option 1: _name<br>Option 2: name1<br>Option 3: 1name<br>Option 4: name_<br>**Question C1_2:**<br>The symbol of annotation in Python is ?<br>**Question C1_3:**<br>Which of the following options are not reserved words?<br>Option 1: not<br>Option 2: and<br>Option 3: or<br>Option 4: xor<br><br>**Question C1_4:**<br>A Python program is given below:<br>print(‘\’’, ‘\’’, sep = ‘A’)<br>What is the result after executing this program? | **Question C1_1:**<br>Which of the following variables is named incorrectly?<br>Option 1: _996apple<br>Option 2: _apple<br>Option 3: 996_apple<br>Option 4: apple_5<br>**Question C1_2:**<br>What symbols do I need to replace ? to print the ‘ symbol print(‘ ? ‘)<br>Option 1: #<br>Option 2: !:<br>Option 3: /<br>Option 4: \<br>**Question C1_3:**<br>A Python program is given below:<br>print(“=“, “=“, sep = “u”)<br>What is the result after executing this program? |
| Concept: C2 | |
| **Question C2_1:**<br>A Python program is given below:<br>mail = ‘apple@gmail’<br>mail = mail.split(‘@’)<br>print(mail[1])<br>What is the result after executing this program?<br>**Question C2_2:**<br>A Python program is given below:<br>mail = ‘meme_@gmail.com’<br>index = mail.find(‘@’)<br>ans = mail[0:index:1]<br>print(ans)<br>What is the result after executing this program?<br>**Question C2_3:**<br>What are the results after executing these arithmetic formulae respectively? (Be careful: 5 and 5.0 are different answers)<br>7/2<br>float(7//2)<br>int(7%2)<br>**Question C2_4:**<br>What are the results after executing these programs? (If the program cannot execute, please fill in unable to execute)<br>L = [‘Python’, ‘Hello world’, 5]<br>L.spilt()<br>L[1].spilt() | **Question C2_1:**<br>Which of the following is the return value of ‘test’.find(‘t’)?<br>Option 1: 3<br>Option 2: -1<br>Option 3: 0<br>Option 4: True<br>**Question C2_2:**<br>Which of the following options has the meaning of squared root?<br>Option 1: **(1/2)<br>Option 2: *2<br>Option 3: //2<br>Option 4: **2<br>**Question C2_3:**<br>A Python program is given below:<br>temp = 7 // 2<br>ans = temp % 2<br>print(ans)<br>What is the result after executing this program?<br>**Question C2_4:**<br>A Python program is given below:<br>temp = ‘中央大學’<br>index = temp.find(‘大’)<br>temp[ index : ]<br>What is the result after executing this program?<br>**Question C2_5:** |

| | |
|---|---|
| L[2].spilt()<br>**Question C2_5:**<br>A Python program is given below:<br>alphabet = ‘abcdefghijklmnopqrstuvwxyz’<br>n = 26<br>print(alphabet[n%2])<br>What is the result after executing this program?<br>**Question C2_6:**<br>Suppose s = ‘abc’, what will the output of s.find(‘z’) be?<br>Option 1: 0<br>Option 2: 2<br>Option 3: 1<br>Option 4: -1 | A Python program is given below:<br>url = “https://www.ncu.edu.tw/tw/index.html”<br>print( url.count(‘/’) )<br>What is the result after executing this program? |
| Concept: C3 | |
| **Question C3_1:**<br>Suppose a = [1,2,3] becomes [4,1,2,3] through a certain program, which of the following functions is needed for the program?<br>Option 1: None of them<br>Option 2: a.append(4)<br>Option 3: a[0] = 4<br>Option 4: a.insert(0,4)<br>**Question C3_2:**<br>Suppose a = [1,2,3], what are the results after executing the following programs respectively?<br>Program 1: myList[2:-1]<br>Program 2: myList[1][-1]<br>**Question C3_3:**<br>Which of the following is the functions that calculates the number of times that a particular elements is inside the list?<br>Option 1: .find()<br>Option 2: .count()<br>Option 3: .replace()<br>Option 4: .len() | **Question C3_1:**<br>A Python program is given below:<br>myList = [ [1,2,3], [4,5,6], 7, [1, [2, 3]] ]<br>Which of the following is the output of len(myList)?<br>Option 1: 5<br>Option 2: 10<br>Option 3: 3<br>Option 4: 4<br>**Question C3_2:**<br>A Python program is given below:<br>mylist = [1, 2, 3]<br>Which of the following is the output of mylist.count(2)?<br>Option 1: True<br>Option 2: -2<br>Option 3: 0<br>Option 4: 1<br>**Question C3_3:**<br>A Python program is given below:<br>temp = [1,2,3]<br>temp.insert(2,4)<br>temp.append(1)<br>print(len(temp))<br>What is the result after executing this program?<br>**Question C3_4:**<br>A Python program is given below:<br>temp = [1,2,5]<br>temp.reverse()<br>print(temp[0])<br>What is the result after executing this program? |
| Concept: C4 | |
| **Question C4_1:**<br>A Python program is given below:<br>mail = ‘anna@gmail.com’<br>if(‘gmail’ in mail):<br>print(True)<br>else:<br>print(False)<br>What is the result after executing this program?<br>**Question C4_2:**<br>Which of the following options will output false?<br>Option 1: print( 1 != 1 and 1 == 1 )<br>Option 2: print(1 != 2)<br>Option 3: print(1 == 1)<br>Option 4: print( 1 != 1 or 1 == 1 ) | **Question C4_1:**<br>Which of the following is not a reserved word for logical judgement?<br>Option 1: or<br>Option 2: not<br>Option 3: and<br>Option 4: break<br>**Question C4_2:**<br>Which of the following options will output false?<br>Option 1: 1 == 1 and 1 != 2<br>Option 2: 1 == 1 or 1 == 3<br>Option 3: ‘a’ == ‘a’<br>Option 4: not(1 == 1 or 1 == 3)<br>**Question C4_3:** |

| | |
|---|---|
| **Question C4_3:**<br>In a conditional program (nested if is not considered), what are the maximum number and minimum number of if and else? | A Python program is given below:<br>score = 59<br>if score > 60:<br>    print('ans1')<br>    if score < 60:<br>        print('ans2')<br>else:<br>    print('ans3')<br>What is the result after executing this program?<br>**Question C4_4:**<br>A Python program is given below:<br>if 'a' in ['ab', 'c']:<br>    print('ans1')<br>else:<br>    print('ans2')<br>What is the result after executing this program? |

## Appendix B

Extracted features for each learning system

| Learning system | Feature | Description |
|---|---|---|
| BookRoll | $f_1$: reading time | Total e-book reading time in class for each concept. |
| | $f_2$: marker amount | Number of markers added to the e-book per concept. |
| | $f_3$: memo amount | Number of memos added to the e-book per concept. |
| | $f_4$: review time | Total e-book reading time after class for each concept. |
| | $f_5$: preview time | Total e-book reading time before class for each concept. |
| Assessment system | $f_6$: time of first assessment | The first time each concept was tested in the assessment system. |
| | $f_7$: correct answer rate in assessment | Correct answer rate for each concept in the assessment system. |
| | $f_8$: number of completed tests | Number of completed tests in the assessment system. |
| Review system | $f_9$: correct answer rate in review | Correct answer rate for each concept in the review system. |
| | $f_{10}$: time of first review | The first time each concept was tested in the review system. |

## Appendix C

Predictive performance of various feature selection algorithms with different models for each concept.

| Concept | Model | Performance: Accuracy/Precision/Recall/F1/AUC | | | |
|---|---|---|---|---|---|
| | | Without feature selection | Feature selection method | | |
| | | | mRMR | Chi2 | Relief |
| $C_1$ | SVM | .70/.80/.70/.64/.63 | .90/.95/.90/.91/.94 | .70/.75/.70/.72/.63 | .80/.84/.80/.76/.67 |
| | LR | .75/.81/.75/.68/.58 | .90/.91/.90/.88/.67 | .80/.84/.80/.76/.67 | .70/.79/.70/.74/.39 |
| | KNN | .65/.78/.65/.55/.56 | .90/.91/.90/.89/.83 | .80/.64/.80/.71/.50 | .80/.93/.80/.84/.89 |
| | DT | .80/.80/.80/.80/.61 | .80/.91/.90/.89/.83 | .80/.80/.80/.80/.69 | .80/.80/.80/.80/.69 |
| $C_2$ | SVM | .78/.77/.78/.76/.71 | .89/.94/.89/.90/.94 | .78/.78/0/78/.78/.75 | .67/.68/.67/.67/.68 |
| | LR | .78/.77/.78/.76/.71 | .94/.95/.94/.94/.83 | .67/.79/.67/.61/.63 | .67/.70/.67/.68/.67 |
| | KNN | .67/.67/.67/.67/.63 | .78/.89/.78/.80/.86 | .78/.78/.78/.78/.68 | .89/.92/.89/.89/.92 |
| | DT | .78/.76/.78/.76/.66 | .89/.91/.89/.89/.90 | .78/.87/.78/.78/.83 | .67/.79/.67/.61/.63 |
| $C_3$ | SVM | .78/.69/.78/.73/.47 | .89/.90/.89/.87/.75 | .78/.83/.78/.74/.67 | .78/.67/.78/.68/.50 |
| | LR | .78/.83/.78/.72/.60 | .89/.90/.89/.87/.67 | .77/.78/.77/.77/.44 | .78/.93/.78/.82/.88 |
| | KNN | .72/.77/.72/.75/.41 | .89/.90/.89/.87/.75 | .89/.94/.89/.90/.94 | .89/.90/.89/.87/.75 |
| | DT | .78/.78/.78/.78/.44 | .89/.94/.89/.90/.94 | .78/.84/.78/.76/.75 | .67/.92/.67/.73/.81 |
| $C_4$ | SVM | .77/.77/.77/.77/.60 | .89/.94/.89/.94/.58 | .78/.78/.78/.78/.68 | .89/.79/.89/.84/.50 |
| | LR | .77/.75/.77/.74/.66 | .89/.90/.89/.87/.75 | .78/.61/.78/.68/.50 | .78/.78/.78/.78/.44 |
| | KNN | .77/.76/.77/.75/.71 | .89/.79/.89/.84/.50 | .67/.70/.67/.68/.67 | .89/.91/.89/.88/.83 |
| | DT | .82/.80/.82/.81/.63 | .89/.91/.89/.88/.83 | .67/.68/.67/.67/.68 | .78/.78/.78/.78/.75 |

## Appendix D

The following are the feedback questions posed to students after the review activity to gauge their opinions on the individualized intervention tutoring approach. Students indicated their agreement by using a 5-point Likert scale.

Q1. The teacher's summary activities for each concept addressed in class were helpful.
Q2. The programming remedial materials provided today will be helpful for exam preparation.
Q3. The programming remedial materials provided today helped me review concepts that I am not familiar with.
Q4. The programming remedial materials have helped me to review my knowledge of programming concepts, and these materials can help me improve my test scores.

## Appendix E

Regression models for learning characteristics, learning strategies, and learning performance.

| Independent variable: Pre-test of learning performance | | | | | | |
|---|---|---|---|---|---|---|
| Dependent variables | β value | Standard error | *t* | $R^2$ | Adjusted $R^2$ | *F* |
| | | | | .32 | .28 | $7.92^{***}$ |
| Constant | 60.62 | 4.34 | $13.98^{***}$ | | | |
| $f_3$ | .24 | .38 | $2.99^{**}$ | | | |
| $f_8$ | 1.15 | .12 | $2.11^{*}$ | | | |
| Independent variables: Learning feature $f_3$ (memo amount) | | | | | | |
| dependent variables | β value | standard error | *t* | $R^2$ | Adjusted $R^2$ | *F* |
| | | | | .25 | .22 | $11.01^{**}$ |
| Constant | -51.30 | 20.40 | $-2.52^{*}$ | | | |
| $LS_9$(Help seeking) | 18.30 | 5.52 | $3.32^{**}$ | | | |
| Independent variables: $LS_9$ (Help seeking) | | | | | | |
| dependent variables | β value | standard error | *t* | $R^2$ | Adjusted $R^2$ | *F* |
| | | | | .13 | .10 | $5.0^{*}$ |
| Constant | 2.57 | .50 | $5.16^{***}$ | | | |
| $LS_1$(Rehearsal) | .27 | .12 | $2.24^{*}$ | | | |

*Note.* $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

## Appendix F

Descriptive statistics of SRL ability pretest and learning performance posttest for high/medium/low groups for the rehearsal, elaboration, organization, and critical thinking strategies

| | | Number | | Mean/*SD* SRL ability pre-test | | Mean/ *SD* Learning performance post-test | |
|---|---|---|---|---|---|---|---|
| | group | $G_C$ | $G_E$ | $G_C$ | $G_E$ | $G_C$ | $G_E$ |
| Rehearsal | $Re_L$ | 17 | 6 | 3.53/.31 | 3.33/.30 | 87.35/5.89 | 9.00/5.48 |
| | $Re_M$ | 20 | 19 | 3.97/.18 | 3.96/.22 | 85.75/11.50 | 91.05/7.56 |
| | $Re_H$ | 5 | 11 | 4.79/.25 | 4.88/.18 | 59.00/3.08 | 87.27/6.07 |
| | total | 42 | 36 | 3.91/.48 | 4.11/.58 | 83.21/15.65 | 89.72/6.86 |
| Elaboration | $Re_L$ | 10 | 19 | 3.67/.14 | 3.86/.26 | 85.50/1.66 | 89.21/7.31 |
| | $Re_M$ | 26 | 7 | 3.98/.20 | 4.00/.27 | 86.54/1.75 | 92.14/7.56 |
| | $Re_H$ | 6 | 10 | 1.85/.26 | 4.72/.25 | 65.00/27.57 | 89.00/5.68 |
| | total | 42 | 36 | 4.02/.41 | 4.11/.45 | 83.21/15.65 | 89.72/6.86 |
| Organization | $Re_L$ | 8 | 4 | 3.00/.29 | 3.13/.48 | 88.75/6.94 | 95.00/4.08 |
| | $Re_M$ | 30 | 20 | 3.82/.23 | 3.98/.21 | 85.00/11.22 | 89.25/7.48 |
| | $Re_H$ | 4 | 12 | 4.73/.32 | 4.75/.19 | 58.75/33.26 | 88.75/6.08 |
| | total | 42 | 36 | 3.76/.64 | 4.12/.56 | 83.21/15.65 | 89.72/6.86 |
| Critical Thinking | $Re_L$ | 19 | 11 | 3.28/.28 | 3.56/.32 | 84.74/11.48 | 88.64/8.09 |
| | $Re_M$ | 17 | 14 | 3.96/.12 | 3.96/.12 | 89.12/7.12 | 9.00/6.20 |
| | $Re_H$ | 6 | 11 | 4.72/.27 | 4.70/.21 | 61.67/26.20 | 9.45/6.88 |
| | total | 42 | 36 | 3.75/.56 | 4.05/.50 | 83.21/15.65 | 89.72/6.86 |

Liang, C., Horikoshi, I., Majumdar, R., Flanagan, B., & Ogata, H. (2023). Towards Predictable Process and Consequence Attributes of Data-Driven Group Work: Primary Analysis for Assisting Teachers with Automatic Group Formation. *Educational Technology & Society*, *26*(4), 90-103. https://doi.org/10.30191/ETS.202310_26(4).0006

# Towards Predictable Process and Consequence Attributes of Data-Driven Group Work: Primary Analysis for Assisting Teachers with Automatic Group Formation

**Changhao Liang[1*], Izumi Horikoshi[2], Rwitajit Majumdar[2], Brendan Flanagan[3] and Hiroaki Ogata[2]**

[1]Graduate School of Informatics, Kyoto University, Japan // [2]Academic Center for Computing and Media Studies, Kyoto University, Japan // [3]Center for Innovative Research and Education in Data Science, Kyoto University, Japan // liang.changhao.84c@st.kyoto-u.ac.jp // horikoshi.izumi.7f@kyoto-u.ac.jp // majumdar.rwitajit.4a@kyoto-u.ac.jp // flanagan.brendanjohn.4n@kyoto-u.ac.jp // hiroaki.ogata@gmail.com

[*]Corresponding author

(Submitted September 26, 2022; Revised April 6, 2023; Accepted April 13, 2023)

**ABSTRACT:** Data-driven platforms with rich data and learning analytics applications provide immense opportunities to support collaborative learning such as algorithmic group formation systems based on learning logs. However, teachers can still get overwhelmed since they have to manually set the parameters to create groups and it takes time to understand the meaning of each indicator. Therefore, it is imperative to explore predictive indicators for algorithmic group formation to release teachers from the dilemma with explainable group formation indicators and recommended settings based on group work purposes. Employing learning logs of group work from a reading-based university course, this study examines how learner indicators from different dimensions before the group work connect to the subsequent group work processes and consequences attributes through correlation analysis. Results find that the reading engagement and previous peer ratings can reveal individual achievement of the group work, and a homogeneous grouping strategy based on reading annotations and previous group work experience can predict desirable group performance for this learning context. In addition, it also proposes the potential of automatic group formation with recommended parameter settings that leverage the results of predictive indicators.

**Keywords:** Group work indicator, GLOBE, Correlation analysis, Group formation, CSCL, Group work prediction, Teacher assistance

## 1. Introduction

Group learning gets increasingly prevalent in pedagogical practice (Dinh et al., 2021) and prevalent online courses nowadays raise impetus to the demand for such interactive activities. Prevalent research on computer-supported collaborative learning (CSCL) (Stahl et al., 2006) and learning analytics (LA) (Siemens, 2012) bring about immense opportunities to scaffold group work nowadays. However, there are still obstacles that hinder teachers from using technical support due to unfamiliarity with digital systems and lack of learning data (Austin et al., 2010; Brusilovsky et al., 2015; van der Velde et al., 2021).

Current LA researches focus on LA tools during the orchestration phase of the group work that investigates the group dynamics for timely intervention or forecasting the learning outcomes (Van Leeuwen et al., 2014). The predictive analysis before the group work start is less discussed, which is equally meaningful for group work organization such as group formation (Wessner & Pfister, 2001).

With the accumulation of abundant learning log data, group formation systems using learning logs advent (Boticki et al., 2019; Liang et al., 2021). However, teachers still need to manually set the parameters to create groups, which can overwhelm them and take time to understand the meaning of each indicator. To simplify such work for teachers from complicated parameter selections, it is imperative to automatically recommend appropriate group formation indicators that are predictive of desirable group work performance. This study aims to present a step towards an automatic group formation system by analyzing the data of a reading-based university course with asynchronous forum group discussion so that the recommendations of parameter selection of similar group learning contexts can be made for teachers.

Throughout the correlation analysis, this study examines how learner indicators from different dimensions before the group work connect to the subsequent group work processes and consequences attributes. Based on the predictive learner indicators, our findings can help teachers with the detection of endangered students and group formation strategies in reading-based learning contexts. Meanwhile, it also contributes to the development of automatic group formation systems in a data-rich environment.

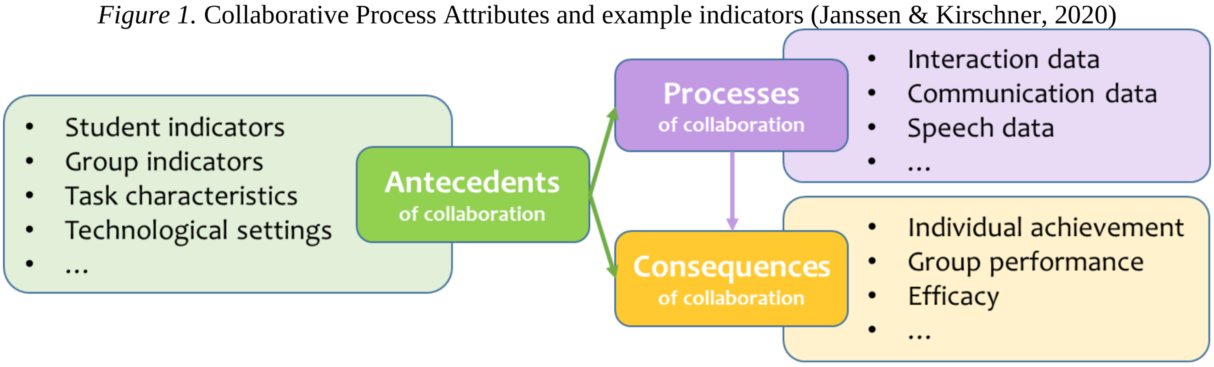
The following sections will first discourse on the research background based on related studies. Then we present correlation analysis based on data from a university reading course to explore the predictive indicators for desirable learning outcomes, followed by the discussion of results and implications for automatic group formation system design and teacher assistance.

## 2. Research background

### 2.1. Group work attributes and indicators

When conducting group work in pedagogy contexts, multiple issues should be considered in different stages (Urhahne et al., 2010). To characterize these issues with data, multiple indicators are proposed which reveal certain aspects of group work. Janssen and Kirschner (2020) put forward the concept of Collaborative Process Attributes that depict collaboration in three constructs: antecedents, processes, and consequences (see Figure 1). Indicators of antecedent attributes can pose an effect on processes and consequences of collaboration. However, which antecedent attributes influence the process and consequences of collaboration more was less discussed in previous studies, though it can be not only instructive for system innovation on automatic grouping, but also assist teachers to set groups appropriately with assorted student model data. In a digital learning environment with abundant learning log data, many of these indicators are recorded as learner models that depict the learning characteristics of students (Brusilovsky et al., 2015).

*Figure 1.* Collaborative Process Attributes and example indicators (Janssen & Kirschner, 2020)




For antecedents of collaboration, Janssen and Kirschner (2020) presented several typical instances based on what it describes. Student and group indicators are frequently-discussed (Saqr et al., 2020) and are prone to vary from group work tasks. Student indicators encompass all domain-specific and domain-independent information and as quantified indicators (Boticki et al., 2019), which can be easily derived from student model attributes under data-driven infrastructures. For example, gender, previous knowledge and task experience, preferences of learning styles, and personalities can be enveloped in the student indicators of group work (Savicki et al., 1996; Abnar et al., 2012; Zheng & Pinkwart, 2014; Sánchez et al., 2021). Group indicators describe characteristics of groups such as group size and intimacy (Amason & Sapienza, 1997; Huckman et al., 2009). Meanwhile, the heterogeneity distribution of student indicators within one group was also highlighted (Xu et al., 2020; Liang et al., 2021), which is closely connected to data-driven algorithmic group formation.

Processes of collaboration are an important part of CSCL research (Strode et al., 2022) since they can offer a holistic picture of the collaborative process that records the evidence during group work. The communication data, no matter in form of oral utterance (Liang et al., 2021) or online forums (Fidalgo-Blanco et al., 2015), assumes widely-used group learning evidence in related studies. Timeline sequence modeling, social network analysis (SNA), and epistemic network analysis (ENA) are conducted to further investigate the interaction data (Fidalgo-Blanco et al., 2015; Hoppe et al., 2021; Kaliisa et al., 2022). Using these interaction data during group work, it is feasible to use machine learning techniques to predict group performance (Cen et al., 2016). However, these data get available only when the current group work has started and the groups have been created.

Consequences of collaboration disclose the outcome of collaborative learning (Janssen & Kirschner, 2020). On the one hand, individual achievement estimates how much one has learned throughout the group work, especially for cognitive skills and knowledge acquisition. On the other hand, group performance is another indicator of collaboration quality, which can include the scores of group presentations and collaboratively composed reports.

Related research investigated the impact of specific student and group indicators in controlled experiments. For instance, previous knowledge and task experience proved to be closely related to group work performance in a collaborative programming context (Rentsch & Klimoski, 2001; Hsu et al., 2021). Similarly, Xu et al. (2020) also found the education level and domain knowledge of users can interactively predict users' knowledge gained in collaborative web searching sessions. In parallel, the heterogeneity of a group also affects the group work performance (Sánchez et al., 2021), and the impact of group heterogeneity can be different depending on the learning context (Liang et al., 2022b).

However, current studies seldom address these indicators from a comprehensive data-driven perspective, and it remains unclear which indicators of antecedent attributes get more predictive of the processes and consequences of collaborative learning in a certain learning context. Under the data-rich group work support system, Liang et al. (2022a) strategized a step to address this issue and provided explainable factors to teachers, while the group-level indicators were not incorporated. As it is significant to consider both student indicators and group indicators simultaneously according to Cress (2008), further studies remain imperative to detect the impact of group compositions.

### 2.2. Group formation based on student model data

Group formation is important since it can determine the quality of group work (Wessner & Pfister, 2001), and it was also found that collaborative learning with properly formed groups outperforms traditional teaching methods (Kyndt et al., 2013). However, creating collaborative learning groups remains challenging in CSCL studies due to the unfamiliarity of students and time-consuming procedures. Teachers can also get stuck due to little exposure to the CSCL tools in their daily routines. When creating groups, we need to determine three issues: the characteristics of group members, the context of the group work, and the group formation techniques according to Maqtary et al. (2019).

The characteristics of students lay the foundation to perform group formation algorithms. These student characteristics correspond to the antecedent attributes in the previous section and can be acquired in online learning platforms where multiple learning log data are accumulated. In the data-rich environment, student model data makes it possible to take student characteristics into account when creating groups.

The context is important as well since the optimal settings of group formation can differ from the purpose and traits of group work activity. For example, learning with peer help calls for heterogeneity of knowledge level, while homogeneous groups perform better in situations that encourage interaction and familiarity of group mates.

Based on different student model data and purposes, manifold techniques were employed for learning group creation. Clustering techniques underpinned by distance measurements are used for homogeneous groupings, such as the K-means algorithm that puts students in the same cluster together in the mobile learning context (Maqtary et al., 2019). In cases where students created abundant learner-generated content, the semantic method can group students (Isotani et al., 2009) based on textual features in terms of knowledge diversity, textual similarity as well as a semantic network of learner's interaction texts (Yoshida et al., 2023). It is hard to express the heterogeneity of groups under the semantic matchmakers in comparable values (Konert et al., 2014).

To deal with group formation from multiple student attributes, Moreno et al. (2012) put forward a genetic algorithm (GA) that can generate different group compositions (heterogeneous or homogeneous) in light of the calculated fitness values. The fitness values can be estimated by distance measures of vectors such as the sum of the squared differences (Moreno et al., 2012), which can reflect the heterogeneity of the student characteristics. In this way, homogeneous groups consisting of similar group members, or heterogeneous groups with dissimilar group members can be determined. The genetic algorithm presents flexibility owing to the fitness functions that can be adjusted to meet various grouping purposes and accommodate assorted input variables as was discussed in Flanagan et al. (2021) and Revelo et al. (2021). Liang et al. (2021) presented a group formation system that enables student models from different data sources underpinned by genetic algorithms and LEAF infrastructure (Flanagan & Ogata, 2017; Ogata et al., 2023) that aggregates multiple learning logs.

GroupAL is another relative project for group formation using a similar technique of vector optimization as GA (Konert et al., 2014). The GroupAL algorithm also provides flexible settings of parameters and criteria (heterogeneous or homogeneous) to meet different learning scenarios. Similar to the fitness function in GA, the optimal group allocation also relies on the defined metrics that depict the distance among participants and pairwise disjoint groups. However, without multiple iterations implemented in GA, GroupAL assigns

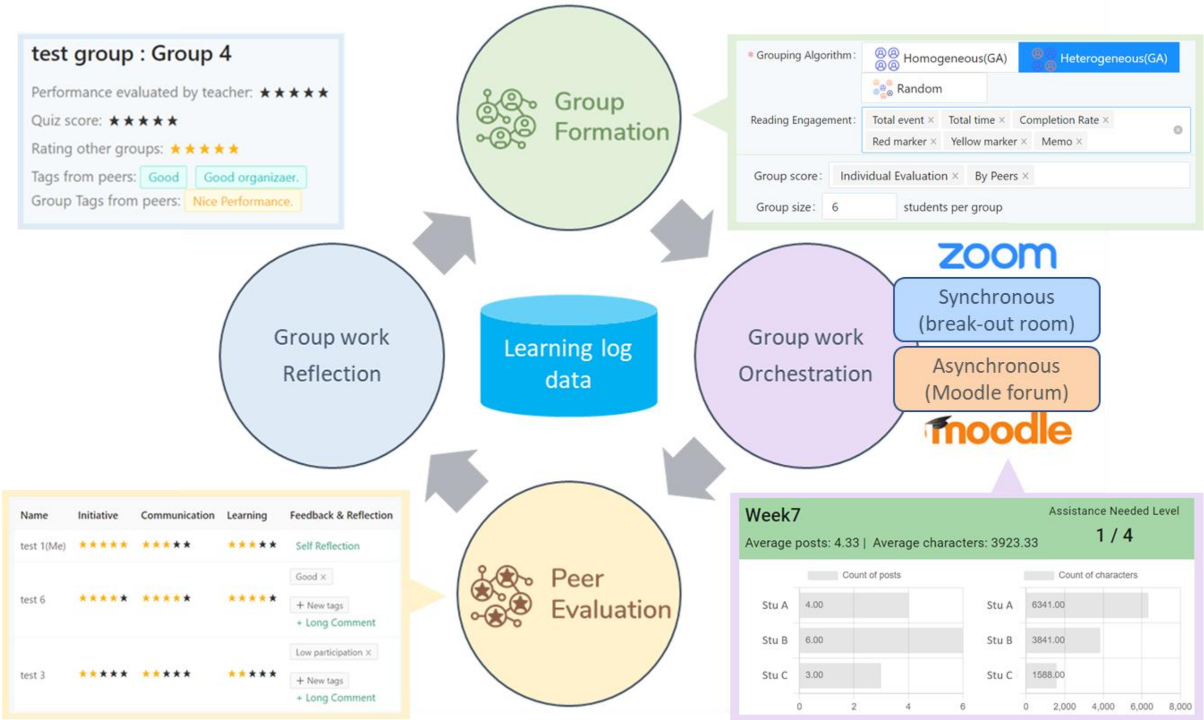
participants to learning groups only once. Under the same criteria and parameter settings, both GroupAL and GA can make different cohorts of groups since both approaches start from a randomized group allocation. Further, there were efforts of data integration to derive data from e-learning systems such as MoodlePeers as extensions of the GroupAL project (Konert et al., 2016).
In previous studies, the impact of algorithmic group formation using several student model indicators with heterogeneous or homogeneous compositions were investigated. However, which indicators play a more significant role to elicit desirable outcomes still deserve further inspection.

### 2.3. Continuous data-driven environment for group work conduction

The division of antecedents, processes, and consequences is not absolute since the previous learning logs reflecting the processes and consequences of collaboration can be employed as antecedents in the next round of group work. Group Learning Orchestration Based on Evidence (GLOBE) presents a data-driven environment (Liang et al., 2021) that enables such data re-usage. It also integrates group work-related learning logs from different sources to scaffold group work, hence suggesting further opportunities to explore the relationships among indicators of group work.

GLOBE utilizes data from Learning and Evidence Analytics Framework (LEAF), an overall technical framework integrating research and production systems to support learning analytic research as well as AI-driven services for effective teaching-learning (Ogata et al., 2023). The data covers learning records from learning management systems (LMS) such as Moodle, and reading interaction logs from learning material distribution platform BookRoll (Ogata et al., 2015).

*Figure 2.* GLOBE framework with continuous data-driven modules (Liang et al., 2021)




As illustrated in Figure 2, data-driven support consists of four phases: group formation, orchestration, evaluation, and reflection. An algorithmic group formation system, a forum discussion dashboard, and a peer evaluation system actuate the GLOBE framework utilizing the learning logs. In the group formation phase, teachers can create groups using multiple indicators, which play the roles of antecedents of collaboration. Then in the orchestration phase, the teacher can check the group work process in the dashboard and give timely interventions. For evaluation, both the teacher and students can give ratings in the evaluation module and check real-time feedback for reflection.

Table 1 lists the current indicators used in GLOBE with its pedagogical implications (the proxy for construct) in related studies. These indicators can be divided into several categories. Reading attributes from BookRoll reading logs (Ogata et al., 2015) can reflect learning engagement and also active reading behaviors (Toyokawa et al., 2021). These reading engagement data can also predict individual learning outcomes (Junco & Clem, 2015; Chen et al. 2021; Yang et al., 2021). Forum attributes talk about the interaction aspects in online group work, which covers passive participation with only view behaviors and positive participation indicated by post behaviors (Fidalgo-Blanco et al., 2015). The ratings from teachers and peers assess the performance of the group presentation and participation within each group, which also suggest the experience of collaborative learning as indicators of subsequent activities (Liang et al., 2022). Finally, the scores from external sources such as test scores and final course grades uploaded to the learning management systems (LMS) are of equal significance as an estimation of the individual learning outcome.

*Table 1*. Group work indicators in GLOBE systems

| Indicators in GLOBE | Description | Collaborative process attributes | Data source | Proxy for construct (what does it convey) |
|---|---|---|---|---|
| *Reading time | Total time spent on the e-book reader | Antecedent | BookRoll | Reading engagement that can predict learning achievement and academic performance (Junco & Clem, 2015; Chen et al., 2021; Yang et al., 2021) |
| *Operation times | Total number of operation times in the e-book reader (e.g., flipping page) | Antecedent | BookRoll | |
| *Completion rate | Percentage of completion of the reading material | Antecedent | BookRoll | |
| *Red markers | Number of annotations of important parts in the reading material | Antecedent | BookRoll | Active reading skills (Khusniyah & Lustyantie, 2017; Toyokawa et al., 2021) |
| *Yellow markers | Number of annotations of difficult parts in the reading material | Antecedent | BookRoll | |
| *Memos | Number of memos in the reading material | Antecedent | BookRoll | |
| Forum views | Times of views in the forum | Process | Moodle forum | Engagement and active interactions (Fidalgo-Blanco et al., 2015) |
| Forum posts | Number of posting messages in the forum | Process | Moodle forum | |
| Forum characters | Number of characters in all the forum posts | Process | Moodle forum | |
| *Teacher's ratings | Teacher's rating scores of the group presentation | Antecedent & Consequence | Evaluation module | Group work experience and task experience |
| *Peer ratings (individual) | Peer rating scores of group members within the group | Antecedent & Consequence | GLOBE | |
| *Peer ratings (group) | Peer rating scores of presentations from other groups | Antecedent & Consequence | GLOBE | |
| Course scores | Test/quiz scores and final grades that reflect the learning outcome | Antecedent & Consequence | LMS | Academic performance and learning outcome |

*Note*. *Heterogeneity of this indicator as an antecedent attribute within a group can be calculated by the squared differences (Flanagan et al., 2021) as a group-level indicator.

## 3. Method

To investigate the impact of each antecedent attribute, we run correlation analysis using an online reading course under the LEAF and GLOBE infrastructure. The study aims to detect the relationship between the antecedent attributes and that in the subsequent phases (processes and consequences), which can be utilized to assist teachers to create groups with a recommendation of optimal group formation settings.

We conducted a single group study with a pulled-in dataset of one university course. During the weekly learning activities in the online learning platforms, their Collaborative Process Attributes were anonymously recorded in the data repository of GLOBE. This study aims to find optimal predictors for desirable group work by analyzing the correlation of the antecedents with processes and consequences attributes of collaboration. The overarching research questions of this study are as follows:

- RQ1: What are the associations among individual-level indicators in different Collaborative Process Attributes?
- RQ2: What are the associations among group-level indicators in different Collaborative Process Attributes?

### 3.1. Research context and participants

The dataset came from a university course “Readings in Humanities and Social Sciences: Education Technology and AI” in Japan in the academic year 2022. On completing this course, students should understand the structure and expressions in academic articles. The course also allowed students to improve their English reading and presentation skills. Weekly reading and group work activities were implemented under the LEAF and GLOBE infrastructure. The course collected abundant data on Collaborative Process Attributes, thus producing enough data samples from real-world settings with routine practices. Hence it holds generalizability (Maissenhaelter et al., 2018) and convenience for extraction of evidence in further analysis (Kuromiya et al., 2020). Thirty-two (32) students registered for the course at the beginning, with 7 students withdrawing midway. 25 students finished the whole course and got a final course grade. 19 students came from the Faculty of Engineering, 3 students came from the Faculty of Integrated Human Study, and the remaining 3 students majored in Pharmacy, Economics, and Science respectively. There were 17 sophomores, 5 junior students, and 3 senior students among the participants.

In this course, group work was conducted several times from week 3 to week 11 across the 15-week semester. Following the GLOBE framework, students were grouped five times by the group formation system (Liang et al., 2021) across the course with different group formation indicators for different academic reading topics (see Table 2).

*Figure 3.* Workflow of the weekly activity implemented in the course

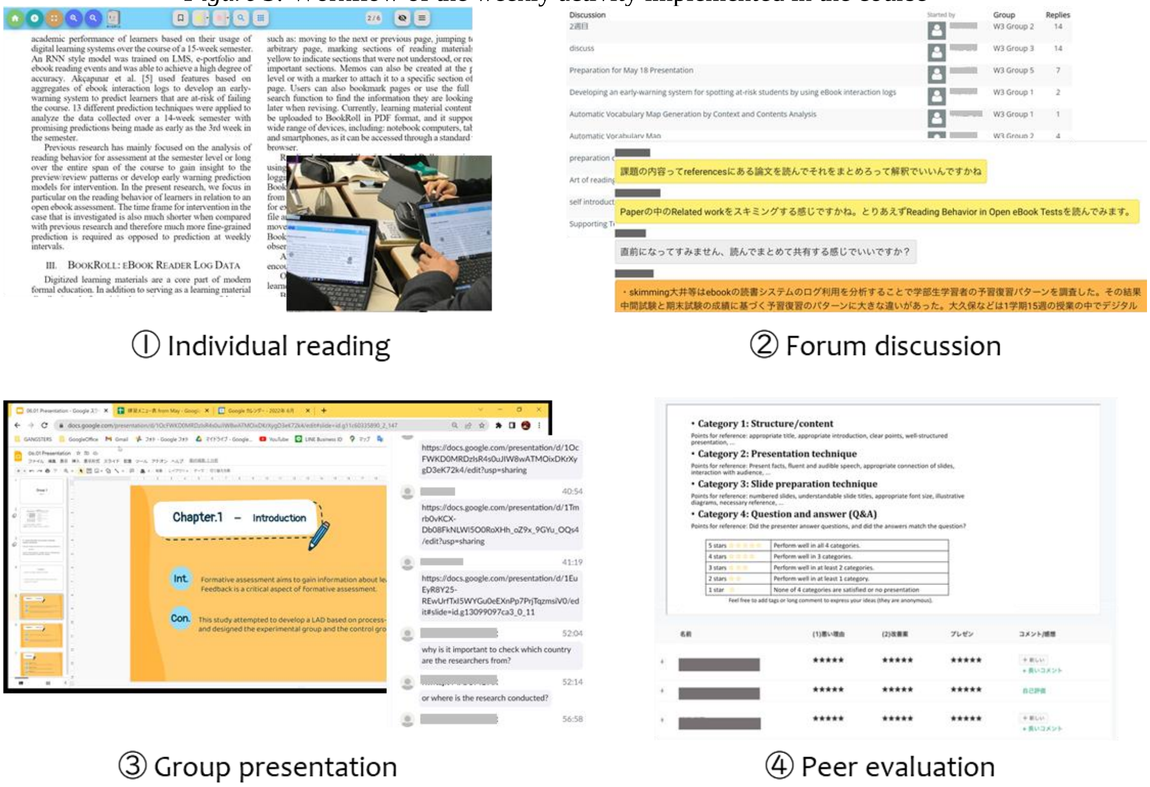



Figure 3 shows the workflow of the weekly activity implemented in the course. For each week, students were required to read several articles on BookRoll, an e-book reading system (Ogata et al., 2015) that can automatically collect learning data. Then, they should share and discuss their reading progress with their group members in the Moodle forum and prepare a brief presentation as a group for the next offline class. During the class, each group made presentations, which were peer-evaluated by the audience (both the instructor and students) in the classroom in the evaluation systems (Liang et al., 2021). In the meantime, asked to make peer ratings on the initiative and communication of their group mates in the peer evaluation system for each week as well.

*Table 2.* Group formation and group work topics in the course

| | Input attributes | Group work topic | # of students | # of groups |
|---|---|---|---|---|
| Week 3-4 | Reading engagement | Fast overview & reading strategy | 32 | 6 |
| Week 5-7 | Reading engagement & previous group ratings and peer ratings | Related work & review design | 32 | 5 |
| Week 8-9 | Reading engagement | Keywords & systematic survey | 25 | 5 |
| Week 10 | Reading engagement & previous group ratings and peer ratings | Using group graphs | 26 | 4 |
| Week 11 | Reading engagement & previous group ratings and peer ratings | Using group graphs & self-directed learning tools | 26 | 4 |

### 3.2. Data collection

The data of 8 group work in 5 group compositions were pulled in for analysis since all of the group work followed the same procedure and identical rating rubrics. The individual indicators of antecedent attributes and process attributes were standardized into the range of 0 to 1 for the group formation input. For group-level indicators, antecedent attributes were estimated by the squared differences (Flanagan et al., 2021) as heterogeneity, and average scores were calculated for some process and consequence attributes (forum posts, forum characters, peer ratings of initiative, and peer ratings of communication). Table 3 summarizes all these indicators involved in the study.

*Table 3.* Indicators used in this study

| Indicator | *N* | Mean | Max | Min |
|---|---|---|---|---|
| **Antecedents** | | | | |
| Reading time | 199 | 0.658 | 1 | 0.08 |
| Operation times | 199 | 0.653 | 1 | 0.06 |
| Completion rate | 199 | 0.483 | 0.65 | 0.05 |
| Red markers | 199 | 0.532 | 1 | 0 |
| Yellow markers | 199 | 0.551 | 1 | 0 |
| Memos | 199 | 0.384 | 1 | 0 |
| *Heterogeneity of reading time | 46 | 0.250 | 0.461 | 0.078 |
| *Heterogeneity of operation times | 46 | 0.239 | 0.409 | 0.035 |
| *Heterogeneity of completion rate | 46 | 0.123 | 0.218 | 0 |
| *Heterogeneity of red markers | 46 | 0.347 | 0.489 | 0.078 |
| *Heterogeneity of yellow markers | 46 | 0.335 | 0.526 | 0.064 |
| *Heterogeneity of memos | 46 | 0.388 | 0.509 | 0 |
| Previous teacher’s ratings | 121 | 0.876 | 1 | 0.6 |
| Previous peer ratings (individual) | 109 | 0.738 | 1 | 0.2 |
| Previous peer ratings (group) | 121 | 0.786 | 0.9 | 0.629 |
| *Heterogeneity of previous teacher’s ratings | 26 | 0.098 | 0.121 | 0.031 |
| *Heterogeneity of previous peer ratings (individual) | 26 | 0.243 | 0.312 | 0.076 |
| *Heterogeneity of previous peer ratings (group) | 26 | 0.091 | 0.132 | 0.017 |
| **Processes** | | | | |
| Forum posts | 114 | 0.301 | 0.99 | 0 |
| Forum characters | 114 | 0.353 | 0.99 | 0 |
| **Consequences** | | | | |
| *Teacher’s ratings | 46 | 4.413 | 5 | 3 |
| Peer ratings of initiative | 199 | 3.658 | 5 | 0.5 |
| Peer ratings of communication | 199 | 3.461 | 5 | 1 |
| *Peer ratings (group) | 46 | 4.055 | 4.667 | 2.857 |
| Final course grades | 25 | 69.8 | 100 | 30 |

*Note.* *Group-level indicators.

### 3.3. Data analysis

We used correlation analysis and calculated the Pearson correlation coefficient for each pair of antecedent-process and antecedent-consequence. To deal with missing values (e.g., in weeks 3-4 and 8-9, previous group ratings and peer ratings as antecedent attributes were not used for group formation), we exclude cases pairwise before the analysis.

According to the research questions, we investigate two levels of indicators in this study. For individual-level indicators, we inspect the correlation among values. Positive relations denote that the higher score of an indicator one possesses, the more predictive of the desired learning outcome this indicator can be, and vice versa. Insignificant correlation means low predictive power in this learning context.

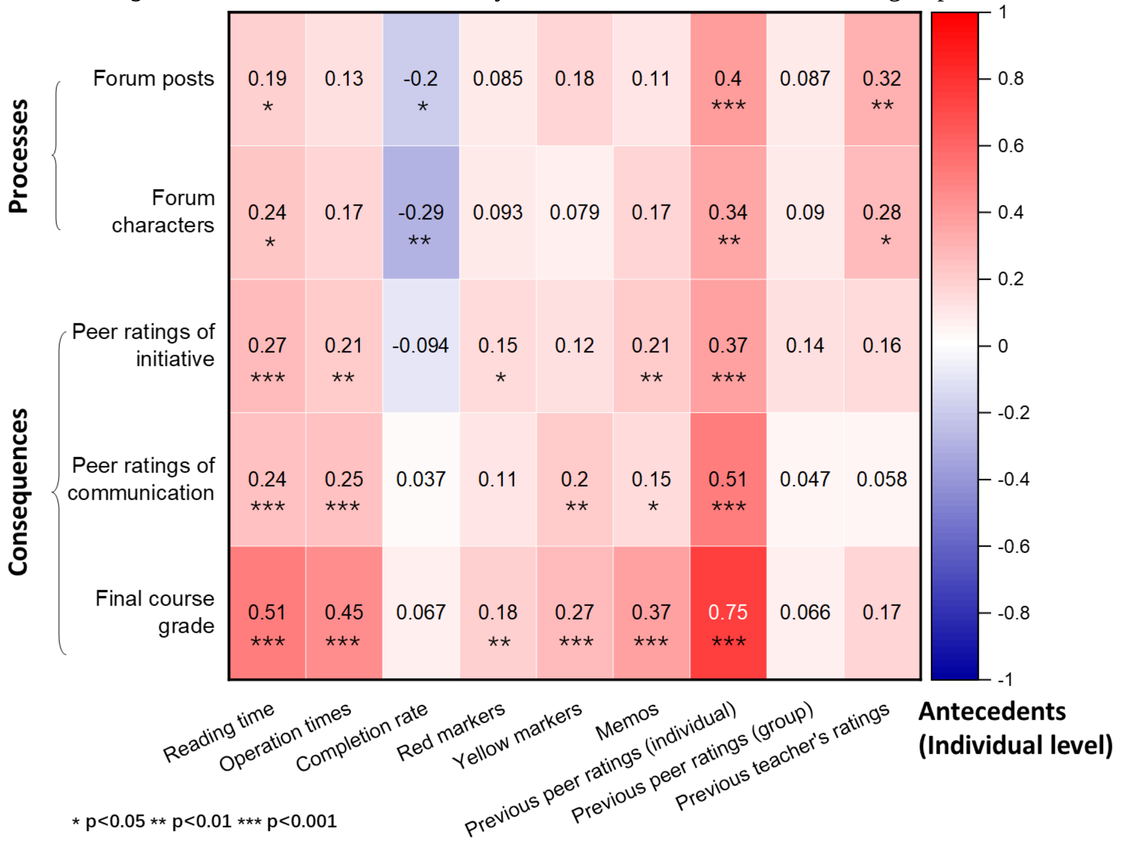
For group-level indicators, we examined their correlations with the group-level indicators of processes and consequences attributes that were calculated by aggregation of each group. The heterogeneity of each indicator as an antecedent attribute within a group is calculated by the squared differences, which are also used in the group formation algorithm to measure the heterogeneity of each group as the fitness function (Flanagan et al., 2021). As for the indicator of heterogeneity, the positive relation coefficient suggests the more heterogeneous the values of a certain indicator within a group, the better performance this group will have. On the contrary, negative correlations connote the more homogeneous the values of a certain indicator in a group, the more desirable the group-level outcome will be.

## 4. Results

### 4.1. Individual-level indicators

Figure 4 is the correlation diagram of individual-level indicators. As can be seen in the diagram, reading time and previous peer ratings for individuals show significant positive associations to all processes and consequences attributes. The association between previous peer ratings for individual and final course grades is strong (> 0.7). Operation times and the number of memos have significant positive correlations to all three consequence attributes, but their associations to process attributes are not found. Conversely, previous teachers' ratings of group work are related to the individual performance of two processes attributes, but not associated with all individual-level consequence scores. Both red markers and yellow markers take close relations to the final course grade. In addition, red markers show a weak significant association with initiative scores of peer ratings while yellow markers are weakly associated with communication scores of peer ratings. The completion rate connotes a weak adverse connection to the process attributes of forum utterance in this study and no significant correlation with all three consequence attributes. Meanwhile, previous peer ratings of group presentations indicate no significant relationship to any individual-level indicators.

*Figure 4.* Results of correlation analysis of individual-level indicators of group work

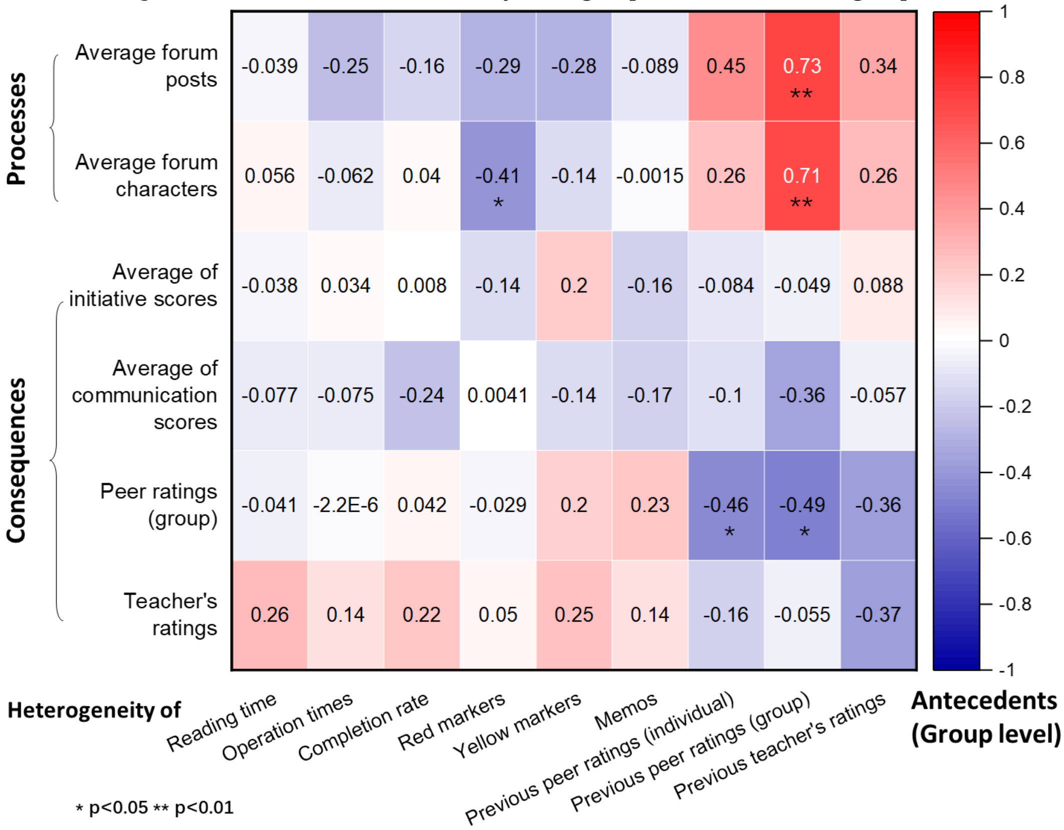


### 4.2. Group level indicators

Figure 5 illustrates the results of correlation analysis for group-level indicators. As a result, positive and strong associations are found between (1) heterogeneity of previous peer ratings (group) and average forum posts and (2) heterogeneity of peer ratings (group) and average forum characters. This means that the heterogeneous composition of these antecedent attributes can contribute to the performance of the group work processes.

Negative correlations are revealed between (1) heterogeneity of red markers and average forum characters, (2) heterogeneity of previous peer ratings (individual) and the peer ratings received of the current group presentation, and (3) heterogeneity of previous peer ratings (group) and the peer ratings received of the current group presentation. These three correlations are moderate. This denotes the potential of the homogeneous composition on these antecedent attributes to scaffold the performance of the group work processes. Apart from the former results, all other correlations are insignificant in statistics.

*Figure 5*. Results of correlation analysis of group-level indicators of group work




## 5. Discussion

### 5.1. Individual-level indicators and individual performance

Compared to the previous study, most correlations in this study remain the same with Liang et al. (2022a). The reading time and previous peer ratings received are still the most predictive indicators that suggest a significant positive correlation with all processes attributes of forum engagement and consequences attributes of peer ratings as well as the final course grade. These results are also in accord with Junco and Clem (2015) and Chen et al. (2021) that found reading time is predictive of the individual learning outcome. The active reading indicators such as memos and markers are also positively associated with desirable learning consequences as Yang et al. (2021) presented. In parallel, the reliability of peer ratings under the peer evaluation system can be approved as well, suggesting that students of the online university course can give a fair assessment to their peers based on rubrics. However, the completion rate showed an adverse association with the forum engagement indicators. This can be caused by the pull-in operation when we aggregate data. Since this study used all data from the course, the overall completion rate got lower due to the abundant reading materials as can be seen in the descriptive statistics in Table 3.

We can also find that, as for previous peer ratings and the teacher's ratings of group presentations, the predictive power is relatively low in that only the teacher's ratings of group presentations have a weak correlation to the forum engagement indicators. Since these two ratings are group-level assessments of previous group work, their reliability can be reduced by social loafing and free riding (Forsell et al., 2020), which can elicit less predictive power when modeling each individual using such scores. Apart from this, it also shows the necessity for analysis of group-level indicators as was mentioned by Cress (2008).

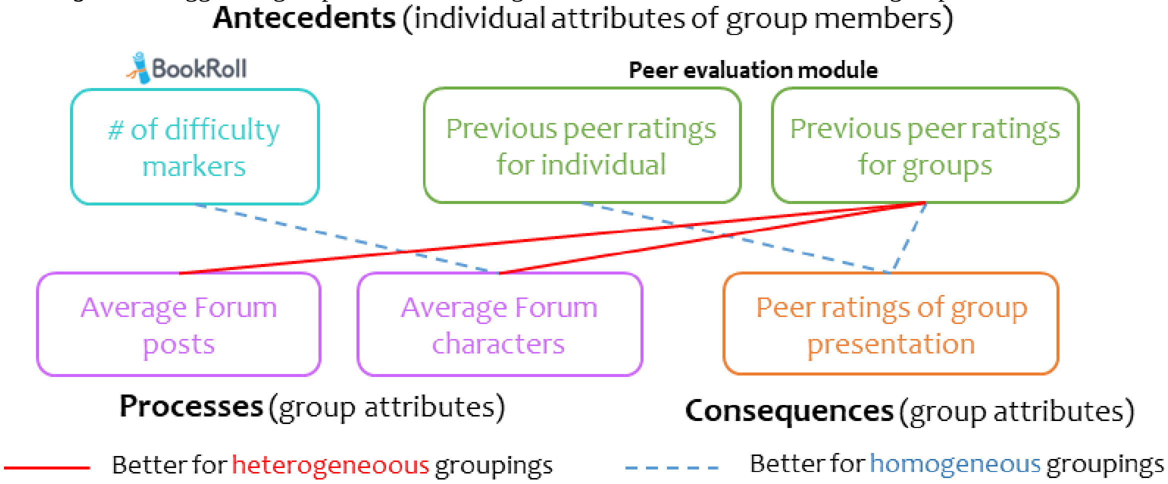
In sum, in the context of reading-based group work, the reading engagement attributes and peer ratings received in previous group work that indicates group work experience are closely connected to the individual performance of subsequent forum discussions and learning outcomes, which can guide group formation settings and intervention suggestions in the similar context such as language learning and academic reading. In parallel, from the participation of reading and previous group work performance, teachers can take timely measures to help these endangered students predicted by the GLOBE system (Liang et al., 2022b).

### 5.2. Group level indicators on group work performance

The group-level analysis focuses on the heterogeneity of each antecedent attribute within each group and aims to explore group dynamics. First, the average forum engagement of a group indicated by posts and characters is strongly positively correlated with its heterogeneity of previous group performance rated by peers. While no correlation was detected at the individual level between these two indicators. These findings support the strategy to heterogeneously group students so that we can guarantee that at least one outperforming student with desirable previous group work experience is assigned to each group, thus avoiding absolute silence in groups with all underperforming students. Such a positive effect of heterogeneous strategy on previous performance indicators agrees with group work in the classroom scenario as well (Liang et al., 2022).

As for annotation data that indicate the records of active reading strategy, we found the groups with more homogeneous red markers indicating highlights tend to have more forum discussions, though for individuals more markers did not indicate more posts. As an indicator of active reading engagement, the effect of grouping students with homogeneous engagement levels agrees with the other research on online courses and MOOCs (Abou-Khalil & Ogata, 2021; Sánchez et al., 2021), which can be explained by reduced social loafing for lack of proactive students to count on (Wichmann et al., 2016). Furthermore, the homogeneous grouping can be more promising when considering the annotated contents, since students with common annotations can show joint interest that can facilitate the interaction of the participants (Toyokawa et al., 2021).

*Figure 6.* Suggested group formation strategies based on correlations between group-level attributes



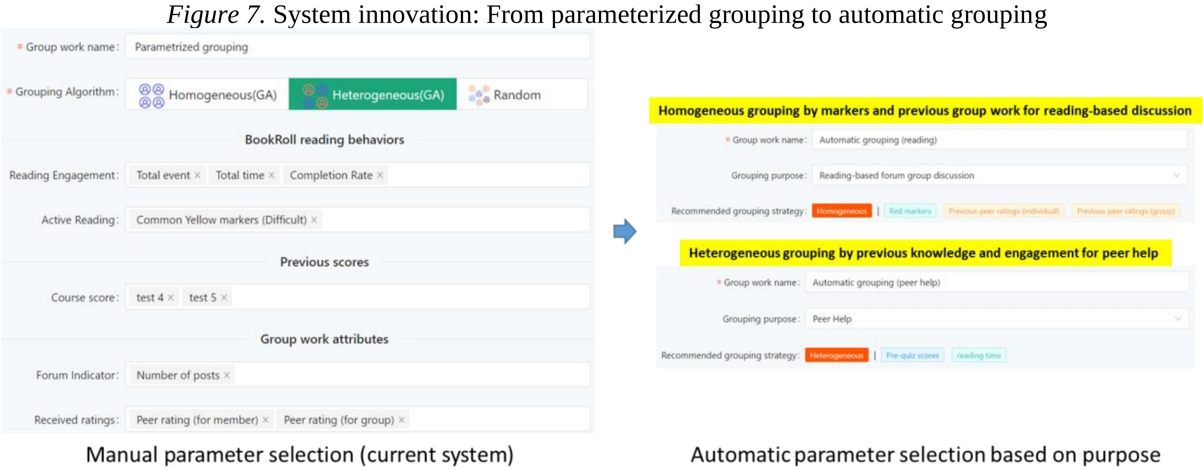

Another finding that deserves our attention is that the heterogeneity of previous ratings, both for individuals and groups, are of moderate negative related to the peer rating scores of the final group presentation. The result denotes that though a group with heterogeneity in the previous group experience tends to have more discussion and engagement when it comes to the cooperative for a group-level output, it can become hard to reach a consensus, thus resulting in undesirable performance on group presentations. The heterogeneous groups with unbalanced knowledge of the task encourage peer help that facilitates individual achievement (Kanika et al., 2022), but it may not contribute to the cooperation and synergistic output of a group. To figure out the reason, further analysis of forum discussions is required to investigate the relationship between processes and consequences indicators of group work in the orchestration phase of GLOBE.

According to our primary analysis (see Figure 6), we have identified appropriate group formation strategies for teacher assistance in the data-driven environment of LEAF. A homogeneous grouping strategy, considering the number of difficulty markers and previous peer ratings, has the potential to enhance the number of forum characters and peer ratings of group presentation. This finding provides guidance for subsequent group formation in the context of active reading-based group work. On the other hand, heterogeneous grouping based on previous peer ratings for groups can facilitate more detailed forum discussions with more characters in forum posts. This strategy can be useful for online courses where online reading and forum discussion are closely connected.

### 5.3. Automatic group formation with optimal indicators to assist teachers

For technical implications, the research provides supportive evidence for the innovation of the current group formation system. Although we only addressed reading-based group discussions herein, similar research on other contexts can be done in the same way under the GLOBE framework. As is shown in Figure 7, teachers have to manually choose multiple indicators when creating groups currently. With the accumulation of evidence from studies on the predictive antecedent in different learning contexts, the strengthened system can automatically select input parameters based on the selected learning purpose and context in the future. For example, homogeneous algorithms with red markers and previous peer ratings are suggested based on the results of this study, as they are associated with better group performance. Conversely, in contexts that underscore individual learning with peer help design, heterogeneous groups with reading engagement and test scores that indicate previous knowledge are recommended in the automatic grouping according to Liang et al. (2021) and Liang et al. (2022b).

*Figure 7.* System innovation: From parameterized grouping to automatic grouping




For pedagogical implications, a pivotal goal of this study is to help teachers to determine the optimal group formation indicators in data-driven digital systems. This study discloses predictive antecedent indicators to the performance of subsequent group work in a forum-supported academic reading course, which can guide teachers in similar contexts. The automatic group formation function will further release teachers from selecting assorted variables in the system and reduce the time for creating groups. Further studies to examine the effectiveness of the automatic grouping will become necessary then.

### 5.4. Limitations

The indicators incorporated in this study are still limited. Under the data-driven platforms, most of the indicators are from learner models that reflect learning-related characteristics, but the social-emotional indicators are less addressed in the current systems. These issues should also be addressed by uploaded scores and social network data as quantitative input for group formation. However, how to incorporate these data with different granularity and formats into the group formation algorithm remains unclear, and deserves future investigation. In parallel, the objective behavior data of previous group work was not used as the antecedent for the next round following the continuous data flow, which may reduce the reliability of previous group work performance indicators.

Meanwhile, though we got a larger sample size using pulled-in data of all group work throughout a semester in a university course compared to the previous study (Liang et al., 2022a), the learning context is confined to reading-based tasks with asynchronous forum discussions. Hence the predictive indicators in other learning conditions and cultures can vary. Therefore, the results of the current study need further validation in other learning scenarios.

## 6. Conclusion and future work

In conclusion, this study investigated the connections between antecedent attributes and the processes/consequences of group work in an asynchronous online reading course. We considered both individual-level and group-level indicators in the correlation analysis and found predictive indicators for algorithmic group formation. The reading engagement and previous peer ratings can reveal individual achievement of the group work, and a homogeneous grouping strategy based on reading annotations and previous group work experience can predict desirable group performance for this learning context. This study also provides avenues for future research to find predictive indicators in more learning contexts, and in turn, orchestrate an automatic group formation system that can mitigate teachers' trivial work from manual grouping. Meanwhile, how to make the antecedent indicators of groups created by algorithms explainable to teachers with adequate illustrations also deserves further consideration.

## Acknowledgement

This work was partly supported by JSPS KAKENHI 20K20131, 22H03902 and 20H01722; NEDO JPNP20006 and JPNP18013 and SPIRITS 2020 of Kyoto University and Support for Pioneering Research Initiated by the Next Generation program operated by the Japan Science and Technology Agency (JST) JPMJSP2110.

## References

Abnar, S., Orooji, F., & Taghiyareh, F. (2012). An Evolutionary algorithm for forming mixed groups of learners in web based collaborative learning environments. In *2012 IEEE international conference on technology enhanced education (ICTEE)* (pp. 1-6). IEEE. https://doi.org/10.1109/ICTEE.2012.6208612

Abou-Khalil, V., & Ogata, H. (2021). Homogeneous student engagement: A Strategy for group formation during online learning. In *International Conference on Collaboration Technologies and Social Computing* (pp. 85-92). Springer, Cham.

Amason, A. C., & Sapienza, H. J. (1997). The Effects of top management team size and interaction norms on cognitive and affective conflict. *Journal of Management*, *23*(4), 495-516.

Austin, R., Smyth, J., Rickard, A., Quirk-Bolt, N., & Metcalfe, N. (2010). Collaborative digital learning in schools: Teacher perceptions of purpose and effectiveness. *Technology, Pedagogy and Education*, *19*(3), 327-343.

Boticki, I., Akçapınar, G., & Ogata, H. (2019). E-book user modelling through learning analytics: The Case of learner engagement and reading styles. *Interactive Learning Environments*, *27*(5-6), 754-765.

Brusilovsky, P., Somyürek, S., Guerra, J., Hosseini, R., Zadorozhny, V., & Durlach, P. J. (2015). Open social student modeling for personalized learning. *IEEE Transactions on Emerging Topics in Computing*, *4*(3), 450-461.

Cen, L., Ruta, D., Powell, L., Hirsch, B., & Ng, J. (2016). Quantitative approach to collaborative learning: Performance prediction, individual assessment, and group composition. *International Journal of Computer-Supported Collaborative Learning*, *11*(2), 187-225.

Chen, C. H., Yang, S. J., Weng, J. X., Ogata, H., & Su, C. Y. (2021). Predicting at-risk university students based on their e-book reading behaviours by using machine learning classifiers. *Australasian Journal of Educational Technology*, *37*(4), 130-144.

Cress, U. (2008). The Need for considering multilevel analysis in CSCL research: an appeal for the use of more advanced statistical methods. *International Journal of Computer-Supported Collaborative Learning*, *3*(1), 69-84.

Dinh, J. V., Schweissing, E. J., Venkatesh, A., Traylor, A. M., Kilcullen, M. P., Perez, J. A., & Salas, E. (2021). The Study of teamwork processes within the dynamic domains of healthcare: A Systematic and taxonomic review. *Frontiers in Communication*, *6*, 617928. https://doi.org/10.3389/fcomm.2021.617928

Fidalgo-Blanco, Á., Sein-Echaluce, M. L., García-Peñalvo, F. J., & Conde, M. Á. (2015). Using learning analytics to improve teamwork assessment. *Computers in Human Behavior*, *47*, 149-156.

Flanagan, B., & Ogata, H. (2017). Integration of learning analytics research and production systems while protecting privacy. In *The 25th International Conference on Computers in Education* (pp. 333-338). Asia-Pacific Society for Computers in Education.

Flanagan, B., Liang, C., Majumdar, R., and Ogata, H. (2021). Towards explainable group formation by knowledge map-based genetic algorithm. In *2021 International Conference on Advanced Learning Technologies (ICALT)* (pp. 370–372). https://doi.org/10.1109/ICALT52272.2021.00119

Forsell, J., Forslund Frykedal, K., & Hammar Chiriac, E. (2020). Group work assessment: Assessing social skills at group level. *Small Group Research*, *51*(1), 87-124.

Hoppe, H. U., Doberstein, D., & Hecking, T. (2021). Using sequence analysis to determine the well-functioning of small groups in large online courses. *International Journal of Artificial Intelligence in Education*, *31*(4), 680-699.

Huckman, R. S., Staats, B. R., & Upton, D. M. (2009). Team familiarity, role experience, and performance: Evidence from Indian software services. *Management science*, *55*(1), 85-100.

Hsu, T. C., Abelson, H., Patton, E., Chen, S. C., & Chang, H. N. (2021). Self-efficacy and behavior patterns of learners using a real-time collaboration system developed for group programming. *International Journal of Computer-Supported Collaborative Learning*, *16*(4), 559-582.

Isotani, S., Inaba, A., Ikeda, M., & Mizoguchi, R. (2009). An Ontology engineering approach to the realization of theory-driven group formation. *International Journal of Computer-Supported Collaborative Learning*, *4*(4), 445-478.

Janssen, J., & Kirschner, P. A. (2020). Applying collaborative cognitive load theory to computer-supported collaborative learning: Towards a research agenda. *Educational Technology Research and Development*, *68*(2), 783-805.

Junco, R., & Clem, C. (2015). Predicting course outcomes with digital textbook usage data. *The Internet and Higher Education*, *27*, 54-63.

Kaliisa, R., Rienties, B., Mørch, A. I., & Kluge, A. (2022). Social learning analytics in computer-supported collaborative learning environments: A Systematic review of empirical studies. *Computers and Education Open*, 100073. https://doi.org/10.1016/j.caeo.2022.100073

Kanika, Chakraverty, S., Chakraborty, P., & Madan, M. (2022). Effect of different grouping arrangements on students' achievement and experience in collaborative learning environment. *Interactive Learning Environments*, 1-13. https://doi.org/10.1080/10494820.2022.2036764

Khusniyah, N. L., & Lustyantie, N. (2017). Improving English reading comprehension ability through Survey, Questions, Read, Record, Recite, Review Strategy (SQ4R). *English language teaching*, *10*(12), 202-211.

Konert, J., Burlak, D., & Steinmetz, R. (2014). The Group formation problem: An Algorithmic approach to learning group formation. In *Open Learning and Teaching in Educational Communities: 9th European Conference on Technology Enhanced Learning, EC-TEL 2014, Graz, Austria, Proceedings 9* (pp. 221-234). Springer International Publishing.

Konert, J., Bellhäuser, H., Röpke, R., Gallwas, E., & Zucik, A. (2016). MoodlePeers: Factors relevant in learning group formation for improved learning outcomes, satisfaction and commitment in E-learning scenarios using GroupAL. In *Adaptive and Adaptable Learning: 11th European Conference on Technology Enhanced Learning, EC-TEL 2016, Lyon, France, Proceedings 11* (pp. 390-396). Springer International Publishing.

Kuromiya, H., Majumdar, R., & Ogata, H. (2020). Fostering evidence-based education with learning analytics. *Educational Technology & Society*, *23*(4), 14-29.

Kyndt, E., Raes, E., Lismont, B., Timmers, F., Cascallar, E., & Dochy, F. (2013). A Meta-analysis of the effects of face-to-face cooperative learning. Do recent studies falsify or verify earlier findings? *Educational research review*, *10*, 133-149.

Liang, C., Majumdar, R. & Ogata, H. (2021). Learning log-based automatic group formation: System design and classroom implementation study. *Research and Practice in Technology Enhanced Learning*, *16*(1), 1–22.

Liang, C., Majumdar, R., Horikoshi, I., Flanagan, B. & Ogata, H. (2022a). Exploring predictive indicators of reading-based online group work for group formation assistance. In *30th International Conference on Computers in Education* (pp. 642-647). Asia-Pacific Society for Computers in Education.

Liang C., Majumdar R., Nakamizo Y., Flanagan B. & Ogata H. (2022b). Algorithmic group formation and group work evaluation in a learning analytics-enhanced environment: Implementation study in a Japanese Junior High School. *Interactive Learning Environments*, 1-24. https://doi.org/10.1080/10494820.2022.2121730

Maissenhaelter, B. E., Woolmore, A. L., & Schlag, P. M. (2018). Real-world evidence research based on big data: Motivation—challenges—success factors. *Der Onkologe*, *24*, 378-389.

Maqtary, N., Mohsen, A., & Bechkoum, K. (2019). Group formation techniques in computer-supported collaborative learning: A Systematic literature review. *Technology, Knowledge and Learning*, *24*(2),169-190.

Moreno, J., Ovalle, D. A., & Vicari, R. M. (2012). A Genetic algorithm approach for group formation in collaborative learning considering multiple student characteristics. *Computers & Education*, *58*(1), 560-569.

Ogata, H., Majumdar, R., & Flanagan, B. (2023). Learning in the digital age: Power of shared learning logs to support sustainable educational practices. *IEICE Transactions on Information and Systems*, *E106-D*(2), 101-109. https://doi.org/10.1587/transinf.2022ETI0002

Ogata, H., Yin, C., Oi, M., Okubo, F., Shimada, A., Kojima, K., & Yamada, M. (2015). E-Book-based learning analytics in university education. In *the 23rd International Conference on Computers in Education (ICCE 2015)* (pp. 401-406). Asia-Pacific Society for Computers in Education.

Rentsch, J. R., & Klimoski, R. J. (2001). Why do 'great minds' think alike?: Antecedents of team member schema agreement. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, *22*(2), 107-120.

Revelo Sánchez, O., Collazos, C. A., & Redondo, M. A. (2021). Automatic group organization for collaborative learning applying genetic algorithm techniques and the big five model. *Mathematics*, *9*(13), 1578. https://doi.org/10.3390/math9131578

Sánchez, O. R., Ordóñez, C. A. C., Duque, M. Á. R., & Pinto, I. I. B. S. (2021). Homogeneous group formation in collaborative learning scenarios: An Approach based on personality traits and genetic algorithms. *IEEE Transactions on Learning Technologies*, *14*(4), 486-499.

Saqr, M., Nouri, J., Vartiainen, H., & Malmberg, J. (2020). What makes an online problem-based group successful? A Learning analytics study using social network analysis. *BMC medical education*, *20*(1), 1-11.

Savicki, V., Kelley, M., & Lingenfelter, D. (1996). Gender, group composition, and task type in small task groups using computer-mediated communication. *Computers in human behavior*, *12*(4), 549-565.

Siemens, G. (2012). Learning analytics: Envisioning a research discipline and a domain of practice. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 4-8). https://doi.org/10.1145/2330601.2330605

Stahl, G., Koschmann, T., & Suthers, D. D. (2006). Computer-supported collaborativelearning: An Historical perspective. In *Cambridge handbook of the learning sciences* (pp. 409–426). Cambridge University Press.

Strode, D., Dingsøyr, T., & Lindsjorn, Y. (2022). A Teamwork effectiveness model for agile software development. *Empirical Software Engineering*, *27*(2), 1-50.

Toyokawa, Y., Majumdar, R., Lecailliez, L., & Ogata, H. (2021). A Flipped model of active reading using a learning analytics-enhanced e-book platform. In *the 29th International Conference on Computers in Education (ICCE 2021)* (pp. 510-515). Asia-Pacific Society for Computers in Education.

Urhahne, D., Schanze, S., Bell, T., Mansfield, A., & Holmes, J. (2010). Role of the teacher in computer-supported collaborative inquiry learning. *International Journal of Science Education*, *32*(2), 221-243.

van der Velde, M., Sense, F., Borst, J., & van Rijn, H. (2021). Alleviating the cold start problem in adaptive learning using data-driven difficulty estimates. *Computational Brain & Behavior*, *4*(2), 231-249.

Van Leeuwen, A., Janssen, J., Erkens, G., & Brekelmans, M. (2014). Supporting teachers in guiding collaborating students: effects of learning analytics in CSCL. *Computers & Education*, *79*, 28-39.

Wessner, M., & Pfister, H. R. (2001). Group formation in computer-supported collaborative learning. In *Proceedings of the 2001 international ACM SIGGROUP conference on supporting group work* (pp. 24-31). https://doi.org/10.1145/500286.500293

Wichmann, A., Hecking, T., Elson, M., Christmann, N., Herrmann, T., & Hoppe, H. U. (2016). Group formation for small-group learning: Are heterogeneous groups more productive? In *proceedings of the 12th international symposium on open collaboration* (pp. 1-4). https://doi.org/10.1145/2957792.2965662

Xu, L., Zhou, X., & Gadiraju, U. (2020). How does team composition affect knowledge gain of users in collaborative web search? In *HT'20: Proceedings of the 31st ACM Conference on Hypertext and Social Media* (pp. 91-100). Association for Computing Machinery (ACM). https://doi.org/10.1145/3372923.3404784

Yang, C. C., Chen, I. Y., Akçapınar, G., Flanagan, B., & Ogata, H. (2021). Using a summarized lecture material recommendation system to enhance students' preclass preparation in a flipped classroom. *Educational Technology & Society*, *24*(2), 1-13.

Yoshida, M., Xiong, C., Liu, Y., & Liu, H. (2023). An Investigation into the formation of learning groups on social media and their growth. *Interactive Learning Environments, 31*(1). https://doi.org/10.1080/10494820.2020.1780268

Zheng, Z., & Pinkwart, N. (2014). A Discrete particle swarm optimization approach to compose heterogeneous learning groups. In *2014 IEEE 14th international conference on advanced learning technologies* (pp. 49-51). IEEE. https://doi.org/10.1109/ICALT.2014.24

McIntyre, N. A. (2023). Guest Editorial: Dynamic Accounts of Digital Divides: Longitudinal Insights into Inequitable Access to Online Learning. *Educational Technology & Society*, *26*(4), 104-107. https://doi.org/10.30191/ETS.202310_26(4).0007

# Guest Editorial: Dynamic Accounts of Digital Divides: Longitudinal Insights into Inequitable Access to Online Learning

**Nora A. McIntyre**
Southampton Education School, University of Southampton, United Kingdom // N.McIntyre@soton.ac.uk

**ABSTRACT:** Educational access is the first step to realising education as a universal human right. Yet, education remains a luxury for many. Contrary to expectations, the advent of online learning has not resolved this fundamental injustice as out-of-school learning and early drop-outs remain widespread phenomena. Rather, digital divides have hit hard in exacerbating social inequalities, with help from Covid-19. This special issue takes a critical look at online learning in its potential to rebalance and hinder universal educational access. It includes four papers focused on differing positions of disadvantage. Based on these papers, this special issue highlights that — although it is possible for disadvantaged learners to compensate for inequities — socioeconomic and infrastructural constraints often prevail, and will continue to unless critical changes are made to the educational ecosystem. Furthermore, the Special Issue calls for future work that takes process-level changes into account in order to generate recommendations that are better grounded in iterative change and causality: only then can actionable and impactful changes be made for a more equitable future of online learning.

**Keywords:** Social justice, Educational access, Disadvantage, Longitudinal research, Online learning

## 1. Access is critical to realising education as a human right

Whilst much research attention has been given to educational gains through effective delivery of education (Reynolds et al., 2014), much less attention has been given to the precursors of gaining access to such education. Yet, educational access is an indispensable pre-requisite for educational engagement and academic gains. Indeed, access to education is not so much a privilege as it is a fundamental human right (United Nations Convention, 1990). Meanwhile, there is widespread recognition that educational access is not universal, with prevalent challenges globally, especially among learners in low-and-middle-income countries and from disasdvantaged backgrounds (Momo et al., 2019).

Systematic, empirical research has consistently shown interest in analysis of inequalities (Berkowitz et al., 2017; Broer et al., 2019). However, growing concern is now emerging beyond inequalities to address educational inequities (Long, 2022; Yerrick et al., 2022). There are fundamental and consequential distinctions between considerations of inequality and inequity. Inequality research is outcome-oriented and concerned, for example, with learner attainment or teacher salary. In contrast, inequity research is concerned with the whole individual and asks questions of, for example, the whole learner, factors of constraints, and rights or privileges that are amiss for significant minorities (Espinoza, 2007).

Once inequities have been highlighted, and as recognition is established of the situation, the potential to disrupt injustice increases, as the narrative is increasingly shared — and owned — by all. The stance of the marginalised normalises as these populations are humanised to, in turn, be better accounted for in future iterations of educational provision (Stornaiuolo & Thomas, 2017). Thus, research into the predictors, mechanisms, and under-represented narratives for obstacles to educational access is critically important if educational research is to hold to education as a human right.

## 2. A critical look at online learning for educational access

Online learning has long promised to increase educational access. With the right ingredients, human development scholars have documented their anticipation for the potential in online learning to meet the needs of underserved populations, to capitalise diversity, and to build society to an unprecedented scale (Stewart, 2004). Indeed, early signs lent support to this trajectory for online learning to reach the unreached (e.g., Moloney & Oakley, 2010).

Yet, online learning does not always increase educational access, or resolve societal divides. In fact, the opposite is often true among the most marginalised, especially as globalisation has proceeded at an alamingly rapid pace (Zondiros, 2008). In part, this is because disparities can be exacerbated by online learning, due to digital divides that are not present in classroom learning (Mathrani et al., 2021). Recent machine learning analysis has found online learning to exacerbate between-country and gender-related disparities in educational access (McIntyre, 2022b). Interestingly, these gender disparities were not found in educational gains made once online learning is accessed (McIntyre, 2022a). Thus, the role of online learning in equalising educational access is a complex one with potential for research to reveal many surprising, complex, and insightful mechanisms.

The COVID-19 pandemic has prompted unprecedented attention to the social divides in educational access via online and blended learning (e.g., Greenhow & Galvin, 2020). The unique affordances and challenges of online learning, along with its psychosocial correlates, have received more attention than ever before. Although much opportunistic research has been conducted since the onset of the pandemic, these research efforts have consisted mostly of snapshot studies from a very specific and ungeneralisable moment in time. Little-to-no research has employed the longitudinal research design, either at data collection or during analysis. There is a dearth of research on students' experiences of online education at the process-level, such as for social interaction (Rasheed et al., 2020).

Accordingly, the Special Issue set out to address the question: does online learning serve to increase educational access for the disadvantaged, or does it instead exacerbate societal divides and feed into the digital divides? In doing so, this issue called for research on access disparities in online learning that adopts a longitudinal design.

## 3. Contribution of papers in this special issue

The following four papers were accepted in this Special Issue. Each one of them shed theoretical and practical light on the potential for online learning to support, or hinder, educational access.

As the first paper in this Special Issue, Li and colleagues (*this issue*) explored the relationship between the internal, individual differences among online learners and their external, environmental disadvantages. Specifically, they investigated the comparative importance of online learners' personalities versus the urban/rural status of their context for learning. They reported extraversion, neuroticism, and agreeableness to make significant contributions to devleopment of digital skills, over and above the rural disadvantage. This brings encouragement, that online learners in disadvantaged, rural settings can potentially overcome socioeconomic and infrastructural constraints to make learning gains.

Nevertheless, the possibility to overcome digital divides to access online learning can be limited.

In their paper, Ng and colleagues (*this issue*) shed light on disadvantaged learners who venture across socioeconomic divides as a daily experience, namely cross-border students. These are students who would normally travel from Mainland China to Hong Kong, China, every day in order to access their school-based education. During the Pandemic, their educational access moved online due to travel restrictions. This online-only educational access set cross-border students' learning back severely, relative to their local peers. Other than the single modality for learning, the lower socioeconomic status shared by these students meant that they were unlikely to possess the advanced digital resources and infrastructure required to sustain online learning. Thus, inequities can overwhelm online learners who cannot reasonably overcome an ecosystem of disadvantage in the way that only few might.

Related is the disadvantaged position of being physically disabled which can prove an overwhelmingly disadvantaged one. AlShawabkeh and colleagues (*this issue*) report research with deaf online learners in Higher Education. Compared with hearing peers, deaf or hard of hearing learners were found to make consistently lower learning gains: this inequity was exacerbated by forced online learning during Covid-19 restrictions. Thus, online learning is more often found to widen rather than reduce societal divides by presenting digital divides into an ecosystem of unequitable educational access.

It is clear that significant efforts are required to tackle the challenges that online learning presents to social justice at multiple levels.

To address this, Lin and colleagues (*this issue*) presented a scaffolding framework by which teachers and learners can be supported to adopt technological tools for learning. The framework recognises the primary importance of social dynamics in the use of any learning resource, including digital ones. The framework then addresses the importance of technological readiness for learning with novel digital tools: in this case, augmented reality on mobile phones. Furthermore, in their analyses, Lin and colleagues examine the difference between the value of mobile augmented reality for learning in rural areas as compared with urban areas. Indeed, the dual scaffolding-embedded framework was found to mitigate the rural disadvantage in online learning. Thus, authors have reported on a rigorous, multifaceted approach to tackling digital divides in online learning.

*Paper 1:* Personality Traits Predict Digital Skills Divide between Urban and Rural College Students: A Longitudinal and Cross-sectional Analysis of Online Learning during COVID-19 Pandemic.
*Authors:* Li Zhao, Yue Liu, and Yu-Sheng Su

*Paper 2:* Inequity Issues in Online learning of Chinese Cross-border Students under the COVID-19 Pandemic: A Longitudinal Study at a Macro-level
*Authors:* Davy Tsz Kit Ng, and Roxanne Xiaoxuan Fang

*Paper 3:* Technology-based learning and digital divide for deaf/non-deaf students during Covid-19: Academic justice lens in higher education.
*Authors:* Abdallah A. AlShawabkeh, Faten F. Kharbat, Ajayeb S. Abu Daabes, and M. Lynn Woolsey

*Paper 4:* Mitigating the urban-rural digital divides: A dual scaffoldings-embedded mobile augmented reality learning approach in the post COVID-19 pandemic.
*Authors:* Xiao-Fan Lin, Juan Jiang, Guoyu Luo, Xiyu Huang, Wenyi Li, Jiayan Zou, Zhaoyang Wang, and Qintai Hu

## 4. Conclusion and future research

This special issue revealed that online learning does not automatically increase educational access or dissolve societal divides. On the contrary, online learning can magnify digital divides to, instead, exacerbate societal inequities. This is especially true when online learning is the main—or sole—channel for accessing education (Ng et al., *this issue*), and even more so when essential assistive technology is not available (AlShawabkeh et al., *this issue*). Although some online learners may have the internal resources required to prevail (Li et al., *this issue*), this is not normally the case for the disadvantaged learner. A holistic framework is required in order to address both social and digital disparities in access to online learning (e.g., Lin et al., *this issue*).

Our evidence derives from rigorous research methods, too. Mixed methods research features in two of the four papers (Ng and AlShawabkeh) to triangulate insights across data types. Triangulation across constructs is carried by the other two papers (Li and Lin). Change over time is at the forefront, such that longitudinal patterns are analysed across all the papers. Thus, together, we present blueprints of how comprehensive research into social justice in online learning should be conducted.

Despite the above insights and achievements, much has still to be done.

This Special Issue pioneers in its efforts to highlight and champion longitudinal research into a critical component of education for social justice. More longitudinal research is needed in order to disentangle the processes and mechanisms for effective online education that is equitable in iterative design and delivery. Such a priority needs to be examined and enacted across the levels of the education system. At the micro level, processes relating to social justice need investigation within sessions of online learning. At the macro level, change over time needs to be examined beyond the two or three time-point approach to longitudinal research.

Furthermore, only two countries were represented in this Special Issue: there is a need for increased interest in this research topic at the global level, with more country settings represented that examine social justice in online learning, and perhaps countries can be compared within single studies to explore complementary strengths across countries in the design and delivery of equitable education through online learning.

Finally, an increased focus on online learning in the form of open educational resources (OERs) may really be where the greatest rewards lie for rebalancing educational inequities in online learning (Adam, 2020; Geith & Vignare, 2008). It is within this channel of online learning provision that there is exceptional scope for learning opportunities that circumvent socioeconomic and infrastructural barriers which can otherwise characterise access to online learning. Therefore, scholars would do well to concentrate research efforts on exploring and developing scalable frameworks for online learning provision as OERs to conprehensively address prevalent barriers for disadvantaged populations around the world.

## Acknowledgment

This Special Issue is highly inter-disciplinary with a research focus that has much room for further research attention and significant global impact. In view of these, the Guest Editor thanks the authors and generous reviewers for their invaluable support towards the completion and success of this Special Issue.

## References

Adam, T. (2020). Between social justice and decolonisation: Exploring South African MOOC designers' conceptualisations and approaches to addressing injustices. *Journal of Interactive Media in Education, 2020*(1), 7. https://doi.org/10.5334/jime.557

Berkowitz, R., Moore, H., Astor, R. A., & Benbenishty, R. (2017). A Research synthesis of the associations between socioeconomic background, inequality, school climate, and academic achievement. *Review of Educational Research, 87*(2), 425–469. https://doi.org/10.3102/0034654316669821

Broer, M., Bai, Y., & Fonseca, F. (2019). A Review of the literature on socioeconomic status and educational achievement. In M. Broer, Y. Bai, & F. Fonseca (Eds.), *Socioeconomic Inequality and Educational Outcomes: Evidence from Twenty Years of TIMSS* (pp. 7–17). Springer International Publishing. https://doi.org/10.1007/978-3-030-11991-1_2

Espinoza, O. (2007). Solving the equity–equality conceptual dilemma: A New model for analysis of the educational process. *Educational Research, 49*(4), 343–363. https://doi.org/10.1080/00131880701717198

Geith, C., & Vignare, K. (2008). Access to education with online learning and open educational resources: Can they close the gap? *Journal of Asynchronous Learning Networks, 12*(1), 105–126.

Greenhow, C., & Galvin, S. (2020). Teaching with social media: Evidence-based strategies for making remote higher education less remote. *Information and Learning Sciences, 121*(7/8), 513–524. https://doi.org/10.1108/ILS-04-2020-0138

Long, D. A. (2022). Equity, not just equality: How equality of educational outcome policies could help narrow excellence and identification gaps. *Gifted Child Quarterly, 66*(2), 105–107. https://doi.org/10.1177/00169862211037944

Mathrani, A., Sarvesh, T., & Umer, R. (2021). Digital divide framework: Online learning in developing countries during the COVID-19 lockdown. *Globalisation, Societies and Education, 0*(0), 1–16. https://doi.org/10.1080/14767724.2021.1981253

McIntyre, N. A. (2022a). Accelerating online learning: Machine learning insights into the importance of cumulative experience, independence, and country setting. *Computers and Education: Artificial Intelligence, 3*, 100106. https://doi.org/10.1016/j.caeai.2022.100106

McIntyre, N. A. (2022b). Access to online learning: Machine learning analysis from a social justice perspective. *Education and Information Technologies, 28*, 3787–3832. https://doi.org/10.1007/s10639-022-11280-5

Moloney, J. F., & Oakley, B. (2010). Scaling online education: Increasing access to higher education. *Journal of Asynchronous Learning Networks, 14*(1), 55–70.

Momo, M. S. M., Cabus, S. J., De Witte, K., & Groot, W. (2019). A Systematic review of the literature on the causes of early school leaving in Africa and Asia. *Review of Education, 7*(3), 496–522. https://doi.org/10.1002/rev3.3134

Rasheed, R. A., Kamsin, A., & Abdullah, N. A. (2020). Challenges in the online component of blended learning: A Systematic review. *Computers & Education, 144*, 103701. https://doi.org/10.1016/j.compedu.2019.103701

Reynolds, D., Sammons, P., De Fraine, B., Van Damme, J., Townsend, T., Teddlie, C., & Stringfield, S. (2014). Educational effectiveness research (EER): A State-of-the-art review. *School Effectiveness and School Improvement, 25*(2), 197–230. https://doi.org/10.1080/09243453.2014.885450

Stewart, B. L. (2004). Online learning: A Strategy for social responsibility in educational access. *The Internet and Higher Education, 7*(4), 299–310.

Stornaiuolo, A., & Thomas, E. E. (2017). Disrupting educational inequalities through youth digital activism. *Review of Research in Education, 41*(1), 337–357. https://doi.org/10.3102/0091732X16687973

United Nations Convention. (1990). *The United Nations Convention on rights of the child*. UNICEF. https://www.unicef.org.uk/what-we-do/un-convention-child-rights/

Yerrick, R., Eastman, M. G., Miles, M. L., Henderson, R., & Nunna, R. (2022). Striving for more: Beyond the guise of objectivity and equality in engineering education. In A. J. Rodriguez & R. L. Suriel (Eds.), *Equity in STEM Education Research: Advocating for Equitable Attention* (pp. 169–197). Springer International Publishing. https://doi.org/10.1007/978-3-031-08150-7_9

Zondiros, D. (2008). Online, distance education and globalisation: Its impact on educational access, inequality and exclusion. *European Journal of Open, Distance and E-Learning, 11*(1).

Lin, X.-F., Jiang, J., Luo, G., Huang, X., Li, W., Zou, J., Wang, Z., & Hu, Q. (2023). Mitigating the Urban-rural Digital Divide: A Dual Scaffolding-embedded Mobile Augmented Reality Learning Approach in the Post COVID-19 Pandemic. *Educational Technology & Society*, *26*(4), 108-122. https://doi.org/10.30191/ETS.202310_26(4).0008

# Mitigating the Urban-rural Digital Divide: A Dual Scaffolding-embedded Mobile Augmented Reality Learning Approach in the Post COVID-19 Pandemic

**Xiao-Fan Lin[1,2,3*], Juan Jiang[1], Guoyu Luo[1], Xiyu Huang[1], Wenyi Li[2], Jiayan Zou[1], Zhaoyang Wang[1] and Qintai Hu[4]**

[1]School of Education Information Technology, South China Normal University, China // [2]GuangDong Engineering Technology Research Center of Smart Learning, Guangdong Provincial Institute of Elementary Education and Information Technology, China // [3]Guangdong Provincial Philosophy and Social Sciences Key Laboratory of Artificial Intelligence and Smart Education, Institute for Artificial Intelligence Education, South China Normal University, China // [4]New Engineering Education Research Center, Guangdong University of Technology, China // linxiaofan@m.scnu.edu.cn // 2013010048@m.scnu.edu.cn // luoguoyusera@163.com // 1036562967@qq.com // liwenyi@pku.edu.cn // 981901151@qq.com // m17754831067@163.com // huqt8@gdut.edu.cn

[*]Corresponding author

**ABSTRACT:** Mitigating the digital divide is essential for the sustainable development of education. To provide distance access which ensures equality in education for both urban and rural students, online learning has been emphasized in the post COVID-19 period. However, some challenges to total online learning have been described, such as isolation and the lack of in-time interaction due to the separation between teachers and students. The mobile augmented reality (AR) learning approach has the potential to combine real-world and online objects to promote interaction in total online learning. However, researchers have found that teachers and students might feel frustrated while handling too much information to deal with complex tasks in AR learning without appropriate scaffolding, which might reinforce the digital divide. Therefore, there is a need to propose a dual scaffolding-embedded AR learning approach to mitigate the urban-rural digital divide rather than a single form of scaffolding for either teachers or students. A quasi-experiment was conducted by recruiting 173 sixth-grade students from four classes of urban and rural schools in southern China. The longitudinal results showed that the proposed approach effectively improved both urban and rural students' learning achievement, higher-order cognition, and self-efficacy. The comparison of the students' learning outcomes indicated that it was helpful to mitigate the digital divide between rural and urban students through the proposed approach. These findings provide insights for school administrators to provide relational and suitable scaffolding-embedded mobile AR learning to support total online education for mitigating the urban and rural digital divide in the post COVID-19 period.

**Keywords:** Digital divide, Mobile, Augmented reality, Scaffolding, Outcomes

## 1. Introduction

Education equity, one of the core values of sustainable development of education, is proposed in Sustainable Development Goal 4 of the 2030 Agenda to "ensure inclusive and equitable quality education" (UNESCO, 2017). One form of educational inequity is the divergent gains from education during the COVID-19 pandemic that exist between urban and rural settings. Researchers have indicated that technology education could help mitigate the urban-rural digital divide by overcoming some of the strong forces that lead to educational inequalities (Zilka, 2021). Scholars have pointed out that all members participating actively and thoughtfully in the online learning environment enabled effective practices during the COVID-19 pandemic (Carrillo & Flores, 2020). Therefore, online education in the post COVID-19 period has the potential to provide the same equality in education to mitigate the digital divide between urban and rural settings.

Scholars have indicated that total online learning brought some challenges due to the unequal access to technical resources, digital literacy, a feeling of nihility, the poor state of autonomous learning, and a lack of in-time interaction with the separation between teachers and students, thus further widening the digital divide between urban and rural settings during the post-COVID-19 period (Laufer et al., 2021; Palau et al., 2021; Romero-Hall, 2021). In other words, total online learning may fail to achieve the expected teaching objectives in the post COVID-19 period if the design of online education is not appropriate. One form of educational inequity is unequal learning gains from online and blended education between urban and rural settings during the COVID-19 pandemic, which can be regarded as the new urban-rural digital divide. The mobile augmented reality (AR) technology-assisted learning approach provides a solution to mitigate the new urban-rural digital divide in total online learning during the COVID-19 pandemic. AR refers to the technology which integrates virtual and real worlds, and which provides an interactive and immersive experience (Seo & So, 2022). Research has found that the mobile AR technology-assisted learning approach could combine real-world and online objects, assisting users in interacting with online objects in real time to reduce loneliness and isolation from total online learning during the COVID-19 pandemic (Lin et al., 2022a). In other words, the mobile AR technology-assisted learning approach offers a significant advantage of overlapping the real world with online information in real time. Therefore, the mobile AR technology-assisted learning approach provides the potential to address the barriers (e.g., the feeling of nihility, the poor state of autonomous learning, and the lack of in-time interaction) for teachers and students during the post COVID-19 period in total online learning due to the separation of teachers and students. In addition, mobile technology-assisted AR is regarded as one technology that can help students develop cognition and motivate them to learn during the learning process (Chou et al., 2021). It has enormous potential for the fair distribution of educational resources to mitigate the urban-rural digital divide. Research has verified that the mobile AR technology-assisted learning approach could reduce the effects of individual differences (Chen & Wang, 2015).

Although the mobile AR technology-assisted learning approach could meet teachers' and students' need to mitigate the urban-rural digital divide, teachers and students may feel frustrated when handling plentiful information to deal with complex tasks in mobile AR learning. Other challenges to mobile AR learning that may cause a new education gap between urban and rural settings have been indicated, such as poor use of AR resources for teaching (Alalwan et al., 2020; Alkhattabi, 2017), and the lack of guidance or prompts for students (Extremera et al., 2020). Previous studies have emphasized the significance of the provision of proper scaffolding to address the challenges to teachers' and students' mobile AR learning (Alkhattabi, 2017; Lin et al., 2020). Multi-dimensional scaffolding, including cognitive scaffolding, metacognitive scaffolding, and peer scaffolding, has the potential to solve the problems of teachers and students when applying the mobile AR technology-assisted learning approach (Hou & Keng, 2021; Lin et al., 2020).

However, as far as we know, no studies regarding both teachers' and students' scaffolding-based AR learning approach to mitigate the urban-rural digital divide in the post COVID-19 period have been presented. Most previous studies have paid attention to using a single form of scaffolding for either teachers or students (Ibanez et al., 2016; Lin et al., 2020; Tsai & Huang, 2014) in AR learning rather than integrating dual scaffolding into AR-based learning to achieve the expected outcomes. To address this research gap, a dual scaffolding-embedded AR learning approach for teachers and students is proposed to narrow the urban-rural digital divide in the post COVID-19 period. Dual scaffolding is defined as scaffolds (e.g., peer scaffolding, problem-assistance scaffolding, and metacognition-awareness scaffolding) that focus on interactions that occur when teachers and students can use in the mobile AR environment. The interactions that are integrated into mobile AR learning can be divided into two types: interaction between teachers and students, and interaction between humans and AR content. The unique attribution of this study was to incorporate dual scaffolding for the interactions that occurred between teachers and students when they engaged in mobile AR learning. Two important aspects of teachers' or students' scaffolding could provide some suggestions for mitigating the urban-rural digital divide in the AR learning process. From the perspective of scaffolding for teachers, Tsai and Huang (2014) indicated that scaffolding is an effective form of assistance for novice and experienced teachers in a mobile AR environment to support teaching material management. From the perspective of scaffolding for students, Hou and Keng (2021) developed a framework of an alternate reality-based board game for AR exploration with multi-dimensional scaffolding, including cognitive scaffolding, metacognitive scaffolding, and peer scaffolding, designed to help students focus on spatial and logical thinking. Therefore, this study developed a dual scaffolding-embedded mobile AR learning approach to mitigate the urban-rural digital divide during the post COVID-19 period. To evaluate the effectiveness of the proposed approach, a quasi-experiment was conducted by using an intelligent AR environment platform. The research questions are as follows:

- Q1: Can the dual scaffolding-embedded mobile AR learning approach benefit students by enhancing their learning achievement to mitigate the urban-rural digital divide?
- Q2: Can the dual scaffolding-embedded mobile AR learning approach enhance students' higher-order cognition to mitigate the urban-rural digital divide?
- Q3: Can the dual scaffolding-embedded mobile AR learning approach promote students' self-efficacy to mitigate the urban-rural digital divide?

## 2. Literature review

### 2.1. Mitigating the urban-rural digital divide

Mitigating the urban-rural digital divide refers to better quality education outcome equity that can ensure students from urban and rural settings, irrespective of their socioeconomic backgrounds, gender, area, or other characteristics, have the digital education they need to achieve certain outcomes (Nachbauer & Kyriakides, 2020). Mitigating the urban-rural digital divide is essential to remove the educational inequity to achieve educational equality. Mitigating the digital divide between urban and rural settings has become a widely-discussed issue involving educational facilities and digital teaching resources (Kuo et al., 2021). It can engage students in urban and rural schools in situating and interacting to achieve academic achievements and develop their cognition. Empirical results regarding education divide diminishment have shown that when teachers exhibit a higher degree of integration between online learning course operation and the digital characteristics of online learning courses, the level of the rural-urban education divide could decrease (Hsieh, 2017). Thus, the effective application of online learning has prompted unprecedented attention to mitigate the digital divide during the COVID-19 pandemic.

However, some scholars have highlighted inequalities related to the digital divide (unequal access to technical resources) and the new digital divide (differing levels of digital skills) (Ritzhaupt et al., 2020). In addition, researchers have discovered that the digital divide was more likely to appear in public schools and rural areas (Palau et al., 2021). For example, teachers had difficulties self-regulating their work for online education (Laufer et al., 2021; Palau et al., 2021). Students may lack digital literacy and in-time interactions, and have poor self-regulation for online learning. Previous researchers tended to choose crucial and convenient factors for testing the effect of mitigating the urban-rural digital divide from the perspective of learning achievement (Charalambous et al., 2018), space accessibility (Wang et al., 2021), and attitude (Lin et al., 2022b). However, there is an urgent need to explore an effective approach to facilitating students' higher-order cognition tendency and self-efficacy to equip students with skills and motivation to overcome the challenges of the urban-rural education gap in the next generation. In other words, the combination of learning achievement, higher-order cognition tendency, and self-efficacy may allow researchers to attain the best of the new while retaining the best of the old when exploring the effect of mitigating the urban-rural digital divide. Students' achievement, recognition ability, self-efficacy, and literacy levels have also been used to measure the urban-rural digital divide (Kyriakides et al., 2021; Penuel & Watkin, 2019). Accordingly, this study utilized students' academic achievement, degree acquisition, ability, and literacy improvement to evaluate the urban-rural digital divide.

### 2.2. Mobile AR learning

Batdi et al. (2018) found that the integration of technology into education to support the teaching and learning process, such as mobile AR technology, as a learning approach, has a positive effect on students' academic achievement. Some previous research has proved the effects of mobile AR technology on promoting students' learning outcomes, such as students' learning achievements (Extremera et al., 2020), positive learning attitudes, and higher-order thinking capacity (Dunleavy et al., 2009), which mitigates the urban-rural digital divide. Mobile AR technology has been shown to facilitate interactive learning in practice, implying that mobile AR technology as information technology has the potential to mitigate the urban-rural digital divide. Lin et al. (2015) developed an AR-assisted learning system to assist junior high school students in learning solid geometry, and found that it could improve students' spatial perceptions. Moreover, AR has a better effect on students with higher self-efficacy (Lin et al., 2020). Despite the potential advantages of AR in mitigating the urban-rural digital divide, most previous studies have focused on the function of interactive situational resources in a real-world environment, implying the main benefits of perceiving and comprehending pre-packaged information (Me & Hsu, 2015). Such a learning experience may confuse users when the real-world target and the digital information mismatch, implying the significance of integrating a teaching support tool into a mobile AR environment (Wu et al., 2018). In this study, AR is considered as a technology that utilizes mobile and context-aware devices (e.g., smartphones, tablets) to enable participants to interact with digital information embedded in a real-world environment in real time (Seo & So, 2022).

Although mobile AR technology-assisted learning could be an effective approach to match the need for weak schools in rural areas to mitigate the urban-rural digital divide, there are some learning and teaching difficulties if the approach is not used properly. In AR-based digital divide between urban and rural settings, two bodies of research can be found on teachers and students to help us identify the learning and teaching difficulties. On the one hand, students may have difficulties with learning effectiveness due to cognitive overload, misconceptions, and the lack of effective reflection prompts. Some students may not complete interactive tasks because they may suffer from cognitive overload when using AR situational interactive resources (Extremera et al., 2020). Cognitive overload may have a negative effect on students' AR learning effectiveness (Akcayir & Akcayir, 2017; Chen, 2020). Chen (2020) has declared that if the refection prompts only give feedback to students at a knowledge level, AR may fail to significantly enhance their learning achievements with refection prompts. In other words, when students receive too much real and online information in the AR learning environment, their learning effectiveness is negatively influenced due to cognitive overload. However, previous researchers have suggested that using multi-dimensional scaffolding (e.g., cognitive, metacognitive, and peer scaffolding) can address students' spatial and logical thinking difficulties in AR-based learning (Hou & Keng, 2021). On the other hand, one important difficulty for teachers in AR-based teaching is the lack of effective supporting scaffolding. Teachers may have trouble addressing the unintended AR technical problems because of three main difficulties: lack of human infrastructure and IT skills, lack of appropriate ICT infrastructure, and resistance to change (Alkhattabi, 2017). To solve the teaching difficulties of lacking effective supporting scaffolding, previous studies have revealed that the combination of scaffolding theory and AR technology could be useful for teachers to access material management with interaction and convenience for using AR effectively (Tsai & Huang, 2014). In other words, scaffolding can be embedded into the AR-based learning environment to solve these problems for teachers and students to support learning and teaching effectiveness for mitigating the urban-rural digital divide.

### 2.3. Scaffolding

Scaffolding is defined as support that enables somebody to solve a problem, carry out a task, or achieve a goal beyond their unassisted efforts (Kim & Hannafin, 2011). Shin et al. (2020) demonstrated that student perceptions of the usefulness of hard, peer, and teacher scaffolds may positively impact students' academic achievement and group performance in inquiry-based technology-enhanced learning activities. Some have discovered that appropriate technology-based scaffolding can potentially encourage more reflective and elaborative discourses, which can assist students' task-related interaction with regard to problem solving (Ak, 2016). In this study, scaffolding can be defined as assistance from a more knowledgeable person who helps learners do a learning task beyond their capability in a technology-enhanced learning environment.

However, to the best of our knowledge, there are no studies integrating dual scaffolding into AR-based learning to mitigate the urban-rural digital divide. Most existing studies focus on using a single form of scaffolding for teachers or students in AR learning to mitigate the urban-rural digital divide, rather than combining AR with dual scaffolding. To address this research gap, a dual scaffolding-embedded mobile AR learning approach is proposed to mitigate the urban-rural digital divide and to promote students' learning achievement, cognition, and self-efficacy. The integration of dual scaffolding for teachers and students and mobile AR learning is the most significant feature of this study. Three categories of studies on scaffolding provide evidence for students or teachers to mitigate the urban-rural digital divide in the mobile AR environment.

On the one hand, researchers have incorporated effective supporting scaffolding into AR to promote teaching and learning effectiveness from teachers' perspectives (Lin et al., 2020; Tsai & Huang, 2014). Lin et al. (2020) indicated a need to apply both reflective scaffolding and technology-assisted (e.g., AR-based) teacher support to help students modify their misunderstandings. Tsai and Huang (2014) emphasized that providing a material management scaffolding tool was effective for novice teachers and experienced teachers in a mobile AR learning environment. However, since the current AR system is usually designed by technology developers, it is difficult to fully meet the needs of teachers when they have to conduct student activities in a mobile AR environment, which indirectly affects students' education outcomes, reduces their learning motivation, and increases their learning confusion. Therefore, the teaching scaffolding for teachers has potential value to mitigate the urban-rural digital divide in a mobile AR-based environment, effectively assisting teachers' teaching.

On the other hand, to address the difficulties of students in terms of their learning effectiveness, scaffolding could be applied to improve the students' learning achievement, cognition, and self-efficacy. There is a great deal of research on scaffolding for students using mobile AR technology. Hou and Keng (2021) developed collaborative scaffolding to guide students' discussion, and cognitive scaffolding in an AR-based educational board game, allowing students to obtain feedback and additional clues for improving their cognitive ability. Scholars have indicated the importance of providing scaffolding to reinforce students' knowledge construction and cognitive process in a mobile AR environment (Ibanez et al., 2016). Sezen-Barrie et al. (2020) indicated that using cognitive scaffolding could promote interactions with students in the classroom across time and space. Therefore, the learning scaffolding for students could help them learn effectively to mitigate the digital divide between urban-rural settings in AR-based learning.

Besides, although a single form of scaffolding could help teachers or students in AR-based learning, there is a gap between teachers and students, which could further reinforce the urban-rural digital divide. Lai et al. (2016) indicated differences between teachers and students in terms of their mobile learning environmental preferences. They found that teachers pay attention to technical problems while students focus on rich and useful learning content. Additionally, the AR learning environment enables students to focus more on memorizing, calculating, and practicing (Cai et al., 2021; Lin et al., 2020). Teachers are engaged in designing more effective activities with an AR-based flipped learning guiding approach and interactions with peers to improve their learning motivation, critical thinking tendency, and group self-efficacy (Chang & Hwang, 2018). Thus, there is a need to engage scaffolding to bridge the gap between teachers and students. This scaffolding should ensure that the teachers' teaching approach could match students' learning requirements to mitigate the digital divide between urban and rural settings (e.g., cognition, problem-solving, and self-efficacy). Scaffolding services for teachers and students in the design of mobile AR systematic situational interactive resources should be able to satisfy the functions of teachers' guidance and students' independent exploration. Accordingly, teachers could help students adjust their learning behaviors and improve their academic achievement through dual-scaffolding strategies. To sum up, the above research evidence verifies that a single form of scaffolding has the potential to match the need of teachers or students to mitigate the urban-rural digital divide.

In light of this, there is a need to conduct a dual scaffolding-embedded mobile AR learning approach to mitigate the urban-rural digital divide. In an attempt to improve this situation, this study integrated learning scaffolding on both the teacher-side and student-side system operation, and incorporated the teaching scaffolding service mechanism in the teacher-side operation to mitigate the urban-rural digital divide. Besides, we developed a dual scaffolding-embedded mobile AR learning system to meet the scaffolding service for teachers' teaching and students' learning in the mobile AR environment during the construction of the approach.

## 3. The dual scaffolding-embedded mobile AR learning approach

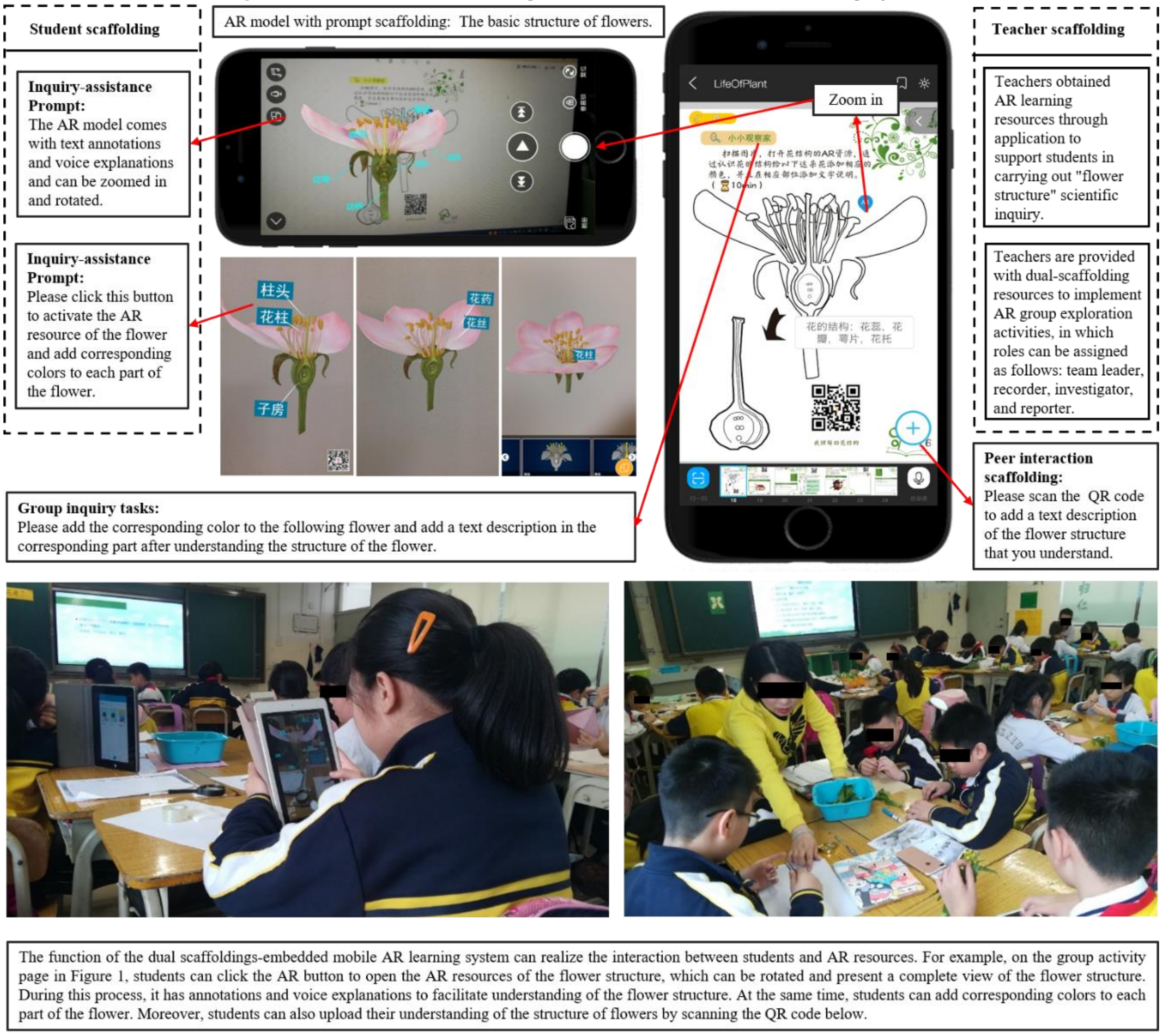
This study proposed a dual scaffolding-embedded mobile AR learning approach, providing the scaffolding service for both teachers and students. Accordingly, a dual scaffolding-embedded mobile AR learning system (Figure 1) was developed based on the proposed approach, including the dual scaffolding mechanism, the mobile AR mechanism, and the cloud database. First, the dual-scaffolding mechanism enabled students to complete the learning tasks with the guidance of peer scaffolding, problem-assistance scaffolding, and metacognition-awareness scaffolding. It could help teachers monitor the progress of the activity and guide each student's learning progress with the assistance of scaffolding of problem situations, guidance, feedback, and reflection.

In addition, the mobile AR mechanism had five functions as follows: (1) embedding many different AR models; (2) greatly shortening the production cycle of AR resources for teachers in urban and rural areas; (3) achieving an AR effect that meets the teaching objectives; (4) testing students' mastery of knowledge; and (5) providing students with a better understanding of the learning knowledge through interacting with the AR model. The teacher used Dream Editor (i.e., AR application production software) to quickly integrate the AR model into completed teaching resources with the mobile AR mechanism; it did not require a professional programming language. Finally, the cloud database enabled various data (e.g., scaffolding service data, AR-related data, etc.) to connect, recording students' learning data and teachers' teaching data with technologies.

The experimental group used a dual scaffolding-embedded mobile AR learning system which can realize the interaction between students and AR resources. For example, on the group activity page in the upper left of Figure 1, students can click the AR button to open the AR resource of the flower structure, which can be rotated to present a complete view of the flower structure. During this process, it has annotations and voice explanations to facilitate understanding of the flower structure. Students can add corresponding colors to each part of the flower. Moreover, students can also upload their understanding of the structure of flowers by scanning the QR code below.

We developed the ISEED 3.0 platform to construct and achieve the approach, matching the requirement of teachers' top-down teaching goals and students' bottom-up learning goals. The ISEED 3.0 platform is the third iteration of the ISEED platform, and can be used for free by teachers and students in Southern China. This platform was purchased by the Provincial Education Department. All the applications are implemented on that platform, and the details about the ISEED platform were documented in our previous publication (Lin et al., 2020). Besides, we put forward the goal-driven approach (e.g., top-down analysis, decomposition, and refinement) and bottom-up approach (e.g., emergence, fusion, and self-organization) as two implementation methods to ensure the efficient dual-scaffolding matching analysis between the teaching and learning scaffolding. The scaffolding service varies from person to person and from time to time; it requires the application of complex network analysis and a looping optimization algorithm of genetic algorithm heuristic rules. Additionally, a multi-dimensional and multi-layered dual-scaffolding analysis the interactive network of which is composed of big data is also needed. Considering the exponential complexity of the call and combination algorithm of teacher-student scaffolding teaching service in a mobile AR environment, the more complexity the problem is, the more difficult finding the optimal solution for a polynomial-time becomes. The processes of population initialization, termination condition judgment, genetic operation, chromosomal evaluation, and new population generation are realized in five iterative cycles. To realize the steady-state processing of the teacher-student scaffolding service, we can obtain the fast combination and call of the optimal dual-scaffolding matching in the current domain of learners.

*Figure 1.* The dual scaffolding-embedded mobile AR learning system




## 4. Method

### 4.1. Participants

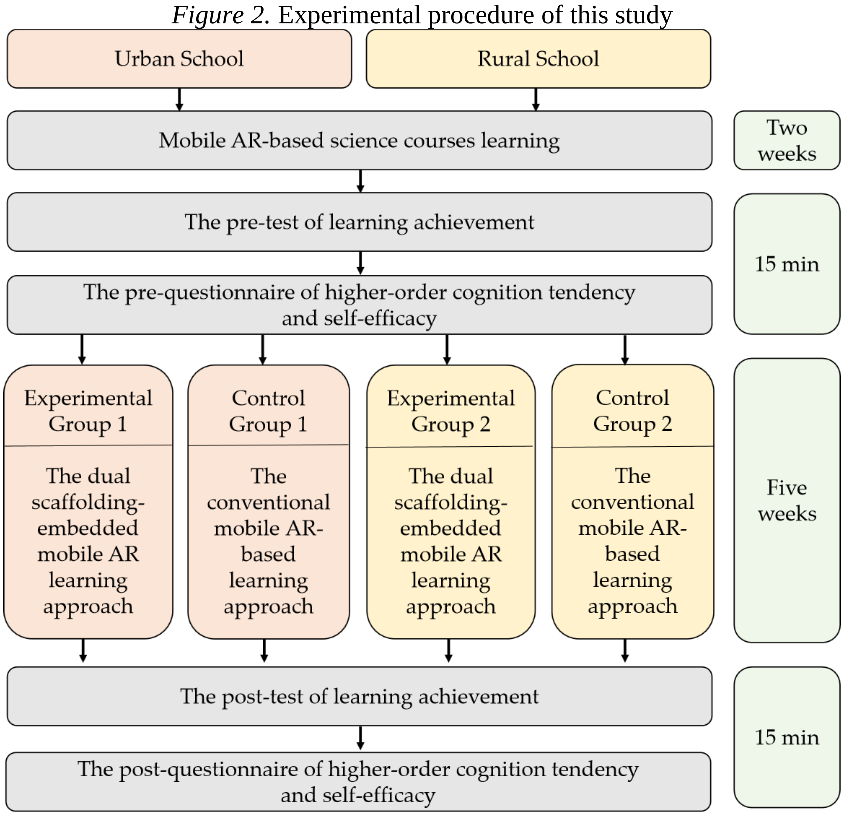
This study recruited 173 sixth-grade students from four classes of urban and rural area schools in southern China. None of the participants had any prior experience of using AR. To ensure that all participants had the same AR learning experience, they were required to complete AR-based science courses. Accordingly, participants were qualified to take a series of more advanced courses, using mobile phones to integrate the dual-scaffolding AR matching mechanism. Therefore, all participants had previous exposure to AR learning projects, in which mobile devices were used for teaching and learning (i.e., the projects occupying more than one-third of the class time).

The students of both schools were divided into two groups. One class was set as the experimental group and learned using the proposed approach, and the other class was set as the control group and learned using the conventional AR-based learning approach. The experimental group consisted of 86 students (mean age = 11.25 years, standard deviation = 1.60). The control group consisted of 87 students (mean age = 11.70 years, standard deviation = 0.98).

### 4.2. Experimental procedure

The experimental procedure of this study is illustrated in Figure 2. Both groups took 2 weeks of AR-based science courses on this subject until they were used to the mobile AR environment. In addition, both groups of students adopted identical mobile AR learning resources to learn the content of flower structure instructed by the same teacher. Furthermore, all the students from each group were required to finish the pre-test of learning achievement and the pre-questionnaire of higher-order cognition tendency and self-efficacy for 15 minutes. Students from both groups were asked to complete the post-test of learning achievement and the post-questionnaires for 15 minutes to examine their learning achievement, higher-order cognition tendency, and self-efficacy with the different learning approaches. The intervention was administered to both groups for 5 weeks. Specifically, both groups received the teacher's feedback to clarify their misconceptions before the next learning activity to achieve the same learning objectives of cognitive concept, inquiry learning, and innovative design with different kinds of assistance scaffolding.

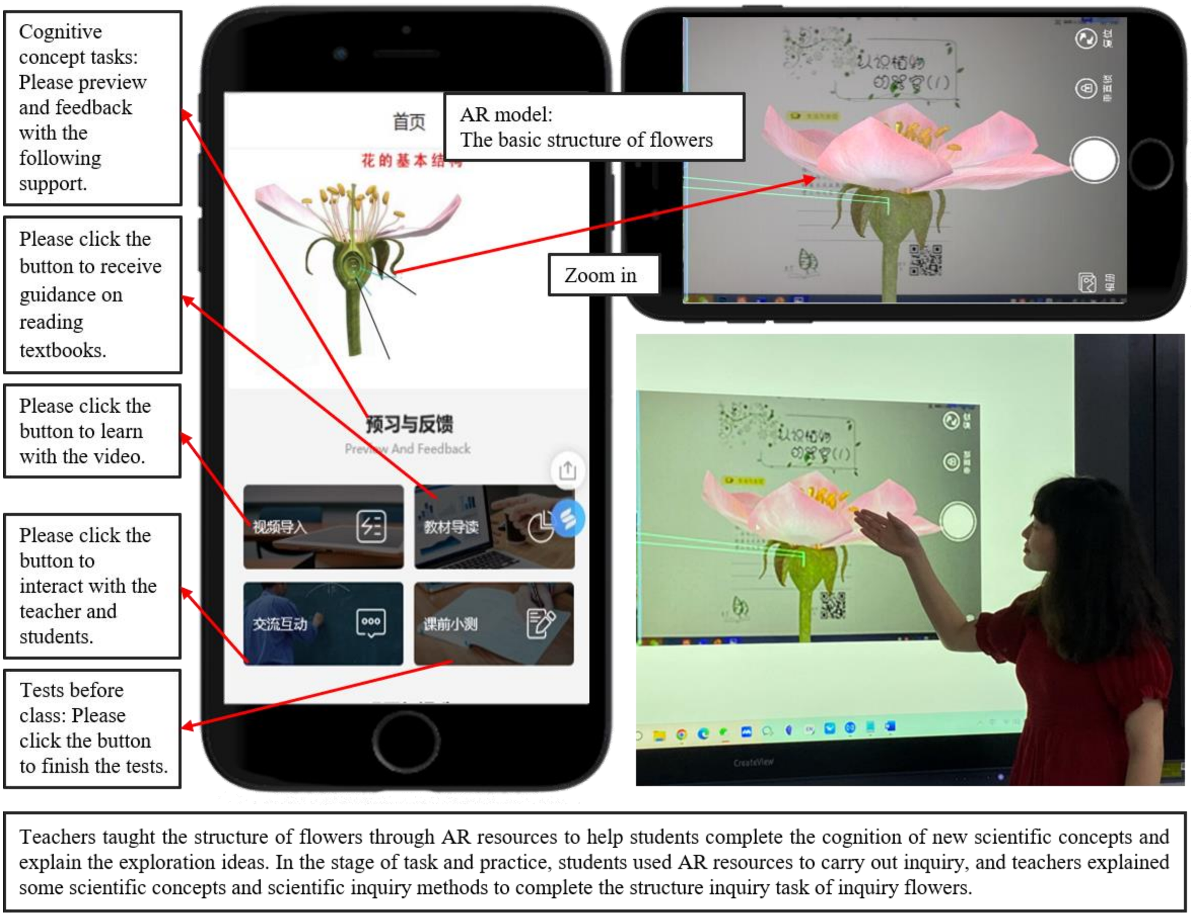
*Figure 2.* Experimental procedure of this study




Each group of students adopted different mobile AR learning approaches with and without dual scaffolding. The experimental group operated the dual scaffolding-embedded AR learning approach. In this experimental group, students participated in peer scaffolding, problem-assistance scaffolding, and metacognition-awareness scaffolding. In addition, the teacher in the experimental group conducted teaching activities with the assistance of problem situations scaffolding, feedback scaffolding, and reflection scaffolding. The control group employed the conventional AR-based learning approach to complete the learning tasks, as demonstrated in Figure 3. In the control group, the students learned new concept knowledge and construction in the mobile AR environment to complete the tasks about the learning content without learning scaffolding. Furthermore, the teacher in the control group implemented teaching activities without the guidance of teaching scaffolding.

The experimental group applied dual-scaffolding to achieve the cognitive concept learning objective. First, the students sought help with problem-assistance scaffolding when they found difficulties using AR resources for concept learning. Second, the teacher addressed students' misunderstandings with online answering and in-class typical problem discussions. Third, the teacher consolidated students' knowledge with the activities, such as explaining words and answering questions. On the contrary, the control group adopted the conventional AR-based learning approach. First, the students were asked to read the AR teaching resources to understand new concepts under the step-by-step teacher guidance. Second, the teacher asked students to answer thinking questions. Third, if students met problems in the class exercises or in absorbing the new knowledge, the teacher could clarify the students' misunderstandings in detail.

*Figure 3.* Conventional mobile AR learning system




Then, in the experimental group, the teacher guided students in each learning group to explore problem-based situation scaffolding. Therefore, students in the experimental group could complete the project observation, group collaborative inquiry tasks, and simulation experiments with metacognition-awareness and peer scaffolding. Unlike the experimental group, the teacher in the control group might conduct thematic observations by guiding students to identify similarities and differences between the AR model and the observation in reality. Then, students in the control group could imitate the teacher's experimental steps to reproduce the specific learning process with AR learning resources for completing the experimental thinking exercises in groups.

Lastly, all students tried to promote the comprehensive application of knowledge. For example, all the students from each group were required to design an outdoor practice for Li Lei (i.e., a person in the task) to help him clarify the similarities and differences among different flowers' structures. Based on this theme, the teacher in the experimental group conducted the innovative design of asking students to design the above outdoor practice with the problem-assistance and metacognition-awareness scaffolding. However, students in the control group needed to use their existing knowledge to design outdoor practice under the teacher's guidance to facilitate their understanding of the abstract concepts.

### 4.3. Data collection

#### *4.3.1. The pre-test and post-test of learning achievement*

This study used knowledge test learning achievements and questionnaires to examine students' learning performance. The pre-test and post-test of learning achievement included questions that three science experts validated; sample questions appearing in each test are respectively shown below:

(a) Pre-test sample question: Li Lei's sunflower seeds have been planted for a week but haven't sprouted yet. How can we solve this problem?
(b) Post-test sample question: Li Lei wanted to pollinate the sunflower artificially, so from which part of the flower would he take out the pollen?

#### *4.3.2. Higher-order cognition tendency questionnaire*

The questionnaire was modeled after the original students' higher-order cognition tendency scale developed by Lai and Hwang (2014), which included three subscales: complex problem-solving tendency (6 items), meta-cognitive awareness (10 items), and creativity tendency (6 items). The questionnaire included 22 items that were rated on a 5-point Likert scale. The Cronbach's alpha values of the subscales were 0.90, 0.87, and 0.91, respectively. This questionnaire was used to investigate students' complex problem-solving tendencies (e.g., I can use the approach learned in AR science class to solve practical scientific problems), meta-cognitive awareness (e.g., In AR science class, I can find out the reasons for the failure of the inquiry experiment), and creativity tendency (e.g., I can design a new experimental idea through the virtual experiment in AR science class to verify my conjectures about some scientific phenomena) in the process of completing tasks in a mobile AR environment.

#### *4.3.3. Self-efficacy questionnaire*

Self-efficacy can be described as an individual sense of self-ability to accomplish the given tasks and achieve the designated performance (Bandura, 1997). Students with high self-efficacy tend to adopt deep learning strategies to gain better academic achievements (Feldon et al., 2018). To further explore whether the introduction of AR situational interactive learning resources used in each group could affect students' learning self-efficacies, the AR learning self-efficacy scale was modeled and modified based on the original learning self-efficacy scale (Lin et al., 2019). An example item is "With the virtual experiment in AR science class, I have understood the scientific concepts taught in the course." The scale has five items, and all items are rated on a 5-point Likert scale. The Cronbach's alpha value was 0.92 in this study. The value indicated that the AR learning self-efficacy scale is reliable to use.

### 4.4. Analysis

The statistical analysis was performed to test the effect of the proposed approach, which was assisted by the integration of educational technology and the dual scaffolding, on urban and rural students' learning achievement, higher-order cognition, and self-efficacy by using SPSS 21 (IBM). The four groups of students were situated to learn in the different areas (i.e., urban or rural) and with the different AR-based learning approaches (i.e., the dual scaffolding-embedded mobile AR learning approach or the conventional mobile AR-based learning approach). To explore the effect of the proposed approach on the four groups, a two-way analysis of covariance (ANCOVA) was conducted on the post-test by considering the pre-test scores as the covariance. Before the ANCOVA test, the homogeneity of variance assumption and regression coefficients were tested to examine whether variances across samples were equal. The partial eta-squared ($\eta 2$) and Cohen's d were calculated to determine the effect size (small < 0.2, medium 0.2 ~ 0.5, and large > 0.8) (Cohen, 1988).

## 5. Results

### 5.1. Learning achievement

Before ANCOVA, Levene's test for learning achievement was not significant ($p > .05$). Therefore, the homogeneity of variance assumption could be verified. In addition, the assumption of homogeneity of regression was not violated ($F = 6.35$, $p > .05$). As shown in Table 1, the result showed that the effect on the interaction between learning approaches and areas was not significant ($p > .05$). Therefore, it was sensible to evaluate the main effects of the dependent variables. The two-way ANCOVA result indicated that a significant effect and a moderate effect size ($\eta 2 > 0.059$) were proved for the learning approaches ($F = 12.23$, $p < .01$, $\eta 2 = 0.13$) and for the area ($F = 5.56$, $p < .05$, $\eta 2 = 0.063$) on students' learning achievement.

*Table 1.* ANCOVA results of students' learning achievement

| Source | *SS* | *df* | *MS* | *F* | $\eta 2$ |
|---|---|---|---|---|---|
| Learning approaches | 820.86 | 1 | 820.86 | $12.23^{**}$ | 0.13 |
| Area | 0.61 | 1 | 0.61 | $5.56^{*}$ | 0.063 |
| Learning approaches × Area | 372.17 | 1 | 372.17 | 0 | 0 |
| Error | 5489.02 | 82 | 66.94 | | |

*Note.* $^{*}p < .05$; $^{**}p < .01$.

*Table 2.* Students' learning achievement in the four groups

| Area | Learning approaches | Adjusted mean | Std. error. | *N* |
|---|---|---|---|---|
| Rural | The dual scaffoldings-embedded mobile AR learning | 87.01 | 1.49 | 43 |
| | The conventional mobile AR-based learning | 79.81 | 1.62 | 43 |
| Urban | The dual scaffoldings-embedded mobile AR learning | 81.78 | 1.88 | 43 |
| | The conventional mobile AR-based learning | 73.41 | 2.14 | 44 |

As shown in Table 2, it was found that the students in the rural area who utilized the dual scaffolding-embedded mobile AR learning approach (the experimental group; adjusted mean = 87.01) outperformed those who used the conventional mobile AR-based learning approach (the control group; adjusted mean = 79.81). Moreover, for the urban area, the experimental group (adjusted mean = 81.78) also performed better than the control group (adjusted mean = 73.41). Consequently, these results indicated that the students with the dual scaffolding-embedded mobile AR learning approach achieved significantly better learning achievement than those with the conventional mobile AR-based learning approach in rural and urban areas. Furthermore, the rural students who received the intervention of the dual scaffolding-embedded mobile AR learning approach in the experimental group (the experimental group; adjusted mean = 87.01) improved their learning achievement in comparison with those who received the same intervention in the urban area (adjusted mean = 81.78). In this context, it was thought that, compared with the urban area, a dual scaffolding-embedded mobile AR learning approach might influence students' learning achievement more positively in the rural area.

### 5.2. Higher-order cognition tendency

We performed two-way ANCOVA to analyze the students' perceived higher-order cognition by adopting learning approaches (i.e., the dual scaffolding-embedded mobile AR learning or the conventional mobile AR-based learning) and areas (rural/urban) as independent variables, while the post questionnaire rating of students' perceived higher-order cognition was the dependent variable, and the pre-test degree was the covariate. Levene's test was not significant ($p > .05$), implying that the items for all groups did discriminate usefully. After verifying that the assumption of homogeneity of regression was not violated ($F = 5.34$, $p > .05$), as shown in Table 3, the effect on the relationship between learning approaches and areas was not significant ($p > .05$). However, it was clear that there are significant differences in the dependent variables regarding learning approaches ($F =10.13$, $p < .01$) and areas ($F = 7.07$, $p < .05$). The ANCOVA result for the experimental group ($\eta 2 = 0.101 > 0.059$) and control group ($\eta 2 = 0.078 > 0.059$) represented a moderate effect size ($\eta 2 = 0.082 > 0.059$). In comparison with the students in the urban area (adjusted mean = 3.72, *SD* = 0.84), it should be noted that those in the rural area (adjusted mean = 4.38, *SD* = 0.89) gave higher ratings for higher-order cognition. Obviously, the students in the dual scaffolding-embedded mobile AR learning approach (adjusted mean = 4.34, *SD* = 0.36) improved their higher-order cognition more significantly than the conventional mobile AR-based learning approach (adjusted mean = 3.78, *SD* = 0.67). These results indicated that the rural area and the dual scaffolding-embedded mobile AR learning approach brought better effects on students' perceived higher-order cognition than the urban area and the conventional mobile AR-based learning approach.

*Table 3.* ANCOVA results of students' perceived higher-order cognition

| Source | *SS* | *df* | *MS* | *F* | $\eta 2$ |
|---|---|---|---|---|---|
| Learning approaches | 5.19 | 1 | 5.19 | $10.13^{**}$ | 0.101 |
| Area | 3.96 | 1 | 3.96 | $7.07^{*}$ | 0.078 |
| Learning approaches × Area | 0.02 | 1 | 0.02 | 0.03 | 0 |
| Error | 47.9 | 83 | 0.58 | | |

*Note.* $^{*}p < .05$; $^{**}p < .01$.

### 5.3. Self-efficacy

A two-way ANCOVA was used to examine the pre-test data (covariate) to the post-test data (dependent variable). The independent variables in this analysis were learning approaches (i.e., the dual scaffolding-embedded mobile AR learning approach and the conventional mobile AR-based learning approach) and areas (i.e., the rural area and urban area). Prior to the ANCOVA test, Levene's test for self-efficacy was not significant ($p > .05$), which indicated the homogeneity of variance assumption was not violated. The homogeneity was applied to ensure that variances across samples were equal ($F = 2.69$, $p > .05$). As presented in Table 4, the analysis of ANCOVA suggested that the main effect of the interaction among learning approaches and area was not significant ($p > .05$), but the results noted a significant difference among learning approaches ($F = 9.26$, $p$

< .01, $\eta2$ = 0.095) and areas ($F$ = 4.39, $p$ < .05, $\eta2$ = 0.053). It also revealed that learning approaches have a moderate effect size ($\eta2$ > 0.059). The results indicated that students' self-efficacy in the rural area (adjusted mean = 3.79, $SD$ = 0.71) was significantly higher than that in the urban area (adjusted mean = 3.26, $SD$ = 0.78). This result revealed that learning approaches in the rural area led to a significantly better improvement in students' self-efficacy than in the urban area. The dual scaffolding-embedded mobile AR learning approach (adjusted mean = 4.12, $SD$ = 0.72) outperformed the conventional mobile AR-based learning approach (adjusted mean = 3.51, $SD$ = 0.68). It was found that students who received the dual scaffolding-embedded mobile AR learning approach significantly improved their self-efficacy.

*Table 4.* ANCOVA results of students' self-efficacy

| Source | *SS* | *df* | *MS* | *F* | $\eta2$ |
|---|---|---|---|---|---|
| Learning approaches | 4.56 | 1 | 4.56 | $9.26^{**}$ | 0.095 |
| Area | 3.46 | 1 | 3.46 | $4.39^{*}$ | 0.053 |
| Learning approaches × Area | 1.47 | 1 | 1.47 | 2.92 | 0.034 |
| Error | 41.7 | 83 | 0.5 | | |

*Note.* $^{*}p < .05$; $^{**}p < .01$.

## 6. Discussion

The current study explored the effect of the dual scaffolding-embedded mobile AR learning approach on students' learning achievement, cognition, and self-efficacy in urban and rural areas. By comparing the achievements of students in urban and rural areas, this study showed that the dual scaffolding-embedded mobile AR learning approach could enhance the achievement of students from both urban and rural areas. This was in line with Hou and Keng (2021), who posited that a multi-dimensional scaffolding-embedded AR learning approach could positively influence students' interaction and learning achievement. The present study further examined the effect of the dual scaffolding-embedded mobile AR learning approach on the higher-order cognition tendency of students from urban and rural areas with three dimensions of complex problem-solving tendency, meta-cognitive awareness, and creativity tendency. The results of higher-order cognition tendency showed the dual scaffolding-embedded mobile AR learning brought better effects than the conventional mobile AR-based learning to mitigate the digital divide between urban and rural settings. This was in good agreement with Ibanez et al. (2016), who found the importance of providing scaffolding for improving students' knowledge construction and cognitive process. Moreover, it was interesting to find that students from rural areas with the dual scaffolding-embedded mobile AR learning approach group had significantly better outcomes than those in the urban area. In other words, it was helpful to mitigate the gap in education outcomes between rural and urban areas with the dual scaffolding-embedded mobile AR learning approach. These findings implied that the dual scaffolding-embedded mobile AR learning approach showed its effectiveness in terms of rural students' learning achievement, higher-order cognition tendency, and self-efficacy, which indicated that it is helpful to mitigate the rural/urban gap in education outcomes with the dual scaffolding-embedded mobile AR learning approach.

For better education outcomes, the interactive situational resources offered by the mobile AR environment should be targeted and pushed to students considering their personalized situations. The study proposed dual scaffolding embedded in the mobile AR environment, which provided matching and guiding schemes to teachers and students with multiple elements instead of only focusing on a single design problem as in existing studies (Hou & Keng, 2021; Tsai & Huang, 2014). For example, although Hou and Keng (2021) investigated predictive factors in dual scaffolding (i.e., peer scaffolding and cognitive scaffolding) for an AR educational board game from students' single perspective, they did not prove the effectiveness of dual scaffolding in an AR learning environment. In contrast to Hou and Keng (2021), this study added value to technology-enhanced scaffolding theory by examining the effects of integrating teacher-student dual scaffolding into AR learning to mitigate the digital divide by promoting both urban and rural students' learning achievement, higher-order cognition, and self-efficacy. Therefore, the study may expand the research on teacher-student scaffolding instruction theory, which advanced understanding of how technology promotes the development of educational science and responded to the lack of effective scaffolding in the current AR education. For example, although Kim and Hannafin (2011) concluded the implication of scaffolding for problem-solving inquiry in technology-enhanced classrooms, given the many dimensions of scaffolding: peer scaffolding, teacher scaffolding, and technology-enhanced scaffolding dimensions, they neglected the value of teacher and student scaffolding. In this regard, the current research is one of the pioneering studies to mitigate the urban-rural digital divide by adopting the dual scaffolding-embedded mobile AR learning approach.

High-quality AR resources should match the application level of teaching scaffolding to promote the equity of education outcomes. However, Li et al. (2020) revealed that it was difficult to provide education resources (e.g., qualified rural teachers and training approaches) based on the rural schools' and teachers' needs. Therefore, we conducted an efficient dual scaffolding-embedded AR learning approach for teachers and students in a mobile AR environment. Unlike previous studies (Hou & Keng, 2021; Lin et al., 2020; Shin et al., 2020) which only considered a single form of scaffolding for teachers or students in the learning process, this study has important practical implications for revealing the role of combining AR with dual scaffolding to mitigate the urban-rural digital divide. For example, although Shin et al. (2020) revealed the role of hard, peer, and teacher scaffolding in technology-enhanced learning, they focused on using teacher scaffolding to enhance students' academic achievement and group performance rather than promoting teachers' teaching effect. This study showed that students in the rural area benefited more greatly than those in the urban area from the dual scaffolding-embedded mobile AR learning approach in terms of their learning achievement, higher-order cognition, and self-efficacy. Accordingly, this study could contribute to providing educational policymakers and school administrators with concrete evidence of how to mitigate the urban-rural digital divide with higher quality and lower cost by incorporating the mobile AR learning approach with the support of teacher-student dual scaffolding in the urban area. In contrast to Lin's et al. (2020) finding that technology-assisted teacher support and reflective scaffolding were more helpful for students to experience, reflect, and inquire in AR-based scientific inquiries, this study indicated much higher values for providing both the learning scaffolding on the student-side system operation and the teaching scaffolding service mechanism on the teacher-side operation to help mitigate the urban-rural digital divide with the mobile AR learning environment. The results of the dual-scaffolding matching mechanism between teachers and students were also generated and tested successfully, so the effectiveness of the research conclusions and outcomes were guaranteed.

## 7. Conclusion

This study aimed at mitigating the urban-rural digital divide by adopting the dual scaffolding-embedded mobile AR learning approach. Specifically, the result showed inconsistencies in the effects of different areas with the dual scaffolding-embedded mobile AR learning approach on the students' learning achievement, cognition, and self-efficacy. This implied that the academic levels, cognitive tendency levels, and self-efficacy levels should be considered when designing or using guiding approaches for helping teachers design teaching activities and students' learning with AR situational interactive resources. In particular, it is important to avoid low-achieving or lower-efficacy students feeling frustrated when they encounter difficulties in the process of completing the interactive tasks with AR situational interactive resources. Therefore, it is suggested that future studies should analyze students' learning behaviors with different cognitive tendency levels and self-efficacy levels to understand their cognitive tendency and self-efficacy in an effective scaffolding-embedded mobile AR environment. These practices might mitigate the urban-rural digital divide with effective scaffolding, including teachers' teaching scaffolding and students' learning support in mobile AR learning in the post COVID-19 pandemic. It could provide insights for school administrators to assist teachers' total online teaching and students' self-regulation.

## 8. Limitations

There were three limitations of the present study that should be noted. First, all measures were self-reported, which did not always report students' truth tendency which might cause conceptions with biases. In future research, it is necessary to use a combination of data collection methods, such as teachers' and students' behaviors collected by the system, and self-reported questionnaires to mitigate the biases. Second, due to the limited nature of the statistical tests, this study only examined the validity of the dual scaffolding-embedded mobile AR learning approach to mitigate the urban-rural digital divide in the post COVID-19 period. However, more nuanced crucial ways in the proposed approach for mitigating the urban-rural digital divide may need further exploration. Given the lack of resources in rural areas, future research could track the proposed approach and investigate the crucial ways influencing the effectiveness of this approach in mitigating the urban-rural digital divide. Third, implementation was for 7 weeks and only examined the effect of the dual scaffolding-embedded mobile AR learning approach on K-12 school students' learning achievement, higher-order cognition tendency, and self-efficacy in science. Future research may conduct relevant experiments on more subjects for a longer period to examine its effectiveness.

## Acknowledgement

This research was funded by The National Natural Science Foundation of China, grant number 62007010. The Science and Technology Projects in Guangzhou, grant number 202102021217; The Major Project of National Social Science Fund of China, grant number 18ZDA334; The Special Funds of Climbing Program regarding the Cultivation of Guangdong College Students' Scientific and Technological Innovation, grant number pdjh2022 b0145; 2021 Annual Education Planning Project of Guangdong Province, grant number 2021JKZQ022; Education Science Planning Project of Guangdong Province (Special Project for Higher Education) "Training of Teachers for Artificial Intelligence Education in K-12 stage in the Era of Artificial Intelligence," grant number 2021GXJK625; The Major Project of National Natural Science Foundation of China, grant number 62237001; South China Normal University teaching quality project" Artificial intelligence teacher training in the context of the new pre-service teacher education," grant number 192.

## References

Ak, S. (2016). The Role of technology-based scaffolding in problem-based online asynchronous discussion. *British Journal of Educational Technology, 47*(4), 680–693. https://doi.org/10.1111/bjet.12254

Akcayir, M., & Akcayir, G. (2017). Advantages and challenges associated with augmented reality crossMark for education: A Systematic review of the literature. *Educational Research Review, 20,* 1–11. https://doi.org/10.1016/j.edurev.2016.11.002

Alalwan, N., Cheng, L., Al-Samarraie, H., Yousef, R., Alzahrani, A. I., & Sarsam, S. M. (2020). Challenges and prospects of virtual reality and augmented reality utilization among primary school teachers: A Developing country perspective. *Studies in Educational Evaluation*, *66,* 100876. https://doi.org/10.1016/j.stueduc.2020.100876

Alkhattabi, M. (2017). Augmented reality as e-learning tool in primary schools' education: Barriers to teachers' adoption. *International Journal of Emerging Technologies in Learning*, *12*(2), 91–100. https://doi.org/10.3991/ijet.v12i02.6158

Bandura, A. (1997). *Self-efficacy: The Exercise of control.* W H Freeman.

Batdi, V., Aslan, A., & Chang, Z. (2018). The Effect of technology supported teaching on students' academic achievement: A Combined meta-analytic and thematic study. *International Journal of Learning Technology, 13*(1), 44–60. http://doi.org/10.1504/IJLT.2018.091632

Cai, S., Liu, C., Wang, T., Liu, E., & Liang, J.-C. (2021). Effects of learning physics using Augmented Reality on students' self-efficacy and conceptions of learning. *British Journal of Educational Technology*, *52*(1), 235–251. https://doi.org/10.1111/bjet.13020

Carrillo, C., & Flores, M.A. (2020). COVID-19 and teacher education: A Literature review of online teaching and learning practices. *European Journal of Teacher Education*, *43*(4), 466-487. https://doi.org/10.1080/02619768.2020.1821184

Chang, S.-C., & Hwang, G.-J. (2018). Impacts of an augmented reality-based flipped learning guiding approach on students' scientific project performance and perceptions. *Computers & Education, 125*, 226–239. https://doi.org/10.1016/j.compedu.2018.06.007

Charalambous, E., Kyriakides, L., & Creemers, B.P.M. (2018). Promoting quality and equity in socially disadvantaged schools: A Group-randomization study. *Studies in Educational Evaluation, 57,* 42–52. https://doi.org/10.1016/j.stueduc.2016.06.001

Chen, C.-H. (2020). Impacts of augmented reality and a digital game on students' science learning with reflection prompts in multimedia learning. *Educational Technology Research and Development, 68*(6), 3057–3076. https://doi.org/10.1007/s11423-020-09834-w

Chen, C.-P., & Wang, C.-H. (2015). Employing augmented-reality-embedded instruction to disperse the imparities of individual differences in earth science learning. *Journal of Science Education and Technology*, *24*(6), 835–847. https://doi.org/10.1007/s10956-015-9567-3

Chou, Y.-Y., Wu, P.-F., Huang, C.-Y., Chang, S.-H., Huang, H.-S., Lin, W.-M., & Lin, M.-L. (2021). Effect of digital learning using augmented reality with multi-dimensional concept map in elementary science course. *Asia Pacific Education Researcher*, 1-11. https://doi.org/10.1007/s40299-021-00580-y

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Lawrence Erlbaum Associates.

Dunleavy, M., Dede, C., & Mitchell, R. (2009). Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. *Journal of Science Education & Technology*, *18*(1), 7–22. https://doi.org/10.1007/s10956-008-9119-1

Extremera, J., Vergara, D., Dávila, L. P., & Rubio, M. P. (2020). Virtual and augmented reality environments to learn the fundamentals of crystallography. *Crystals*, *10*(6), 456. https://doi.org/10.3390/cryst10060456

Feldon, D. F., Franco, J., Chao, J., Peugh, J., & Maahs-Fladung, C. (2018). Self-efficacy change associated with a cognitive load-based intervention in an undergraduate biology course. *Learning & Instruction*, *56*, 64–72. https://doi.org/10.1016/j.learninstruc.2018.04.007

Hou, H.-T., & Keng, S.-H. (2021). A Dual-scaffolding framework integrating peer-scaffolding and cognitive-scaffolding for an augmented reality-based educational board game: An Analysis of learners' collective flow state and collaborative learning behavioral patterns. *Journal of Educational Computing Research*, *59*(3), 547–573. https://doi.org/10.1177/0735633120969409

Hsieh, M. Y. (2017). An Empirical study of education divide diminishment through online learning courses. *EURASIA Journal of Mathematics, Science and Technology Education*, *13*(7), 3189-3208. https://doi.org/10.12973/eurasia.2017.00712a

Ibanez, M.-B., Di-Serio, A., Villaran-Molina, D., & Delgado-Kloos, C. (2016). Support for augmented reality simulation systems: The Effects of scaffolding on learning outcomes and behavior patterns. *IEEE Transactions on Learning Technologies*, *9*(1), 46–56. http://doi.org/10.1109/TLT.2015.2445761

Kim, M. C., & Hannafin, M. J. (2011). Scaffolding problem solving in technology-enhanced learning environments (TELEs): Bridging research and theory with practice. *Computers & Education*, *56*(2), 403-417. https://doi.org/10.1016/j.compedu.2010.08.024

Kuo, T. M., Tsai, C.-C., & Wang, J.-C. (2021). Linking web-based learning self-efficacy and learning engagement in MOOCs: The Role of online academic hardiness. *Internet & Higher Education*, *51*, 100819. https://doi.org/10.1016/j.iheduc.2021.100819

Kyriakides, L., Antoniou, P., & Dimosthenous, A. (2021). Does the duration of school interventions matter? The Effectiveness and sustainability of using the dynamic approach to promote quality and equity. *School Effectiveness and School Improvement*, *32*(4), 607-630. https://doi.org/10.1080/09243453.2021.1923534

Lai, C.-L., & Hwang, G.-J. (2014). Effects of mobile learning time on students' conception of collaboration, communication, complex problem–solving, meta–cognitive awareness and creativity. *International Journal of Mobile Learning and Organisation*, *8(*3-4), 276–291. http://doi.org/10.1504/IJMLO.2014.067029

Lai, C.-L., Hwang, G.-J., Liang, J.-C., & Tsai, C.-C. (2016). Differences between mobile learning environmental preferences of high school teachers and students in Taiwan: A Structural equation model analysis. *Educational Technology Research and Development*, *64*(3), 533–554. https://doi.org/10.1007/s11423-016-9432-y

Laufer, M., Leiser, A., Deacon, B., de Brichambaut, P. P., Fecher, B., Kobsda, C., & Hesse, F. (2021). Digital higher education: A Divider or bridge builder? Leadership perspectives on edtech in a COVID-19 reality. *International Journal of Educational Technology in Higher Education*, *18*(1), 1-17. https://doi.org/10.1186/s41239-021-00287-6

Li, J., Shi, Z., & Xue, E. (2020). The Problems, needs and strategies of rural teacher development at deep poverty areas in China: Rural schooling stakeholder perspectives. *International Journal of Educational Research*, *99*, 101496. https://doi.org/10.1016/j.ijer.2019.101496

Lin, H.-C. K., Chen, M.-C., & Chang, C.-K. (2015). Assessing the effectiveness of learning solid geometry by using an augmented reality-assisted learning system. *Interactive Learning Environments*, *23*(6), 799–810. https://doi.org/10.1080/10494820.2013.817435

Lin, X.-F., Hwang, G.-J., Wang, J., Zhou, Y., Li, W., Liu, J., & Liang, Z.-M. (2022a). Effects of a contextualised reflective mechanism-based augmented reality learning model on students' scientific inquiry learning performances, behavioural patterns, and higher order thinking. *Interactive Learning Environments*, 1-21. https://doi.org/10.1080/10494820.2022.2057546

Lin, X. -F., Liang, Z. M., Chan, K. K., Li, W., & Ling, X. (2022b). Effects of contextual interactive healthcare training on caregivers of patients with suspected COVID‐19 infection: Anxiety, learning achievements, perceived support and self‐efficacies during quarantine. *Journal of Computer Assisted Learning*, *38*(3), 731-742. https://doi.org/10.1111/jcal.12644

Lin, X.-F., Tang, D., Lin, X., Liang, Z. M., & Tsai, C. C. (2019). An Exploration of primary school students' perceived learning practices and associated self-efficacies regarding mobile-assisted seamless science learning. *International Journal of Science Education*, *41*(18), 2675-2695. https://doi.org/10.1080/09500693.2019.1693081

Lin, X.-F., Tang, D., Shen, W., Liang, Z.M., Tang, Y., & Tsai, C.-C. (2020). Exploring the relationship between perceived technology-assisted teacher support and technology-embedded scientific inquiry: the mediation effect of hardiness. *International Journal of Science Education*, *42*(8), 1225-1252. https://doi.org/10.1080/09500693.2020.1755475

Me, F., & Hsu, Y.-C. (2015). Mobile augmented-reality artifact creation as a component of mobile computer-supported collaborative learning. *The Internet & Higher Education*, *26*, 33–41. https://doi.org/10.1016/j.iheduc.2015.04.003

Nachbauer, M., & Kyriakides, L. (2020). A Review and evaluation of approaches to measure equity in educational outcomes. *School Effectiveness and School Improvement*, *31*(2), 306–331. https://doi.org/10.1080/09243453.2019.1672757

Palau, R., Fuentes, M., Mogas, J., & Cebrian, G. (2021). Analysis of the implementation of teaching and learning processes at Catalan schools during the Covid-19 lockdown. *Technology Pedagogy and Education*, *30*(1), 183–199. https://doi.org/10.1080/1475939X.2020.1863855

Penuel, W. R., & Watkins, D. A. (2019). Assessment to promote equity and epistemic justice: A Use-case of a research-practice partnership in science education. *The Annals of the American Academy of Political and Social Science*, *683*(1), 201–216. https://doi.org/10.1177/0002716219843249

Ritzhaupt, A. D., Cheng, L., Luo, W., & Hohlfeld, T. N. (2020). The Digital divide in formal educational settings: The Past, present, and future relevance. In M. J. Bishop, E. Boling, J. Elen, & V. Svihla (Eds.), *Handbook of Research in Educational Communications and Technology* (pp. 483–504). Springer. https://doi.org/10.1007/978-3-030-36119-8_23

Romero-Hall, E. (2021). Current initiatives, barriers, and opportunities for networked learning in Latin America. *Educational Technology Research and Development*, *69*(4), 2267–2283. https://doi.org/10.1007/s11423-021-09965-8

Sezen-Barrie, A., Stapleton, M. K., & Marbach-Ad, G. (2020). Science teachers' sensemaking of the use of epistemic tools to scaffold students' knowledge (re)construction in classrooms. *Journal of Research in Science Teaching*, *57*(7), 1058–1092. http://doi.org/10.1002/tea.21621

Seo, M.-H., & So, H.-J. (2022). Developing a gesture-based AR exhibit: Differently-guided experiences for complex conceptual learning in science. *Educational Technology & Society*, *25*(4), 15-28.

Shin, S., Brush, T. A., & Glazewski, K. D. (2020). Examining the hard, peer, and teacher scaffolding framework in inquiry-based technology-enhanced learning environments: impact on academic achievement and group performance. *Educational Technology Research and Development*, *68*(5), 2423–2447. https://doi.org/10.1007/s11423-020-09763-8

Tsai, C.-H., & Huang, J.-Y. A (2014). A Mobile augmented reality based scaffolding platform for outdoor fieldtrip learning. In *Proceedings of the 3rd IIAI International Conference on Advanced Applied Informatics (IIAI-AAI)* (pp. 307–312). http://doi.org/10.1109/IIAI-AAI.2014.70

UNESCO. (2017). Education for sustainable development goals: Learning objectives. In *The Global Education 2030 Agenda*. UNESCO.

Wang, Y. H., Liu, Y. L., Xing, L. J., & Zhang, Z. Y. (2021). An Improved accessibility-based model to evaluate educational equity: A Case study in the city of Wuhan. *IISPRS International Journal of Geo-Information*, *10*(7), 458. https://doi.org/10.3390/ijgi10070458

Wu, P.-H., Hwang, G.-J., Yang, M.-L., & Chen, C.-H. (2018). Impacts of integrating the repertory grid into an augmented reality-based learning design on students' learning achievements, cognitive load and degree of satisfaction. *Interactive Learning Environments*, *26*(2), 221–234. https://doi.org/10.1080/10494820.2017.1294608

Zilka, G. C. (2021). Attitudes of preservice kindergarten teachers toward the integration of computers and the reduction of the digital divide in kindergartens. *Educational Technology Research and Development*, *69*(2), 711-731. https://doi.org/10.1007/s11423-021-09982-7

Ng, D. T. K., & Fang, X. (2023). Inequity Issues in Online learning of Chinese Cross-border Students under the COVID-19 Pandemic: A Longitudinal Study at a Macro-level. *Educational Technology & Society, 26*(4), 123-135. https://doi.org/10.30191/ETS.202310_26(4).0009

# Inequity Issues in Online learning of Chinese Cross-border Students under the COVID-19 Pandemic: A Longitudinal Study at a Macro-level

**Davy Tsz Kit Ng[1] and Xiaoxuan Fang[2*]**
[1]Faculty of Education, The University of Hong Kong, Hong Kong // [2]Department of Curriculum and Instruction, The Education University of Hong Kong, Hong Kong // davyngtk@connect.hku.hk // roxannexxfang@gmail.com
[*]Corresponding author

**ABSTRACT:** Due to the varied social inequities, the pandemic has prompted unprecedented attention to the social divides in learning gains via online/blended learning. It has been identified that approximately 27,000 Chinese cross-border students live on the Mainland but attend Hong Kong schools every day. The pandemic has restricted the passage of the Shenzhen-Hong Kong border, and these students suffer greatly as they may still be forced to attend online lessons at home for over two years, even if schools return to face-to-face teaching. This study is a 2-year longitudinal case study with a purposeful sampling of a primary school that has over 70% of Chinese cross-border students. A mixed-methods approach was adopted to examine students' online learning experiences via teachers' interviews. Regarding their learning challenges, 35 students were invited to complete a 5-point Online Self-regulated Learning Questionnaire (OSLQ) and an open-ended opinion survey, which was then triangulated with teachers' feedback. First, this study has identified three major challenges that cross-border students suffer the most in online learning: self-regulation strategies, technical challenges, and social interaction. Second, several teaching strategies have been found to offer additional support (e.g., after-school classes and learning centers in Shenzhen) for cross-border students to sustain their online learning and alleviate their learning challenges. Based on these findings, this study contributes to documenting the learning challenges for cross-border students during the pandemic, which provides insights for other schools to facilitate cross-border education and develop appropriate online learning strategies for cross-border students.

**Keywords:** Cross-border students, Online learning, Pandemic, Social divide

## 1. Introduction

Due to the COVID-19 pandemic, approximately 1.6 million learners across more than 200 countries have been significantly impacted by the shift from face-to-face to online learning modes (Pokhrel & Chhetri, 2021; Ng et al., 2020). After two years of extended online learning, the digital divide in society has widened, primarily due to varied social inequities. For instance, Mathrani et al. (2021) reported that some developing countries (e.g., India, Pakistan, and Bangladesh) have experienced exacerbated digital inequity because the quality of online learning relies heavily on students' technological abilities and digital devices. Recent research has revealed how millions of children's educational trajectories have been disrupted and damaged during the pandemic due to social inequity (e.g., Bacher-Hicks et al., 2021; Jæger & Blaabæk, 2020). Such inequity results in unequal learning gains due to government policies, school background, teacher qualifications and experience, students' learning resources, and technologies. Additionally, other factors such as special educational needs, socioeconomic status, self-regulated learning skills, family support, and learning environments also significantly impact students' learning outcomes (Blundell et al., 2020).

During the pandemic, children from low-income families are at greater risk of exclusion from online learning due to insufficient internet or digital devices. Students in rural areas may also face poor internet connections, which can negatively impact their online learning. These issues have been widely discussed in educational studies. However, among these factors, the phenomenon of cross-border learning has been identified but has yet to be explored in academic literature. Cross-border student issues are common across the globe, such as Singapore-Malaysia and USA-Mexico cross-border students who live in their own countries and receive education in neighboring countries (Orraca et al., 2017; Yuen & Cheung, 2014). In Hong Kong, around 27,000 Chinese cross-border students live on the Mainland but attend Hong Kong schools every day (Legislative Council, 2021). However, the pandemic has restricted the passage of the Shenzhen-Hong Kong border, and these students are greatly affected as they may still be forced to attend online lessons at home for over two years, even if schools return to face-to-face teaching. This situation has arisen due to the Hong Kong government's 2011 policy allowing doubly non-permanent resident pregnant women to give birth in Hong Kong (Chan, 2018). These cross-border students are Hong Kong permanent residents, but their parents are Mainland residents who live in the Mainland. Therefore, these students crossed the Shenzhen-Hong Kong border every day before the COVID-19 pandemic. This issue highlights the need for further research on the challenges faced by cross-border students during the pandemic, as well as the development of appropriate online learning strategies to support their education.

Due to the pandemic, Chinese cross-border students were unable to return to their schools for nearly two years (Legislative Council, 2021). The Hong Kong government and schools have developed innovative ways to help these students interact with their classmates, such as setting up learning centers on the Mainland or supporting their learning at home through online platforms. However, these policies cannot be consistently implemented as they depend on the pandemic situation. When the Education Bureau (EDB) announced a whole-school resumption of half-day from May 2021 to January 2022, with the pandemic under control, cross-border students still had to adhere to travel restrictions between Mainland China and Hong Kong. As per the rules, students could only stay on the Mainland to sustain online learning. Without teachers' face-to-face support, students had to independently carry out learning activities in online lessons and self-paced learning tasks, which relied heavily on their motivation and self-regulation. As a result, Chinese cross-border students faced numerous challenges related to online learning during the pandemic, such as social interaction, self-regulation, teachers' teaching strategies, government and school policies, and parental support. It is essential to address these challenges to ensure that cross-border students receive a quality education and support their academic development. Further research is needed to identify effective strategies to address the challenges faced by these students and develop appropriate policies to facilitate cross-border education.

Numerous studies have focused on online learning during the COVID-19 pandemic, reporting on the challenges of learning and teaching caused by remote learning (Adedoyin & Soykan, 2020; Carrillo & Flores, 2020; Rasheed et al., 2020). However, there is a lack of research on online learning experiences at the process level, such as social interaction, for both students and teachers (Rasheed et al., 2020). Therefore, it is crucial to study online learning using a longitudinal approach to analyze the process level, such as the meso-level with measures taken over several months. To address these gaps, this article aims to investigate the longitudinal approach of learning perceptions of Chinese cross-border students at the macro-level (two-year online learning). The study aims to uncover how cross-border students and their teachers perceive their online learning experiences over two years during the pandemic. Furthermore, this study also analyzes the educational policies for cross-border students implemented by the Hong Kong government and schools. By providing a holistic picture, we can understand how cross-border students, their teachers, the Hong Kong government, and schools experience different online learning challenges during COVID-19. The following research questions are identified:

- What are the major challenges that cross-border students perceive in their online learning?
- What strategies have been used to handle cross-border students' online learning challenges, and how have these strategies been employed?

## 2. Literature review

### 2.1. The context of Chinese cross-border students in Hong Kong

Influenced by globalization, some developed countries/regions with a well-known reputation in the education domain have attracted students from other countries (Knight, 2005). Many students aspire to pursue their education in these countries/regions (Knight, 2005). In such scenarios, cross-border/boundary students live in their home country but need to cross the border to study in other countries/regions (Chan et al., 2020). For instance, owing to the similar cultural heritage and high education quality in Singapore, most Malaysian families believe that Singapore offers an international perspective and provides more future education opportunities for their children (Yuen & Cheung, 2014). Consequently, many Malaysian students cross the border to attend schools in Singapore. In another comparable case, Mexican students cross the boundary to study in the USA because the educational quality of the USA is superior to that of Mexico (Orraca et al., 2017). The Mexico-USA cross-border students undergo a similar experience to most cross-border students in other countries, such as adapting to different languages and cultures in two countries/regions (Piedra & Araujo, 2012).

Among Asian countries/regions, there is a unique scenario between Shenzhen (a city in Mainland China) and Hong Kong, where thousands of students are in line to cross the border every day (Chan et al., 2020; Chan, 2018). These students (approximately 25,000) were born in Hong Kong but live in Shenzhen because their parents are non-permanent residents of Hong Kong. This phenomenon arose due to a policy implemented in 2011 that allowed Mainland pregnant women (doubly non-permanent resident pregnant women) to give birth in Hong Kong (Chan, 2018). At the end of 2012, the Hong Kong government announced the "zero doubly non-permanent resident quota" policy, which significantly reduced the number of cross-border students after that year (Policy Address, 2013). Despite this, these students face difficulties when seeking enrollment in Mainland schools because they are permanent residents of Hong Kong rather than Mainland citizens (Chan et al., 2020). If cross-border students prefer to study on the Mainland, they need to attend international or Hong Kong children's schools with high tuition fees, which most families cannot afford. Therefore, these Mainland parents send their children to Hong Kong local schools, from kindergarten to secondary school (Leung & Waters, 2022). One of the influencing factors is that Hong Kong education has an excellent global reputation and provides more international opportunities for children to advance their development. Therefore, most Mainland parents believe that Hong Kong's education is superior to that of the Mainland.

### 2.2. Online learning situation for cross-border students

During the pandemic, the government expressed concerns about the online learning situations for cross-border students in Hong Kong. According to a government document, there were approximately 27,000 students (as of October 2020) who lived in Mainland China but had to cross the border every day to receive K-12 education in Hong Kong before the COVID-19 pandemic (Legislative Council, 2021). However, the pandemic disrupted their learning when various social isolation policies were announced. These cross-border students could not attend physical lessons in Hong Kong, even though schools started half-day face-to-face learning after May 24, 2020 (Legislative Council, 2021). The document listed six dimensions of support measures to facilitate cross-border students' online learning: (1) supporting cross-border students to learn online at home, (2) addressing students' emotional needs, (3) flexible assessment that facilitates cross-border students' learning progress, (4) arrangements for primary one and secondary one cross-border students' admission, (5) offering additional learning courses and services for cross-border students at temporary centers in Shenzhen, and (6) other supports for the renewal of home return permits (Legislative Council, 2021). For example, the Hong Kong government has arranged two temporary learning courses in Shenzhen (Legislative Council, 2021). One is the "Psychosocial Support Course" in the Futian district, and the other is the "Learning Support Course" in the Nanshan district. The two entrusted agencies provide additional services for cross-border students and their parents, including practical information, learning materials, emotional support materials, free borrowing of books, recreational supplies and facilities, and arranging social activities such as networking events and sharing sessions to facilitate their social interaction (Legislative Council, 2021). These measures provide cross-border students with additional resources to facilitate their distance learning via face-to-face group teaching and social activities to meet their cognitive and social needs.

According to the report, parents of cross-border students understood that their children would unavoidably lag behind, compared to other local students, in terms of learning progress and academic performance (International Social Service, 2020). The report showed that over 50% of the families of 3,000 cross-border students worried about their online learning situation due to class suspension arrangements. Several reasons were proposed. First, cross-border students could not participate in face-to-face lessons due to the border closure between Hong Kong and the Mainland during the pandemic. Moreover, they could hardly use some e-learning platforms in the Mainland due to internet censorship (which is not specific to the pandemic), and many domain names are controlled or blocked. These include some popular websites for learning purposes, such as Google Suites, YouTube, and Wikipedia, which cause a disjuncture between the two digital spaces across the Hong Kong-Mainland border (Leung & Waters, 2022). Furthermore, 64% of the parents were concerned about discrimination their children might face when they eventually returned to school. They hoped that schools should minimize the negative impact caused by the pandemic, provide enough time for their children to prepare for class resumption, and offer additional online learning support (International Social Service, 2020).

The learning challenge faced by cross-border students during online learning has captured the interest of scholars. Two recent research papers have discussed various learning difficulties and inequities among cross-border students. The pandemic has exacerbated digital inequity between local and cross-border students (Chan et al., 2020; Leung & Waters, 2022). The studies have identified challenges such as accessing learning materials, stress among cross-border students and their parents, sense of belonging, and online schooling experiences (Chan et al., 2020; Leung & Waters, 2022). These challenges demonstrate that cross-border students face various learning difficulties compared to local students and offer a rationale for this study.

### 2.3. Challenges of learning and teaching in online learning during the COVID-19 pandemic

The transition from face-to-face to online learning has posed direct and deep challenges for students. Many prior studies have concluded multiple challenges of online learning (e.g., Adedoyin & Soykan, 2020; Rasheed et al., 2020). For example, Adedoyin and Soykan (2020) mentioned seven challenges in online learning during the COVID-19 pandemic, including technology, socio-economic factors, digital competence, assessment, and supervision. Particularly, the quality of technology is essential and compulsory for students' online learning. Also, students' digital competence is directly related to problem-solving skills, information management, and collaboration with effectiveness and efficiency (Ferrari, 2012). Students with low digital competence struggle in online learning, which is not good for academic performance. In another research study, Rasheed et al. (2020) reported that students' challenges in online learning included self-regulation, technological literacy and competency, students' isolation, technological sufficiency, and technological complexity challenges. The most severe challenge was discovered to be self-regulation, which reflects students' procrastination behavior (AlJarrah et al., 2018) due to the flexibility and autonomy offered by online learning. Thus, online learning requires students to have more self-control to cover the disadvantages of independent learning and fewer online interaction characteristics to avoid procrastination (Hong et al., 2021). Another influencing factor is the online help-seeking challenge. Broadbent (2017) mentioned that students could not get suitable help when not face-to-face. Er et al. (2015) also suggested that students were challenged in their technological skills when seeking help in online learning. These factors will directly influence students' self-regulated learning because they will be unable to operate the online learning system using the correct technological skills. Other factors influencing students' self-regulation challenges are a lack of self-regulation skills (Chuang et al., 2018), limited preparation before class (Broadbent, 2017), poor time management skills (Akçayır & Akçayır, 2018), and improper utilization of online peer learning strategies (Cakiroglu & Ozturk, 2017).

Apart from students' learning challenges, there are also teaching challenges faced by educators, who play the most important role in online teaching during the COVID-19 pandemic. Several studies have analyzed the challenges of teachers' teaching during the pandemic period (Rasheed et al., 2020; Carrillo & Flores, 2020). For example, Rasheed et al. (2020) conducted a systematic review of challenges in online learning, pointing out a set of challenges for teachers during online teaching, including teachers' technological literacy and competency, online video, technological operations, and teachers' beliefs. Among these challenges, the major ones are teachers' technological literacy and competency. Rasheed et al. (2020) mentioned several reasons for this, such as a lack of experience with using the online platform for creative instruction (Cheng & Chau, 2016), a lack of confidence and time to learn technology applications (Brown, 2016), and a lack of technological competence (Boelens et al., 2017). Another example is Carrillo and Flores (2020), who conducted a literature review on teacher education in online learning under COVID-19. They mentioned teachers' limited competence in using digital instructional formation (Huber & Helm, 2020) and insufficient experience in online teaching. Thus, it is essential to promote professional training to enhance teachers' technological literacy and competence. Otherwise, online learning may cause more and more disadvantages for teachers and students, such as ineffective teaching outcomes and low learning efficiency.

### 2.4. Self-regulated learning

This study used a questionnaire adopted from the Online Self-Regulated Learning Questionnaire to measure learners' self-regulation, which is widely used in secondary schools across various disciplines (e.g., Alten et al., 2021; Lau, 2021; Zalli et al., 2020). This questionnaire was developed based on the 14 self-regulation theories proposed by Zimmerman and Schunk (2011), which include three important stages: forethought and planning, performance control, and self-reflection (Zimmerman, 2000). The first stage requires students to analyze the learning task, set goals for positive actions, seek more information and resources for the specific task content, and establish a suitable environment for studying. The second stage involves note-taking, organizing and memorizing facts/knowledge, seeking assistance and learning support, and managing time well. The final stage focuses on students' reflection during and after completing their learning tasks so that they can evaluate their performance, learning strategies, and outcomes according to the goals set (Fung et al., 2018).

The questionnaire used in this study consists of 24 items with six dimensions of self-regulated learning, including goal-setting, environmental structuring, task strategies, time management, help-seeking, and self-evaluation, developed by Lau (2021). The validity and reliability of this questionnaire were reported to be good among 417 Hong Kong students (Grades 4-9, aged 9-15), and researchers were recommended to use it across different learning settings. In our study, the school context, such as location and cultural background, is also in Hong Kong, and our student participants are from grade 5 with an age of 10. Hence, the questionnaire is suitable for this study to examine online self-regulation behaviors among cross-border students.

The six aspects of the questionnaire are explained as follows:

- Self-regulated learning (SRL) involves students' efforts to manage their learning processes oriented towards achieving goals (Zimmerman & Schunk, 2011).
- Goal-setting involves students' development of an action plan to motivate and guide them towards a goal (e.g., staying attentive during the online class, achieving good performance on the exam) (Cho & Shen,

2013; Latham & Locke, 1991). It helps students commit their thoughts, emotions, and behavior towards attaining the goal.

- Environmental structuring examines how the online learning environment can be rearranged to avoid learning distractions and enhance students' motivation and achievement (Yen et al., 2016).
- Task strategies refer to the positive behaviors students develop using various strategies to overcome their online learning challenges (e.g., collaboration, note-taking, reading aloud, using technologies to facilitate learning) (Barnard-Brak et al., 2010).
- Time management consists of scheduling and distributing students' time for learning (Yen et al., 2016).
- Help-seeking allows students to find teachers, classmates, and parents for help, whereas self-evaluation encourages them to summarize and evaluate their learning progress in online courses to examine their understanding (Whipp & Chiarelli, 2004).

## 3. Methods

### 3.1. Participants

This article outlines the research design of a longitudinal study in Hong Kong aimed at describing the impact of cross-border learning on cross-border students. This study utilized mixed methods to collect data through open-ended surveys and questionnaire responses from 35 primary five cross-border students (17 boys and 18 girls) and semi-structured interviews of their teachers ($n$ = 3). Purposeful sampling was used to identify a primary school that was reported to have 70% (around 140) of its students residing in the Mainland but studying in Hong Kong. In other words, all participating students were cross-border students who claimed that the pandemic had restricted their travel across the border, and attending online lessons was their only option. Primary five students were chosen because they undergo a significant transition from primary four to five. They need to participate in internal assessments for secondary school place allocation, which can cause great examination pressure on most primary students (Berry, 2011).

The researchers collaborated with three teachers from the school who were responsible for the online teaching transition for over two years. These teachers voluntarily agreed to participate in individual interviews to provide a comprehensive picture of the cross-border students' online teaching experience and how they overcame any teaching challenges. Based on their observations, the researchers conducted interviews with the teachers to understand cross-border students' learning experiences, challenges, and other learning challenges they faced during the pandemic. Table 1 displays the demographic information of the teacher interviewees.

*Table 1.* Demographic information of the teacher interviewees

| Name (fictitious) | Gender | Teaching subject | Years of teaching | Job title(s) |
|---|---|---|---|---|
| Chan | Male | Mathematics | 8 | Curriculum Coordinator and IT Officer |
| Chu | Female | English | 4 | Headteacher |
| Wong | Female | Chinese | 7 | Chairman of the Chinese Division |

### 3.2. Data collection and analysis

This study examined students' self-regulation of online learning within two periods (from Jan 2020 to July 2020 and Sep 2021 to Jan 2022). To analyze the learning challenges for cross-border students, the study revealed how students perceived their online learning challenges through an opinion survey, followed by a questionnaire. Qualitative questions (e.g., What challenges do you encounter when learning online?; Do you employ private tutors to support your learning?; How do your parents, schools, and community centers support your learning?) were asked to identify major learning issues students faced. Students then self-reported in the survey what challenges they encountered when learning online. The first and second authors worked together to categorize students' feedback into various key challenges, including lacking self-regulation, poor academic performance, lacking parental support, being distracted, meeting technical challenges, lacking electronic equipment and internet network, and lacking social interaction and communication. To ensure reliability, students' feedback was coded according to the identified challenges (see Table 2). The inter-rater reliability between the researchers was found to be good, with a Cohen's kappa coefficient of 0.77 (Watkins & Pacheco, 2000).

According to the survey, self-regulation was found to be the most significant challenge. This was further examined based on students' questionnaire responses to understand their perceived ability and its six related behaviors. The Chinese version of the Online Self-Regulated Learning Questionnaire (OLSQ) was adopted and evaluated by previous research (Fung et al., 2018). The OLSQ comprises 24 items with a 5-point Likert scale, examining six aspects of online learning: goal setting, environmental structuring, task strategies, time management, help-seeking, and self-evaluation.

To triangulate the quantitative results, qualitative data was collected to provide a fuller picture of their online learning affecting their academic outcomes and social life via students' and teachers' interviews. Throughout the interviews, students' online learning challenges and teaching strategies were documented via interviews to understand what types of support teachers and the school offered to facilitate cross-border students' online learning according to the two research questions. A set of interview protocols was used to facilitate researchers to categorize into themes that would greatly impact cross-border education. After conducting the interviews, their recordings and reflection were transcribed and coded. Inductive thematic analysis was adopted and grouped patterns of similar meanings into two themes based on the research questions: cross-border students' online learning challenges and strategies to facilitate cross-border students' online learning for cross-border students' online learning.

Likewise, the two researchers then further coded and analyzed the dialogues according to the themes to ensure reliability (see Table 2). To establish coding reliability, all of the dialogues were analyzed by the two researchers. Disagreements were settled through discussion. Cohen's kappa coefficient (0.86) was found to be good to show inter-rater reliability between the researchers (Watkins & Pacheco, 2000). With qualitative and quantitative methods and elements, this study can triangulate and combine results into a comprehensive picture of cross-border education.

*Table 2.* Coding scheme of research questions

| Topics | Themes | Definitions/ Descriptions | Sample sentences |
|---|---|---|---|
| RQ1. Cross-border students' online learning challenges | Self-regulated learning | Students' efforts to manage their learning processes are oriented to achieving goals (Zimmerman & Schunk, 2011). | *"Students have no big improvement in self-regulated learning. We can see that self-regulation is really a challenge for cross-border students…"* |
| | Technical challenges | Students might not have enough technical equipment and Internet access at home. | *"Some students don't have enough digital devices…"*<br>*"One of the most serious problems is technical difficulties…"* |
| | Social interaction | The opportunities for students to interact, communicate, and collaborate with classmates and teachers. | *"There is a lack of interaction with the teachers and local students in Hong Kong."* |
| RQ2. Strategies to facilitate cross-border students' online learning | Cognitive support | Support that enables students to construct knowledge in an online learning environment (Garrison, 2007). | *"We used Kahoot (a gamified platform), Seesaw (a digital mind map platform), and Pallet (a social media and collaborative tool)."* |
| | Social support | Support that enables students to socialize with others through online learning activities (Garrison, 2007). | *"We tried hard to make students socialize by using the functions in Zoom, such as group discussion, voting, drawing, and other interactive functions."* |

## 4. Results and discussion

### 4.1. RQ1. What are the major challenges that cross-border students perceive in their online learning?

Opinion surveys have identified three major challenges (i.e., self-regulation, technical challenges, and social interaction). Among the 35 cross-border students, 15 (42.9%) claimed they were not satisfied with the online learning experience compared to the face-to-face classes they attended before the pandemic. Most of them declared that they lacked self-regulation due to an inactive online learning atmosphere ($n$ = 22, 62.9%), negatively impacting their academic performance ($n$ = 16, 45.7%). One reason students suggested is that they could easily be distracted during online learning at home ($n$ = 17, 48.6%), especially without much parental support ($n$ = 16, 45.7%). Furthermore, students encountered various technical difficulties, such as a lack of electronic equipment and internet network ($n$ = 16, 45.7%) and low social interaction and communication with their teachers/classmates ($n$ = 15, 42.9%).

To examine how prolonged online learning would affect students' self-regulation, this study administered the QLSQ to understand students' perceived ability and the six aspects of self-regulation behavior (goal setting, environmental structuring, task strategies, time management, help-seeking, self-evaluation) ($p$ > .05). Although other local students also faced the three online learning challenges, cross-border students had no face-to-face learning opportunities throughout the two years, which exacerbated the negative impact of online learning. The results indicated no significant learning improvement in the pre-test and post-test for all items (see Table 3). In other words, cross-border students' improvement is not significant even though the teachers have worked very hard to support cross-border students' online learning. These results aligned with the above-mentioned opinion survey results and were triangulated with the support of qualitative dialogues.

*Table 3.* Self-regulation and perceived abilities of cross-border students in their pre and post-tests

| | Mean (T1) | *SD* (T1) | Mean (T2) | *SD* (T2) | Sig. level (SE) |
|---|---|---|---|---|---|
| Goal setting | 3.55 | 0.89 | 3.59 | 0.92 | 0.87 (0.21) |
| Environmental structuring | 3.71 | 1.01 | 4.00 | 0.90 | 0.16 (0.20) |
| Task strategies | 3.44 | 1.10 | 3.80 | 1.02 | 0.16 (0.25) |
| Time management | 3.37 | 1.02 | 3.70 | 1.07 | 0.19 (0.24) |
| Help-seeking | 3.29 | 1.01 | 3.47 | 1.01 | 0.13 (0.23) |
| Self-evaluation | 3.38 | 0.99 | 3.74 | 0.96 | 0.30 (0.22) |
| Perceived ability | 3.43 | 1.08 | 3.66 | 0.84 | 0.44 (0.23) |

#### *4.1.1. Self-regulation strategies*

Over two years of online teaching, cross-border students unavoidably fell behind in their learning progress compared to other students. One possible reason is that social isolation policies prevented cross-border students from returning to schools, even when all schools in Hong Kong resumed half-day face-to-face teaching from May 2021 to January 2022. As a result, cross-border students could only continue studying online at home without any physical lessons, while other local students could still receive face-to-face or blended lessons at school. The opinion survey revealed that the biggest challenge perceived by cross-border students was a lack of self-regulation skills.

A teacher believed the situation was understandable, "Students have no big improvement in self-regulated learning. We can see that self-regulation is a challenge for cross-border students. The results do not surprise us. We cannot expect students to have great improvements." Then, the authors interviewed teachers to identify why students were not self-regulated. A teacher mentioned that this was associated with parental support, "Not all students have their parents to guide them to learn at home. It is not easy for young children to attend web-conferencing lessons and study self-paced materials alone. They easily get distracted and do other interesting things while having lessons. Those students who have their parents tutoring them will have better academic achievements. However, most parents have full-time jobs, so online learning relies on students' self-regulated learning." Another teacher expressed sympathy for the cross-border students since the pandemic created an unequal educational opportunity. The teachers had difficulty narrowing the inevitable and unfair gap, "We have never physically met the students. Supporting them in a solely online environment for two years is not easy. We have done our best to improve our teaching quality, but we cannot force our students to stay attentive (and self-regulated) during the online lessons." All these conversations provided evidence from teachers' perspectives that students tended to have weak self-regulation skills to stay motivated and engaged in online learning.

Effective SRL strategies are critical in online learning, given the high degree of student autonomy from the instructor's physical absence (Barak, 2022). Students with greater SRL abilities tend to have a better goal orientation, academic self-efficacy, and regulations in the learning contexts, making them have higher learning abilities (Cho & Shen, 2013). Furthermore, students with SRL skills have positive learning behaviors such as supporting peer work, recommending their notes, helping each other, and displaying a high level of assessment (Wong et al., 2019). As such, teachers need to develop cross-border students' self-regulation strategies, such as asking students to consider how they learn online, providing pacing support, monitoring engagement, and supporting families (Carter Jr et al., 2020).

#### *4.1.2. Technical challenges*

Some cross-border students have reported a lack of necessary technical equipment and internet access at home. While local students could borrow equipment and receive roaming data passes from their schools, the school hardly supported cross-border students in this regard, as it was impossible to deliver equipment to their homes on the Mainland. For instance, some students reported not having tablet devices to facilitate their learning, and their network connection was unsatisfactory. Mr. Chan commented, “Some students do not have enough digital devices. They can only use WeChat or Zoom to facilitate their online learning via mobile phones. For those who do not have digital devices at home, it is impossible to deliver computers or laptops. Local students could easily borrow equipment, but cross-border students could not get these devices. We can seek support from the community center on the Mainland.”

Even students with sufficient digital equipment have reported that the school’s solely online technical support could hardly solve their problems. Mr. Chan, a teacher, explained, “One of the most significant problems is technical difficulties... When they need technical support, it is hard to support them online. We can only rely on their parents to guide them on using e-learning tools at home so they can continue with online learning smoothly. However, not all parents are technologically literate enough to offer timely support.” Other studies have also found that distance learners may face more technical challenges than local online learners, such as a lack of digital equipment, slow internet speeds, and difficulty setting up learning platforms (Yan et al., 2021). Local students could still return to school and gain support, but distance learners such as cross-border students can only resolve their problems online.

Due to internet censorship in the Mainland, as discussed in the literature review, many domain names are controlled, making it difficult for cross-border students to access educational websites and applications (Leung & Waters, 2022). This affects the quality, motivation, and overall experience of their online learning. Students have reported that some websites, even those for educational purposes, are prohibited on the Mainland, making it challenging to access online materials set by their teachers. Mr. Chan commented, “We need to consider if cross-border students can access websites at home. Some websites are blocked on the Mainland... Sometimes, students report to us that they cannot access learning materials. In such cases, we need to send them through alternative ways like email or WeChat.”

#### *4.1.3. Social interaction*

The third challenge faced by students is the difficulty in interacting and socializing with their classmates through online learning activities (Richardson et al., 2017). This is unavoidable as cross-border students can only participate in online learning even when schools have resumed face-to-face classes. Social support could provide learning opportunities for students to co-construct knowledge through varied pedagogies such as collaborative project-based learning and discussion. However, cross-border students could only interact with classmates online. A student shared their evaluations of social support, “I cannot learn much and can’t meet new friends in my school. There is a lack of interaction between teachers and local students in Hong Kong. They virtually appear there.” A teacher also pointed out that some activities are better conducted face-to-face (e.g., music, sports, and arts), which require many social interactions. Cross-border students could only watch video demonstrations or participate in Zoom sessions, and not all activities can be transformed online due to their hands-on nature.” Although the school tried to implement some practices to improve students’ learning experiences, it still did not work due to the pandemic. Another teacher stated, “We thought about arranging institutions in Shenzhen to organize group classes and other extracurricular activities, but we did not implement them eventually because the pandemic became severe.” This is consistent with previous findings that social isolation strongly impacts young students (Zhang et al., 2020). Therefore, due to the pandemic, online learning interaction plays a vital role in managing students’ stress due to social isolation.

### 4.2. RQ2. What strategies have been used to handle cross-border students’ online learning challenges, and how have these strategies been employed?

In the previous section, this study highlighted that cross-border students faced additional challenges (such as a purely online learning environment, lack of socialization, and restricted access to learning resources due to geographical limitations) that local students did not encounter during their online learning. Inspired by Garrison’s (2007) Communities of Inquiry in Online Learning, teaching strategies could include curriculum planning, material preparation, teaching and learning to construct students’ knowledge and enhance their socialization. This section aims to evaluate the teaching strategies designed to alleviate cross-border online learning challenges in terms of cognitive and social support.

Proper cognitive support can provide students with a better learning experience than delivering content through presentation support tools (Bernard et al., 2014). To support cross-border students' academic performance, teachers have employed various strategies, such as using social media, collaborative tools, recorded videos, mind maps, and gamification, to enhance students' engagement and motivation. One teacher reflected on her teaching, "We used Kahoot (a gamified platform), Seesaw (a digital mind map platform), and Pallet (a social media and collaborative tool). We also tried various web-conferencing platforms like Teams and Zoom to facilitate synchronous learning. We want to make the lessons more engaging and interesting." To address the issue of unequal learning opportunities and provide diversified digital learning resources and pedagogies, teachers provided students with supplementary learning activities. A teacher mentioned, "You cannot imagine how young children 'go to school' purely online for two years. They lack self-regulation and learning autonomy, which has worsened their academic performance. Online learning requires great effort to stay motivated, attentive, and self-directed. It is hard for young children, especially when their parents have little time to accompany them. No matter how many make-up classes we offer, if they do not take the initiative in their learning, they do not have a good learning performance."

However, despite the cognitive support measures such as additional lessons and digital tools given to students, they still struggled to stay engaged and motivated. A teacher attributed this to the lack of parental support. "Most parents need to work, so cross-border students have no one to take care of them. We called their parents to gain their support (e.g., employing private tutors). We hope the pandemic will be over soon, and the government will relax the policy to allow cross-border students to travel and study here." Overall, it is evident that cross-border students were not satisfied with the online learning experience, even though teachers worked hard to enhance their academic achievement through cognitive support. Therefore, teachers sought support from parents to further alleviate the challenges faced by cross-border students.

Social support is crucial for students to gain knowledge and socialize with classmates through well-designed online learning activities (Richardson et al., 2017). It provides many opportunities in an online learning environment, such as enhancing student motivation and participation, actual and perceived abilities, learning satisfaction, and retention in schooling (Ng, 2022a; Ng, 2022b). Two teachers provided feedback, "We tried hard to make students socialize by using the functions in Zoom, such as group discussions, voting, drawing, and other interactive functions." "Some introverted students seem to have greater socialization opportunities. Compared to face-to-face teaching, they actively participate in our e-learning activities and online courses. They seldom speak during classes because they are unfamiliar with the local language. With online communication tools (e.g., social media, instant messaging software), they can type and interact using online tools that help them communicate." These measures can facilitate students' interaction, socialization, and collaboration to maintain learning and assessment with peers in an online environment. Social support is echoed among students and teachers as their feedback and contributions during lessons are social capital as resources that empower online learners (McIntyre, 2021).

Although teachers have employed strategies similar to those used for other local students, cross-border students believed that these strategies were not very useful for their learning. Some students employed private tutoring ($n$ = 18, 51.4%), but other cross-border students did not have many resources for private tutoring. Some claimed they were not equipped with enough technological resources ($n$ = 16, 45.7%) and did not have the knowledge/skills to use IT for online learning ($n$ = 6). The saying "fair is not always equal" applies here. To reduce the cross-border students' learning gap, additional resources, including strategies to promote self-regulated learning, technological resources, and cognitive and social support, should be offered to facilitate cross-border students' learning. However, due to geographical restrictions, they may never be able to obtain the same learning experiences as local students.

## 5. Conclusion

The pandemic has brought unprecedented attention to educational inequity and challenges through online/blended learning (Ng et al., 2020). Researchers have conducted numerous studies to document snapshot studies and the major types of social inequities for various reasons (such as special education needs, socioeconomic status) (Chan et al., 2020; Leung & Waters, 2022). Cross-border education is rarely discussed in prior studies, but it is a critical issue that educators should be aware of. Due to prolonged online teaching, cross-border students in Hong Kong suffered more in terms of academic performance and socialization due to travel restrictions. They were forced to participate in purely online learning for over two years. While other local students in Hong Kong could engage in face-to-face teaching, club activities, community service, and competitions, cross-border students could only participate in these activities purely online. Overall, cross-border students could not enjoy an equal opportunity to learn cognitively and socially compared to their counterparts.

To gain an understanding of cross-border education, the research team worked with three teachers and empathized with their stress and helplessness for two years. Strategies were applied to facilitate learning, including government, school, and teacher support (such as setting up community centers on the Mainland and using web-conferencing software/social media to communicate). Although their school devised creative solutions to help cross-border students overcome their online learning challenges, most students claimed that teacher support was insufficient. Cross-border students still face various difficulties, such as a lack of technological resources, motivation, and self-regulation. According to interviews and questionnaires, two concerns were raised by most students, self-regulation and social interaction, and they attributed their failure to low perceived ability. Firstly, solely online learning relied heavily on students' motivation and self-regulation. Without teacher or parent supervision, engaging cross-border students in the online learning environment was not easy. Secondly, the school was not only a learning location for students to acquire knowledge and skills but also a community for students to socialize and enjoy campus life.

Furthermore, dropout rates and prolonged online learning have increased teachers' stress and worry. They claimed that they felt exhausted from making contingency plans for both cross-border and local students. When students applied to transfer schools, their school would have the risk of getting suspended due to an inadequate student-teacher ratio. Despite the difficulties perceived by teachers and students during the pandemic, they improved their technical skills to communicate, learn, and collaborate in an online environment. In our study, the qualitative measurement allowed cross-border students' and their teachers' voices to be showcased and enabled us to explore how other countries perceive similar issues in facilitating cross-border education.

## 6. Research limitations and recommendations for future research

Several limitations were identified in this study. Firstly, only one primary school was selected, which may not provide enough information to generalize the overall picture of the cross-border students' situation in Hong Kong. Secondly, the limited number of teacher and student interviewees may not represent the perspectives of other teachers in other cross-border students' schools. More interviews could generate important themes to make the arguments more convincing. Thirdly, this study lacks comparisons to show the inequity between cross-border students and local students. Future studies should compare the learning differences between the two student groups in Hong Kong. Moreover, feedback from other stakeholders (such as parents and principals) should be collected to fully understand their perspectives on the cross-border students' issue. With more evidence-based support, the government, schools, and educators could understand the challenges cross-border students face in their online learning journey and design appropriate strategies and interventions to facilitate their online learning. In this way, cross-border students could narrow their educational gap and continue their studies after class suspensions.


## Acknowledgement

We thank the teachers for bringing their expertise and assistance throughout the publication of this study.


## References


Adedoyin, O. B., & Soykan, E. (2020). Covid-19 pandemic and online learning: the challenges and opportunities. *Interactive Learning Environments*, *31*(2), 863-875. https://doi.org/10.1080/10494820.2020.1813180

Akçayır, G., & Akçayır, M. (2018). The flipped classroom: A review of its advantages and challenges. *Computers & Education*, *126*, 334-345. https://doi.org/10.1016/j.compedu.2018.07.021

AlJarrah, A., Thomas, M. K., & Shehab, M. (2018). Investigating temporal access in a flipped classroom: Procrastination persists. *International Journal of Educational Technology in Higher Education*, *15*(1), 1–18.

Alten, D. C., Phielix, C., Janssen, J., & Kester, L. (2021). Secondary students' online self-regulated learning during flipped learning: A Latent profile analysis. *Computers in Human Behavior*, *118*, 106676. https://doi.org/10.1016/j.chb.2020.106676

Bacher-Hicks, A., Goodman, J., & Mulhern, C. (2021). Inequality in household adaptation to schooling shocks: Covid-induced online learning engagement in real time. *Journal of Public Economics*, *193*, 104345. https://doi.org/10.1016/j.jpubeco.2020.104345

Barak, M. E. M. (2022). *Managing diversity: Toward a globally inclusive workplace*. Sage Publications. https://www.amberton.edu/media/Syllabi/Fall%202022/Graduate/HRT6590_E1.pdf

Barnard-Brak, L., Paton, V. O., & Lan, W. Y. (2010). Profiles in self-regulated learning in the online learning environment. *International Review of Research in Open and Distributed Learning*, *11*(1), 61-80.

Bernard, R. M., Borokhovski, E., Schmid, R. F., Tamim, R. M., & Abrami, P. C. (2014). A Meta-analysis of blended learning and technology use in higher education: From the general to the applied. *Journal of Computing in Higher Education*, *26*(1), 87-122.

Berry, R. (2011). Assessment trends in Hong Kong: Seeking to establish formative assessment in an examination culture. *Assessment in Education: Principles, Policy & Practice*, *18*(2), 199-211.

Blundell, R., Costa Dias, M., Joyce, R., & Xu, X. (2020). COVID‐19 and inequalities. *Fiscal Studies*, *41*(2), 291-319.

Boelens, L., De Wever, B., & Voet, M. (2017). Four key challenges to the design of blended learning: A Systematic literature review. *Educational Research Review*, *22*, 1–18. https://doi.org/10.1016/j.edurev.2017.06.001

Broadbent, J. (2017). Comparing online and blended learner's self-regulated learning strategies and academic performance. *The Internet and Higher Education*, *33*, 24-32. https://doi.org/10.1016/j.iheduc.2017.01.004

Brown, M. G. (2016). Blended instructional practice: A Review of the empirical literature on instructors' adoption and use of online tools in face-to-face teaching. *The Internet and Higher Education*, *31*, 1-10. https://doi.org/10.1016/j.iheduc.2016.05.001

Cakiroglu, U., & Ozturk, M. (2017). Flipped classroom with problem-based activities: Exploring self-regulated learning in a programming language course. *Journal of Educational Technology & Society*, *20*(1), 337-349.

Carrillo, C., & Flores, M. A. (2020). COVID-19 and teacher education: a literature review of online teaching and learning practices. *European Journal of Teacher Education*, *43*(4), 466-487.

Carter Jr, R. A., Rice, M., Yang, S., & Jackson, H. A. (2020). Self-regulated learning in online learning environments: strategies for remote learning. *Information and Learning Sciences*, 121(5), 321-329.

Chan. P. W. K. (2018). Cross-border education for pupils of kindergartens and schools: The Case of Hong Kong. *International Education Journal*, *17*(3), 90–107.

Chan, A. K. W., Chiu, M. M., Yang, S., & Ngan, L. L. S. (2020). Mobility, belongingness and schooling experiences of Chinese cross-border students. *Children and Youth Services Review*, *111*, 104870. https://doi.org/10.1016/j.childyouth.2020.104870

Cheng, G., & Chau, J. (2016). Exploring the relationships between learning styles, online participation, learning achievement and course satisfaction: An Empirical study of a blended learning course. *British Journal of Educational Technology*, *47*(2), 257–278.

Cho, M. H., & Shen, D. (2013). Self-regulation in online learning. *Distance Education*, *34*(3), 290-301.

Chuang, H., Weng, C., & Chen, C. (2018). Which students benefit most from a flipped classroom approach to language learning? *British Journal of Educational Technology*, *49*(1), 56–68. https://doi.org/10.1111/bjet.12530

Er, E., Kopcha, T. J., Orey, M., & Dustman, W. (2015). Exploring college students' online help-seeking behavior in a flipped classroom with a web-based help-seeking tool. *Australasian Journal of Educational Technology*, *31*(5). https://doi.org/10.14742/ajet.2527

Ferrari, A. (2012). *Digital competence in practice: An Analysis of frameworks*. Publications Office of the European Union. https://doi.org/10.2791/82116

Fung, J. J., Yuen, M., & Yuen, A. H. (2018). Validity evidence for a Chinese version of the online self-regulated learning questionnaire with average students and mathematically talented students. *Measurement and Evaluation in Counseling and Development*, *51*(2), 111-124.

Garrison, D. R. (2007). Online community of inquiry review: Social, cognitive, and teaching presence issues. *Journal of Asynchronous Learning Networks, 11*(1), 61–72.

Hong, J. C., Lee, Y. F., & Ye, J. H. (2021). Procrastination predicts online self-regulated learning and online learning ineffectiveness during the coronavirus lockdown. *Personality and individual differences*, *174*, 1-8. https://doi.org/10.1016/j.paid.2021.110673

Huber, S. G., & Helm, C. (2020). COVID-19 and schooling: Evaluation, assessment and accountability in times of crises—reacting quickly to explore key issues for policy, practice and research with the school barometer. *Educational Assessment, Evaluation and Accountability*, *32*(2), 237–270.

International Social Service. (2020). *Over half cross border students and families concern about inequality and difficulty in online learning over class suspension arrangement.* https://www.isshk.org/en/news/latest_news_detail/93

Jæger, M. M., & Blaabæk, E. H. (2020). Inequality in learning opportunities during Covid-19: Evidence from library takeout. *Research in Social Stratification and Mobility*, *68*, 100524. https://doi.org/10.1016/j.rssm.2020.100524

Knight, J. (2005). Cross-border education: not just students on the move. *International Higher Education*, *41*(2), 1–3.

Latham, G. P., & Locke, E. A. (1991). Self-regulation through goal setting. *Organizational behavior and human decision processes*, *50*(2), 212-247.

Lau, K. L. (2021). Adaptation and validation of a Chinese online self-regulated learning questionnaire. *Journal of Psychoeducational Assessment, 40*(3), 07342829211059979. https://doi.org/10.1177/07342829211059979

Legislative Council (2021). *Education promoting home-school co-operation and parent education.* https://www.legco.gov.hk/yr20-21/english/panels/ed/papers/ed20210702cb4-1175-3-e.pdf

Leung, M. W., & Waters, J. L. (2022). Bridging home and school in cross-border education: The Role of intermediary spaces in the in/exclusion of Mainland Chinese students and their families in Hong Kong. *Urban Studies*, *59*(16) 3271–3291. https://doi.org/10.1177/00420980221084894

Mathrani, A., Sarvesh, T., & Umer, R. (2021). Digital divide framework: Online learning in developing countries during the COVID-19 lockdown. *Globalisation, Societies and Education*, *20*(5). https://doi.org/10.1080/14767724.2021.1981253

McIntyre, N. (2021). *What does post-pandemic educational research tell us about online learning experiences?* https://www.insendi.com/news-and-updates/what-does-post-pandemic-educational-research-tell-us-about-online-learning-experiences

Ng, T. K., Chu, S., Li, X., Reynolds, R., & Chan, M. (2020). Business (teaching) as usual amid the COVID-19 pandemic: A Case study of online teaching practice in Hong Kong. *Journal of Information Technology Education: Research*, *19*(1), 775-802.

Ng, T. K. (2022a). Online lab design for aviation engineering students in higher education: A Pilot study. *Interactive Learning Environments*, 1-18. https://doi.org/10.1080/10494820.2022.2034888

Ng, T. K. (2022b). Online aviation learning experience during the COVID-19 pandemic in Hong Kong and Mainland China. *British Journal of Educational Technology*, *53*(3), 443-474.

Orraca, P., Rocha, D., & Vargas, E. (2017). Cross-border school enrolment: Associated factors in the U.S.–Mexico borderlands. *The Social Science Journal (Fort Collins)*, *54*(4), 389–402.

Piedra, M., & Araujo, B. E. (2012). Literacies crossing borders: Transfronterizo literacy practices of students in a dual language program on the USA-Mexico border. *Language and Intercultural Communication*, *12*(3), 214–229.

Pokhrel, S., & Chhetri, R. (2021). A Literature review on impact of COVID-19 pandemic on teaching and learning. *Higher Education for the Future*, *8*(1), 133–141.

Policy Address. (2013). *Hong Kong Government Policy Address 2013*. https://www.policyaddress.gov.hk/2013/sim/p12.html

Rasheed, R., Kamsin, A., & Abdullah, N. A. (2020). Challenges in the online component of blended learning: A Systematic review. *Computers and Education*, *144*, 103701. https://doi.org/10.1016/j.compedu.2019.103701

Richardson, J. C., Maeda, Y., Lv, J., & Caskurlu, S. (2017). Social presence in relation to students' satisfaction and learning in the online environment: A Meta-analysis. *Computers in Human Behavior*, *71*, 402-417. https://doi.org/10.1016/j.chb.2017.02.001

Watkins, M. W., & Pacheco, M. (2000). Interobserver agreement in behavioral research: Importance and calculation. *Journal of Behavioral Education*, *10*(4), 205-212.

Whipp, J. L., & Chiarelli, S. (2004). Self-regulation in a web-based course: A Case study. *Educational Technology Research and Development*, *52*(4), 5-21.

Wong, J., Baars, M., Davis, D., Van Der Zee, T., Houben, G. J., & Paas, F. (2019). Supporting self-regulated learning in online learning environments and MOOCs: A Systematic review. *International Journal of Human–Computer Interaction*, *35*(4-5), 356-373.

Yan, L., Whitelock-Wainwright, A., Guan, Q., Wen, G., Gašević, D., & Chen, G. (2021). Students' experience of online learning during the COVID-19 pandemic: A Province-wide survey study. *British Journal of Educational Technology*, *52*(5), 2038-2057.

Yen, C. J., Tu, C. H., Sujo-Montes, L., & Sealander, K. (2016). A Predictor for PLE management: Impacts of self-regulated online learning on student's learning skills. *Journal of Educational Technology Development & Exchange*, *9*(1), 29-48.

Yuen, C. Y., & Cheung, A. C. (2014). School engagement and parental involvement: The Case of cross-border students in Singapore. *The Australian Educational Researcher, 41*(1), 89-107.

Zalli, M. M. M., Nordin, H., & Hashim, R. A. (2020). Online self-regulated learning strategies in MOOCs: A Measurement model. *International Journal of Engineering and Technology, 15*(8), 255-263.

Zhang, X., Zhu, W., Kang, S., Qiu, L., Lu, Z., & Sun, Y. (2020). Association between physical activity and mood states of children and adolescents in social isolation during the COVID-19 epidemic. *International Journal of Environmental Research and Public Health*, *17*(20), 7666. https://doi.org/10.3390/ijerph17207666

Zimmerman, B. J. (2000). Self-efficacy: An Essential motive to learn. *Contemporary Educational Psychology*, *25*(1), 82–91.

Zimmerman, B. J., & Schunk, D. H. (2011). *Handbook of self-regulation of learning and performance*. Routledge/Taylor & Francis Group. https://psycnet.apa.org/record/2011-12365-000

AlShawabkeh, A., Kharbat, F., Daabes, A. A., & Woolsey, M. L. (2023). Technology-based Learning and the Digital Divide for Deaf/Hearing Students During Covid-19: Academic Justice Lens in Higher Education. *Educational Technology & Society*, *26*(4), 136-149. https://doi.org/10.30191/ETS.202310_26(4).0010

# Technology-based Learning and the Digital Divide for Deaf/Hearing Students During Covid-19: Academic Justice Lens in Higher Education

**Abdallah AlShawabkeh[1], Faten Kharbat[2*], Ajayeb Abu Daabes[3] and M Lynn Woolsey[4]**
[1]College of Business, Al Ain University, UAE // [2]College of Engineering, Al Ain University, UAE // [3]College of Business, Liwa College of Technology, UAE // [4]Campbell County Public Schools TN, USA //
abdallah.alshawabkeh@aau.ac.ae // faten.kharbat@aau.ac.ae // ajayeb.daabes@ect.ac.ae //
drmarylynnwoolsey@gmail.com
[*]Corresponding author

**ABSTRACT:** This longitudinal mixed method study investigated the education experiences among 38 Deaf/Hard of Hearing (DHH) students, 44 of their hearing peers, and three lecturers in two programs at a postsecondary institution in Abu Dhabi, UAE. Longitudinal data were collected at three points in time, summer 2020, winter 2021, and spring 2021. To investigate the differences between the two groups, GPA marks and responses to designed questionnaires for the different moments were collected from students. Repeated-measures ANOVA, the fixed effect regression approach, and one-way ANOVA were used to analyze the quantitative data. Qualitative data were collected through semi-structured interviews and personal communication with the students and the lecturers. Interviews were conducted with the addition of a sign language interpreter familiar to the DHH students. Results revealed a significant distinct digital divide between the two groups in academic achievement, degree of adaptation, and general perceptions of their education. The results of this study suggest that DHH students in postsecondary settings may be at risk of underachieving in the distance learning study mode. As a result, higher education institutions, universities, and curriculum designers need to take steps to improve the environment of distance learning platforms considering the diversity of students' needs. This will ensure that students with hearing loss have equal access to lectures, course content, and their peers. Promoting equity and justice in the learning environment aims to bridge the knowledge gap and prepare DHH students for their future careers.

**Keywords:** Longitudinal study, Digital divide, Covid-19 crisis, Academic justice, Deaf/Hard of hearing students

## 1. Introduction

Simply defined, the digital divide is the "gulf" between the people who have what they need to access the internet and those who do not have what they need to access the internet. It may have been simpler in the 1990s when the term digital divide began to take hold (DeMartino & Weiser, 2021; Dijk, 2005). In the notion of "have what they need," there is an inherent grouping and/or individualizing of the digital divide. Do groups and individuals have the competency to effectively use the technology (Wu et al., 2018)? How do sociodemographic and socioeconomic indicators such as age, income, gender, education, and disability status affect an individual or group's place in the digital divide (Goodall & Ward, 2010; Jones et al., 2009; Rice & Haythornthwatte, 2006)?

Additionally, inherent to the digital divide, are educational justice (Corcoran et al., 2021; Xiao, 2021), equity (Gorski, 2005; Williams et al., 2021), inclusion (Long & Kowalske, 2022; Shaw, 2021), and social justice, which is a process as well as a goal (Xiao, 2021). The goal is to empower all groups in society to have full and equal involvement that is mutually tailored to fulfill their needs (Shin, 2020). The process promotes equity-mindedness by critically reflecting on how to overcome systemic injustices and pave the path to social justice (Leithwood et al., 2019). Addressing these issues is one of the answers to closing the digital divide (The National Association for Multicultural Education, 2020) During the Covid-19 pandemic, a spotlight was cast on the digital divide on many fronts. This paper addresses the effects of the digital divide in terms of access to distance learning by deaf and hard of hearing students attending virtual classes at a university.

All students had access to the internet. The university education was free for all students. Age, income, and gender do not appear to play a part in the digital divide in this study. Components of the digital divide that challenge deaf and hard-of-hearing students include the competency to use technology, equity, inclusion, and educational and social justice. The competency to use the technology is relevant to the professors and to the individuals in information technology who designed the platform for instruction. All students had the right to continue their education during the pandemic. Not all students can succeed with equity.

Literature focused on DHH students during the pandemic is limited. As evidence, in SciVal (Scopus), statistics show that only 15 studies were conducted between 2018 and 2022 at the time of writing this study addressing "deaf," "higher education," and "Covid-19" jointly; none of them was longitudinal (SciVal, 2022).

This study addresses the issue of appraising the educational experience and the academic achievement of DHH and hearing students through the lens of components of the digital divide that affected university students with hearing loss. It also, demonstrates the perceptions of students and lecturers when implementing the distance learning model of education during the Covid-19 pandemic by answering the following questions:

- Did Covid-19 affect academic achievement for DHH and hearing students equally over time?
  If yes,
- Did both groups adapt at the same rate during the pandemic?
- Did digital equality contribute to the perception of equity amongst both groups over time?

This study provides rich knowledge and information about technology-based learning by using longitudinal quantitative and qualitative approaches to enhance and deepen the understanding of different perspectives and experiences over an extended period and change.

### 1.1. Context of the study

The United Arab Emirates (UAE), also called the Emirates, is an ambitious young country located at the eastern end of the Arabian Peninsula. There is a priority to ensure Emirati citizens with disabilities have rights and are protected and afforded equal opportunities in all areas of life (UAE, 2021). According to the World Health Organization, approximately 11 percent of the UAE population has a disability (WHO, 2020). Federal Law protects Emirati citizens with disabilities (also called People of Determination) No. (29), 2006. They may not be discriminated against in any way. For example, the law provides for "equal care, rights, and opportunities in education, health care, training, and rehabilitation" (UAE, 2020). The UAE also guarantees, "the right to request, receive and communicate information on an equal footing with others" (UAE, 2020, sec. 7). The law was passed in 2006. Nevertheless, executing this 16-year-old law for students who are DHH remains a challenge, particularly in higher education.

Higher education institutions are considered integral to the UAE learning management system. DHH undergraduate students' bachelor's degree enrollment depends on various factors, including high school grades and performance on standardized tests. Students enter with a range of communication styles/preferences, academic backgrounds, cognition abilities, and digital literacy. These qualities often play a crucial role in their success and response to online education (Lynn et al., 2020). Sign language is unlike most other languages; it has no written component, and so DHH students have the extra challenge of learning the language without its written component to boost their learning process.

The courses offered for DHH students are limited to two specific programs. The programs take eight semesters plus an internship semester to be completed. The programs aspire to actively prepare special education teachers and socially qualified cadres acquainted with knowledge, science, and modern technology. DHH students in the UAE, depending on the severity of their hearing loss, learn written Arabic and written English. Their preferred mode of communication is a mixture of Arabic Sign Language with some features of the developing Emirati Sign Language and American Sign Language. In the university setting, they take classes with a sign language interpreter who is typically familiar with these signed languages.

A Moodle-based LMS was established in the university early in 2013 to be used by all students and all faculty. The university provided consistent training for new students and faculty every semester. However, the LMS was only used to support students during the semester before the pandemic. All classes were provided face-to-face with no exceptions. During the Covid-f19 pandemic, in the middle of the second semester of the academic year 2019/2020, UAE education institutions migrated from face-to-face to complete distance learning in response to instructions from the Ministry of Education. Relying on distance learning during Covid-19 was the only method that could reach the students in the UAE. The use of distance learning may have unexpectedly exacerbated the challenges of "technical infrastructure, competencies, and pedagogies for distance learning" (Marinoni et al., 2020, p. 2) facing students in the world, in general, and DHH students in particular.

## 2. Material and methods

### 2.1. Participants

The target population for this research consisted of all the students admitted before the pandemic in the academic year 2018/2019 in two programs at one of the UAE universities, where 38 DHH students and 44 hearing students registered in the same programs taking the same courses at the same pointed semesters.

### 2.2. Methods and data collection

This 12-month longitudinal study investigated the digital divide between two groups of students (DHH and hearing students) across three measures times. Data collection began at the beginning of the Covid-19 pandemic during the summer semester (June 2020) (T1). A short time after the learning process was "stable" and a new academic year started (January 2021) (T2), and a third time at the end of the academic year where the teaching remained online but the exams were on campus (June 2021) (T3). To represent a starting point reference, we chose December 2018 (T0) to track the changes in student performance during these times. Table 1 shows the details for each instrument used during each moment and highlights the main respondents for each instrument along with specific data at each time.

*Table 1.* Data collection instruments over time

| Instrument | Time | Respondents | Studied moments and respondents' numbers |
|---|---|---|---|
| GPA Marks | T0, T1, T2, T3 | DHH students | (38: 21 Male, 17 Female) (T0), (T1), (T2), (T3) |
| | | Hearing Students | (44: 19 Male, 25 Female) (T0), (T1), (T2), (T3) |
| Interviews | T1, T3 | DHH students | (15) (T1) |
| | | Hearing Students | (10) (T1) |
| | | Lecturers | 3 (T1), 3 (T3) |
| Questionnaires | T2, T3 | DHH students | 20 (T2), 31 (T3) |
| | | Hearing Students | 28 (T2), 38 (T3) |

Three main data collection activities were undertaken: the calculated grade point average (GPA) marks, questionnaires, and in-depth interviews with both students and lecturers. Student GPAs were extracted from the registration system for each moment. The interviews were designed to probe the emerging pandemic issues in-depth detail. Thereafter, the questionnaire was constructed to provide a comprehensive overview of students' perception of their experiences with distance learning, with a particular focus on pedagogy and student learning. Figure 1 below demonstrates the data collection process within the study.

## 3. Data analysis and results

The section started by analyzing the GPA marks of DHH and hearing over four semesters. This was followed by analyzing the designed interviews and questionnaires at three different moments in time for students and lecturers. SPSS V27 package was used to analyze quantitative data. To explore the divide between the two groups of students in the different moments, repeated-measures ANOVA, the fixed effect regression approach, and one-way ANOVA were used. A significance level of $p < .05$ was used throughout the study. Manual analysis was employed to analyze the qualitative data, with the data being classified and line coded about emerging themes.

### 3.1. GPA grades analysis

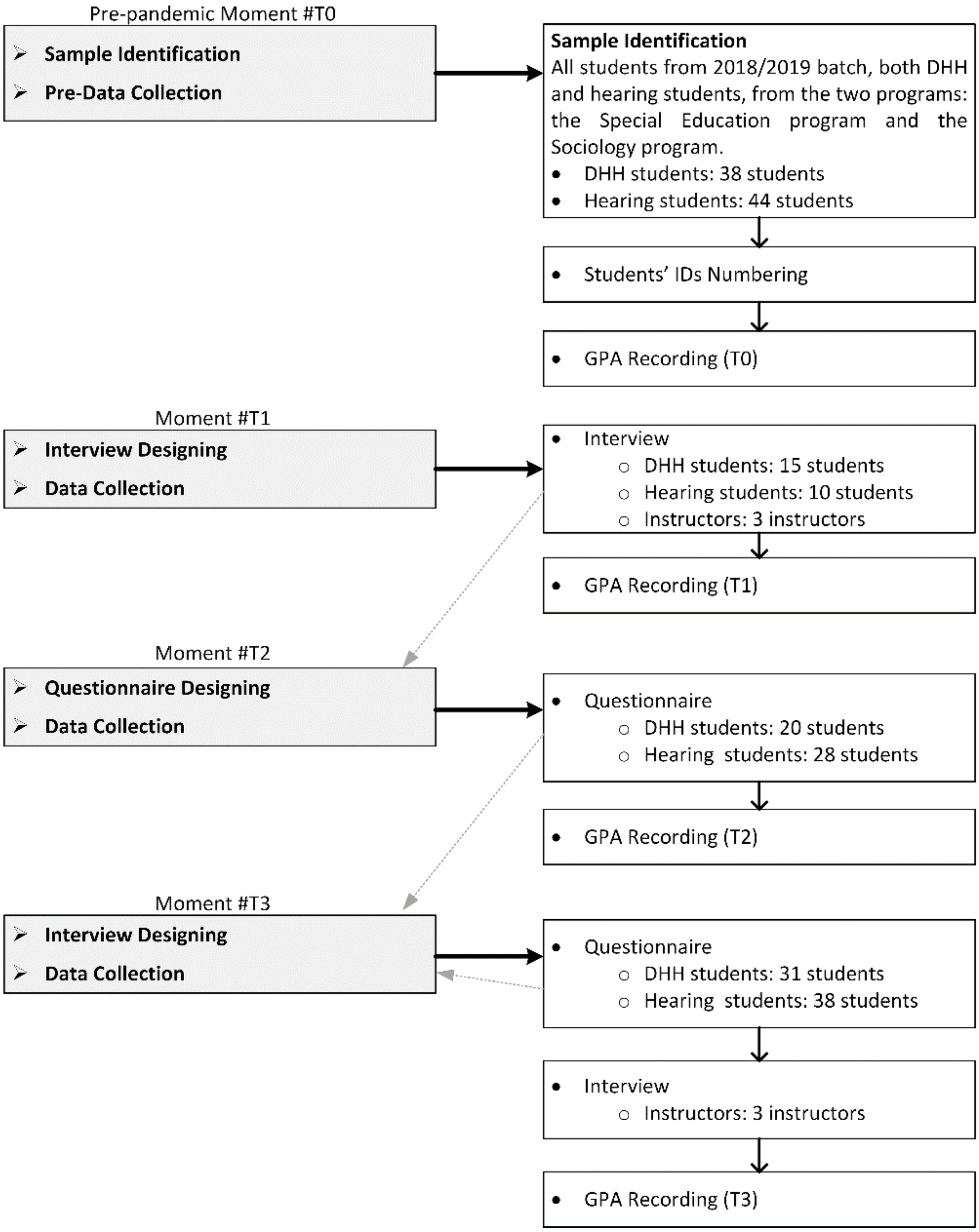
Table 2 shows the descriptive results for the GPA of DHH and hearing over four semesters. Results show decreasing in GPA means for DHH students. Conversely, the GPA scores for hearing students showed an increase over all time points.

One-way repeated measures of ANOVA showed that there were significant differences in the GPA means between the DHH students over time $F(4,320) = 10.706$, $p < .001$.The partial $\eta^2 = .118$ revealed a medium toward large effect. The Wilks' lambda was calculated to be.626 with an associated level of statistical significance of $p < .001$, leading to a conclusion that there was a strongly significant difference between the GPAs of DHH and hearing students over the time of the pandemic. There was a statistically significant difference in the academic performance between DHH and hearing students during the three semesters, $F(4,77) = 11.489$, $p < .001$.

*Table 2.* Descriptive statistics for the GPA over the different semesters

| | Hearing status | Mean | Standard deviation | *N* |
|---|---|---|---|---|
| T0 | Hearing | 2.07 | .79 | 44 |
| | DHH | 2.73 | .45 | 38 |
| T1 | Hearing | 2.18 | .80 | 44 |
| | DHH | 2.68 | .45 | 38 |
| T2 | Hearing | 2.38 | .74 | 44 |
| | DHH | 2.63 | .43 | 38 |
| T3 | Hearing | 2.56 | .48 | 44 |
| | DHH | 2.66 | .42 | 38 |

*Figure 1.* Data collection moments



The fixed effect regression approach was applied for DHH and hearing students separately. The R-square change indicated the increment in *R*-square as a result of adding in the time-varying predictors. It showed that adding in the time-varying predictors for the hearing students added an additional 6.1% of the total variation to students' GPA marks, with $F(4,172) = 10.350$, $p < .001$. The *F*-test indicated that the variation accounted for was significant, $F(47,172) = 1.914$, $p < .001$. While only 0.6% was added addition to the DHH students' GPA marks, with $F(4,148) = 2.734$, $p = .031$. The *F*-test indicated that the variation accounted for was significant, $F(41,148) = 40.981$, $p < .001$.

To investigate the direction of the above significance, hierarchical regression was used for hearing students, T2 ($b = 0.309$, $SE = 0.089$, $p = .001$) and T3 ($b = 0.486$, $SE = .089$, $p = < .001$). Both were positive and significant variables in the model. T1 emerged as positive but non-significant ($b = 0.111$, $SE = 0.089$, $p = .218$). In contrast, for DHH students, T2 ($b = -0.098$, $SE = .032$, $p = .003$) was a negative and significant variable in the model. T1 and T3 emerged as a negative but non-significant ($b = -0.044$, $SE = .032$, $p = .171$) (b = -0.061, $SE = .032$, $p = .062$), respectively variable in the model.

### 3.2. Interviews and questionnaires

The interviews for students were conducted in the first moment (T1) (as shown in Figure 1), while the questionnaires were consequently conducted in the second and third moments. Regarding the instructors, they were interviewed in the first and third moments. Table 3 shows the descriptive of the qualitative instruments used in this study along with their reliability results.

*Table 3*. A Description of the interviews and questionnaires respondent numbers

| Instrument | Respondents | T1 | T2 | T3 |
|---|---|---|---|---|
| Interview | DHH student | 15 (5 male,10 female) ($n$ = 15, 39.4%) | | |
| | Hearing students | 10 (4 male, 6 female) ($n$ = 10, 22.7%) | | |
| | Lecturers | 3 (3 male) ($n$ = 3, 60%) | | 3 (3 male) ($n$ = 3,60%) |
| Questionnaire | DHH students | | 20 (8 male, 12 female) ($n$ = 20, 52.6%) | 31(20 male, 11 female) ($n$ = 31,70.5%) |
| | Hearing students | | 28 (10 male, 18 female) ($n$ = 28,63.6%) | 38 (17 male, 21 female) ($n$ = 38,86.3%) |
| Questionnaire's Reliability | | | Cronbach's Alpha =.902 | Cronbach's Alpha =.916 |

#### *3.2.1. Reliability and validity*

Reliability and validity were considered carefully during the research process for both the interviews and questionnaires. First, for the interviews' reliability, internal homogeneity and external heterogeneity were used (Braun & Clarke, 2006). Several meetings took place to go through the analysis process. Any discrepancies found were resolved with specific data examples. The intercoder agreement through the analysis process reached 80%. Through several meetings of data analysis discussions, 100% agreement was reached, resulting in well-established themes (see section 3.2.2). For the validity purpose, to confirm and clarify the transcripts, transcripts were sent by email to all students and required written feedback.

Second, the reliability measure was tested using Cronbach's alpha for the questionnaires. According to the results of this test, Cronbach's Alpha =.902 and.916 for the first and second questionnaires, respectively. To validate the questionnaire, two steps were followed. First, two experts in technology and special education were asked to review the questionnaire to evaluate how well the questions effectively captured the objectives of this study. That step resulted in a few changes that were reflected in the questionnaire used in the second step. Thereafter, a pilot study was conducted before distributing the main questionnaire. A small sample of students from both groups ($n$ = 10) was contacted to go through the questionnaire which resulted in removing or editing some of the questions.

### *3.2.2. Interviews analysis*

This section shows the details of the thematic analysis of the students' interviews along with the results from the lecturers' interviews.

**Students' interviews**

The initial data collection began by designing and conducting an interview with DHH students, hearing students, and lecturers. Data were collected through a semi-structured interview format where the questions were prepared as a result of the related literature review; e.g., Alsadoon & Turkestani, 2020; McKeown & McKeown, 2019; Vos et al., 2015.

Below are the semi-structured interview questions for students:

- Describe the changes you faced when the University switched to distance learning.
- What changes helped you learn?
- What changes made it more difficult to learn?
- Describe what would be included in the best and most effective distance-learning environment for you.
- What quick changes could your professor/university make that would help you the most?
- For an online class to be perfect for you, what would you include and what would you eliminate?
- What else would you like us to know?

For the first moment (T1) in June 2020, of the 38 DHH students, five male and ten female students (*n* = 15, 39.4%) agreed to participate through in-person interviews on MS Teams. An Arabic Sign Language interpreter was present in the conference call to facilitate communication. Of the 44 hearing students, six female and four male students (*n* = 10, 22.7%) agreed to participate through in-person interviews on MS Teams. For ease of communication, the DHH students were interviewed at different times than the hearing students.

A content analysis was applied. A thematic analysis approach from Vaismoradi et al. (2013) was used. Two authors participated simultaneously in the analysis. The analysis resulted in five themes: course content, technology, delivery approach, assessment methods, and social interaction. In these themes, students described the challenges they faced during the Covid-19 pandemic.

**A. Course contents**

Both groups of students stated the number of slides per lecture was overwhelming. According to one of the hearing students, *"The material is too large and intimidating."* Students noted frustration with the text-based resources which were not suitable for the DHH students without support: *"We felt the need for different types of content like images and not just focusing on text. In addition to this, using texts only is not helpful for deaf students."* The DHH students raised another important issue *"the discussion approach was missing because of the large volume of content the instructor needs to deliver [sic]."*

**B. Delivery approach**

Both groups emphasized changes regarding online delivery. One hearing student said: *"Nothing changed in the teaching style, and yet everything had changed in our teaching environment.*" Moreover, DHH students complained about the online delivery approach using video conferencing. Compared with the physical classroom, the instructor's presence was missing on the screen. This is an important distinction and difference between the delivery mode for the hearing students and the DHH students. According to the DHH students, when video conferencing was used in course delivery, they only saw the interpreter. They could not see the instructor, which was a key communicative barrier for them. This is important to explain. Typically, when using a program like MS Teams, there are spaces for the faces of people. In this case, the screens were set to show only the interpreter. The lecturer was not able to be seen. One DHH student said, *"Indirect communication with our teacher has caused some lack of understanding, which negatively affected my learning."* Another student said, *"It is a big challenge. It is frustrating because we feel our questions are not addressed properly by the interpreter, and the time it takes to repeat the question again online makes it longer than usual, which is frustrating for us and stops us from asking any more questions."*

### C. Technology

The DHH students expressed their concerns about gaps in their IT skills. Although they had completed a course in computer skills in the previous semester, they felt overwhelmed. *"I was unsure about this new method of learning as I knew it requires certain IT skills which I don't have."* Having said this, some DHH students appreciated the opportunity to learn new technology. One DHH student stated, *"I am happy because we are now [sic] better able to use computers and internet than before."* Hearing students were more confident in using the new technology required. A common concern was the poor quality of learning opportunities if they did not have access to high-quality, smart devices.

### D. Assessment methods

Both groups shared similar concerns regarding assessments. Assessments did not change when classes migrated to distance learning. They still included assignments, quizzes, and exams. Hearing students reported problems handling the long quizzes and the assignments' instructions. One hearing student clarified, *"I am terrified by the long instructions associated with the assignments. We are not used to this."* A very challenging issue for the DHH students was the unavailability of direct communication with the lecturer, particularly during the exams. One DHH student explained, *"assignments given to us are too long and there are too many instructions we need to follow. Not being able to contact the instructor directly affected the quality of our answers."*

### E. Social interaction

The DHH students shared grave concerns. DHH students conveyed their sadness at not being able to socialize with each other; for instance, "*at the beginning of the pandemic, we as a group, asked the university to give us the opportunity to come to classes rather than using online distance learning, as coming to university is one of the few opportunities we get to socialize.*" Of course, their request was not permitted since the educational facilities were all shut down. Hearing students expressed similar difficulty in limited socialization but to a much lesser extent. One student stated, "*we missed our gatherings and our chats, but we used other means to communicate with our close colleagues*." DHH students reported that they have been lonely and isolated during COVID.

## Lecturers' interviews

### A. First moment (June 2020)

Lecturers' perceptions of online distance learning during Covid-19 were canvassed from three open questions asked during semi-structured interviews. They were asked to:

- Describe your preparation for teaching DHH students in the face-to-face model vs for online distance learning.
- Describe your perception of teaching DHH students in the face-to-face model vs for online distance learning.
- How do you think DHH students perceive their online distance learning experience vs hearing students?

Five lecturers were invited to participate in this study. Three lecturers agreed and were interviewed. In general, lecturers have rather negative experiences with teaching DHH students using distance learning. They reported that this was due to several challenges of using the new technology to align between students, interpreters, and course content. One instructor stated that: *"not only new technology but also a very huge challenge to manage between the content and the interpreter speed."*

At the same time, lecturers were positive about learning new techniques to help DHH students. One of the lecturers commented on the teaching strategies he used during the pandemic. "*During my teaching experience, this is my first time teaching deaf students. I found this experience rewarding, as I had to revise my methods of teaching to cope with their needs, which has helped me acquire new experience."*

Concerning the second question about their perception of teaching DHH students face-to-face compared with online distance learning, despite feeling unprepared, one of the instructors stated: *"I found teaching English as a foreign language to deaf students very difficult. I used images from the internet to link with words, but I cannot* *always easily find the correct ones. I taught English 1 to the same group before in class, and it was a challenge to teach them comprehension, but with online distance learning I found it even more difficult.*"

Finally, lecturers reported their concerns about their DHH students' abilities to cope with distance learning. They recognized the gaps in IT skills in the DHH students. According to one of the instructors, "*many times; during online classes; at least two or three students asked for IT help as they can't perform some activities requested from them*." The quality of their learning, according to the lecturers; was affected negatively.

### B. Third moment (June 2021)

To provide a contextualized view of the issues that arose from the first and second questionnaires with students, another interview took place with lecturers who taught DHH and hearing students (at least two times between T1 and T3). The main questions were:

- What is the difference between the first and second-time experience of teaching DHH and hearing students using online distance learning tools during Covid-19; and
- What are their perceptions of the effect of online learning on lecture attendance and communication patterns of DHH students in comparison with their peers of hearing students?

The three lecturer participants described more confidence in teaching students (DHH and hearing) compared with the first time they used the system. According to one of the participants, *"compared with the first time, I have become more confident in using technology in teaching online classes, especially for hearing students*." Another instructor indicated his happiness in learning new techniques as a result of the feedback the authors shared with them from the first round of this study. According to the lecturer, the DHH students improved on their second learning experience with the lecturer compared with the first time. Additionally, two out of three lecturers indicated that fewer students reported difficulty using the technology the first time. It seems that their motivation and confidence in using digital tools have enhanced and empowered them to adapt to the current environment.

Although there were some bright spots in instruction, some lecturers relayed their concerns that hearing students were not attending lectures regularly compared with the DHH students. When lecturers talked to students about their lack of engagement, students said they felt the online attendance was optional compared to the face-to-face classes where students had to attend physically. Moreover, even with DHH students, lecturers were concerned that the quality of the student's learning could be reduced due to missing direct face-to-face communication with the instructor, and the opportunity to work as a group with their peers in class.

To conclude, two of the three lecturers agreed that online distance learning was more beneficial to hearing students compared to the DHH students who showed their lack of confidence in learning using online distance learning tools. According to one of the instructors, *"there was a mismatch between deaf and hearing students in their experience to achieving better results and motivation and support to learn*."

### 3.2.3. The questionnaire analysis

The questionnaire for this study was developed based on the first interview's identified dimensions, which led to the development of the questionnaire which was based on the literature (Baber, 2022; Jacobson & Mackey, 2013; Resta & Laferrière, 2008; Woo et al., 2016). The questionnaire contained five themes and 13 related criteria as shown in Table 4. Respondents indicated their level of agreement with each of the 63 statements on a Likert scale from 1 to 7. Another three questions were added to the third moment's questionnaire to identify the students' general experience perception of using online distance learning during the pandemic.

*Table 4.* The factors and related criteria of the questionnaire

| Themes | Criteria |
|---|---|
| Course content | Content quality |
| | Online material usage |
| Technology | Accessibility |
| | System quality |
| | Computer playfulness |
| | Computer self-efficacy |
| | Perceived ease of use |

| | |
|---|---|
| Delivery approach | Effectiveness |
| | Usability |
| Assessment methods | Exams |
| | Assignments |
| Social interaction | Lack of social |
| | Perception of social distance |

**A. DHH vs hearing students in the second moment (January 2021)**

For the second moment (T2), in January of 2021, the 38 DHH students and 44 hearing undergraduate students were sent an email invitation, with a short description of the study, information about confidentiality, and a link to the questionnaire. Of the 82 respondents, 20 DHH students and 28 hearing students responded to the questions of the questionnaire.

The 20 DHH participants had average perceptions of content, technology, assessment, social interaction, and delivery approach of 4.64 (*SD* = 1.01), 4.78 (*SD* =.83), 4.89 (*SD* = 1.08), 5.12 (*SD* =.94), 4.79 (*SD* =.87), respectively. The 28 hearing students had average perceptions of content, technology, assessment, social interaction, and delivery approach of 6.01 (*SD* = .85), 6.34 (*SD* = .85), 6.17 (*SD* = .96), 6.09 (*SD* = .86), 5.78 (*SD* = 1.05), respectively. The general experience had an average of 5.45 (*SD* =.87) for DHH students and 6.45 (*SD* = 1.12) for hearing students.

The results for the one-way ANOVA test show that the effect of the health condition of students, therefore, was significant, $p \leq .001$, in every area (content, technology, assessment, social interaction, and delivery approach). ANOVA results show that was a significant difference between the perspectives of the two groups of students, $F(1,46) = 11.073$, $p = .002$.

**B. DHH students vs hearing students in the third moment (June 2021)**

For the third moment (T3), in June of 2021, the students were sent an email invitation, with a short description of the study, information about confidentiality, and a link to the questionnaire. Of the 82 respondents, 31 DHH students and 38 hearing students responded to the questions of the questionnaire.

The results for the one-way ANOVA test between the two groups of students in the second moment showed that the 31 participants the DHH students had average perceptions of content, technology, assessment, social interaction, and delivery approach of 4.21 (*SD* = .74), 4.51 (*SD* = .88), 4.74 (*SD* = 1.19), 4.80 (*SD* = 1.09), 4.10 (*SD* = .73), respectively. The 38 hearing students had average perceptions of content, technology, assessment, social interaction, and delivery approach of 5.92 (*SD* = 1.08), 6.11 (*SD* = 1.13), 5.90 (*SD* = 1.44), 6.04 (*SD* = 1.13), and 5.41 (*SD* = 1.51), respectively. The general experience had an average of 4.60 (*SD* = .94) for DHH students and 5.66 (*SD* = 1.98) for hearing students.

The effect of the hearing status of students was significant, $p \leq .001$, in every area (content, technology, assessment, social interaction, and delivery approach). The ANOVA results show that was a significant difference between the perspectives of the two groups of students, $F(1,67) = 7.477$, $p = .008$.

**C. Comparison of questionnaire analysis between the two moments for each group**

One-way ANOVA results show that there was a significant difference between the perspectives of the DHH students only in the delivery approach dimension perception; T2: Mean = 4.79 (*SD* = .87), T3: Mean = 4.10 (*SD* = .73), $F(1,49) = 9.112$, $p$ =. 004. However, one-way ANOVA results show that there was no significant difference between the perspectives of the hearing in any of their perceptions.

It is interesting to note that DHH students were more positive in the second moment than in the third moment despite the accumulated knowledge they were supposed to have during the time between the two questionnaires. Generally speaking, DHH students were less satisfied than the hearing students and they also were less satisfied with time.

## 4. Discussion

This one-year longitudinal study investigated the experiences of DHH and hearing students and three of their lecturers regarding distance learning during the Covid-19 lockdowns. Exploration of the digital divide in higher education is an important predictor to students' success (Azubuike et al., 2021; Sabeghi et al., 2021; Song et al., 2021). Overall, the results in this study indicate differences between DHH and hearing students in regard to the changes in their GPAs and perceptions of the distance learning experience. The following points of discussion will go through the three main questions of this study one by one.

### 4.1. Research Question 1: Did COVID-19 affect the academic achievement of DHH students and hearing students equally over time?

To answer this question, the GPA analysis over approximately a year was conducted to show if both groups were equally affected. Results showed that academic achievement was affected for both groups; however, the effect was not in the same direction nor were they at the same level. The academic achievement of the DHH students was significantly better than the hearing students before the pandemic. This contradicts the literature where Shaw (2021) suggested that "disabled students … tend to have worse post-degree outcomes presents a social justice issue" (p. 13). In the current study case, DHH students tend to have better results and outcomes regarding their GPAs. DHH students did not have concerns regarding their health conditions, nor did they report facing any negative stigma within their on-campus learning environment. However, results show a significant variation in the means of the students' GPAs per semester considering both groups of students. The GPAs of the hearing students increased over time while the GPAs of the DHH students dropped for the first two moments but showed a rebound by the third moment. This phenomenon could be a result of student development of familiarity and confidence with the platform and instructional model. These results highlight a gap in performance between the two categories of students. DHH students' learning seemed to be more negatively affected by the pandemic; their GPAs decreased significantly in T1 and decreased even further in T2, with a total variation of 0.6%. However, the GPA for the hearing students increased significantly during the interval of Covid-19 with a total positive variation of 6.1%, implying that they had advantages throughout distance learning. These results are not surprising, considering the challenges and confusion of distance learning amongst DHH students during this period. This difference was picked up previously (Lambert & Czerniewicz, 2020) for DHH and hearing students with different social backgrounds, GPAs, course grades, and genders. Recognizing that digital inequality can reinforce social inequities, it is necessary to investigate how such imbalances perpetuate the inability to compete academically (Moldavan et al., 2021).

### 4.2. Research Question 2: Did both groups adapt at the same rate during the pandemic?

The simple answer is no, they did not adapt at the same rate. However, none of the students failed. The DHH students began the experience of distance learning feeling unsure about the technology. The inability to see the lecturer was reported as a major barrier. Although the hearing students showed higher academic achievement and more positive perceptions toward distance learning, in the third moment, they considered online instruction to be optional. It was in the third moment that the DHH students' GPAs rose but insignificantly. Studies using DHH students as participants revealed practice with technology and an increase in experience support their academic successes (e.g., McKeown & McKeown, 2019). Studies have found links between students' attitudes and their general adaptation in, and outside of the classroom (Wang & Holcombe, 2010; Wu et al., 2018). It is possible that part of the difference in adaptation between the two groups is associated with how students viewed themselves and their universities, which may be reflected in their social environment and the time they invest in their studies. This may be derived from the sense of self-confidence and engagement in constructive academic relationships (Høigaard et al., 2015).

The DHH students formally requested in-person classes even when the university shuttered. By nature, students using sign language are limited in easy communication by the number of people who can effectively sign with them. DHH students who use sign language are already at risk for loneliness, isolation, and challenges in mental health (Bott & Saunders, 2021; Tigwell et al., 2020). In addition, studies have found links between students' attitudes and their general adaptation in, and outside of the classroom (Wang & Holcombe, 2010; Wu et al., 2018). DHH students' GPA could be a proxy measure for their academic achievement. Their initial experience was associated with underachievement and low adaptation during distance learning intervals especially the first and second-time points with a significant reduction in GPA. Thus, early attempts by students to adapt could have affected their experiences and attitudes, as well as academic achievement. This result is in line with the contention that early life experiences have a significant impact on university students' future development and early perceptions of students toward the campus climate and university examination performance are associated with later GPA (Sakız et al., 2021).

It seems that DHH students and their lecturers reported the lack of DHH students' technological abilities, missing face-to-face communication, and the challenge of social distancing. This is considered a major barrier within their physical environment. The difference in adaptation while using online learning contradicts Lambert and Czerniewicz (2020) who found using online courses succeeded to enhance the educational experience and equity. However, the current study echoed the findings of Tigwell et al. (2020), which showed that the Covid-19 pandemic presented physical, social, and emotional challenges for students with disabilities. In conclusion, referring to the social model of disabilities, limitations were imposed on DHH students affecting their adaptability during the pandemic.

### 4.3. Research Question 3: Did digital equality contribute to the perception of equity amongst both groups over time?

The Covid-19 pandemic exposed the digital divide between the two groups of students. There are longstanding digital inequities that may have exacerbated other inequities, especially those related to disabilities as highlighted by (Gao & Hayes, 2021). In this study, it has been shown that the perception of DHH students declined over time while the hearing students had no significant changes. The impact of the pandemic has been multiplied by online education, highlighting even more the negative impact of a digital divide in higher education. As stated by Sabeghi (2021, p. 2): "online education has created a gap in the gap, worsening the negative impacts." The reality is that the platforms designed for distance learning were not designed for easy access for students who relied on sign language, particularly for synchronous courses. Since the environment in this case study did not consider the concept of diversity, according to the social model of disability the environment imposed limitations on the DHH students due to the physical and social barriers; therefore, there is a high possibility that students would not get what they need and they might face many barriers that could affect their study.

Another major dimension is the lecturers involved in teaching the students. Teaching at the postsecondary level is challenging and the availability and value of lecturers should not be underestimated. According to Diez et al. (2015), lecturers play a pivotal role in student learning which can be supported or hindered by them. When postsecondary institutions migrated to distance learning, it occurred quickly and without the requisite training. Many lecturers reported feeling overwhelmed and underprepared to teach remotely (DeMartino & Weiser, 2021). The quick move out of face-to-face instruction resulted in no time to learn those strategies which was the source of challenges for lecturers (Marinoni et al., 2020). Researching teaching strategies particularly effective for students enrolled in courses that are based on a distance learning model is a recommendation. This can, perhaps, reduce the digital divide.

In addition to planning for distance learning, lecturers can greatly improve lessons through planning by using the Universal Design for Learning (UDL) framework. Based on more than 1000 research studies on effective practice (CAST, 2022b), all strategies begin with access and equity. Briefly, UDL is comprised of three pillars. Using UDL, lecturers can ensure planning designed to emphasize equity through designing lessons that "provide multiple means of engagement; provide multiple means of representation; [and] provide multiple means of action and expression" (CAST, 2022a, para. 1). Social justice is based on equity, access, participation, and harmony (Crethar & Winterowd, 2012). The lessons planned using UDL are equity-based, research-based, and have a focus on social justice.

## 5. Conclusion and implications

This study was conducted over three points of time during the Covid-19 pandemic to examine the educational experiences of Deaf and Hard of Hearing (DHH) students in relation to their hearing student peers through a social justice lens for more than a one-time slot. These semesters represent the beginning, middle, and end of the student's experience during the pandemic, with the early year before the pandemic serving as a reference point (the "Zero time point"). The study found that there was a significant difference between the two groups in terms of educational outcomes, even though they were using the same program design, course design, delivery mode, assessment tools, and technology. The results of the study suggest that the priority for DHH students is not inclusion and equality per se, but rather a more equitable learning environment.

These findings have implications for postsecondary education, as they suggest that DHH students are at risk of underachieving in this setting. Higher education institutions, universities, and curriculum designers should therefore work to improve distance learning platforms so that students with hearing loss can have equitable access to lecturers, course content, and their peers. Universities should also develop strategies to reduce inequities and adapt their educational systems to meet the needs of, more broadly, all students with disabilities. To achieve this, it will be necessary to engage in a continuous process of collaboration between DHH students, deaf parents, deaf teachers, sign language interpreters, universities, and Higher Education Institutions to construct equitable learning environments. Equity and justice in the learning environment are the targets to manage the knowledge gap and prepare DHH students for their future careers.

The Covid-19 pandemic presented a unique opportunity to consider how the needs of postsecondary students who were used to face-to-face classes could be met through distance learning. DHH students who use sign language were particularly affected. Based on both the quantitative and qualitative data, the following recommendations are made:

- Provide specific instruction to lecturers on the very specific needs of DHH students in the classroom, accommodations and how to prepare for instruction (Alsadoon & Turkestani, 2020; McKeown & McKeown, 2019). This activity, when conducted properly, can address the challenges noted in the abundance of course content, the overwhelming assessment process, and the frustrations with the delivery of the assessments.
- Consult with the DHH students, the lecturers, and the IT department to ensure students can see the interpreter and the lecturer on the screen. The collaboration of some student representatives and the IT department can help to eliminate frustrations related to the technology.
- Incorporate Universal Design for Learning. Planning lessons using this platform can increase equity to achieve a higher level of social justice.

## References


Alsadoon, E., & Turkestani, M. (2020). Virtual classrooms for hearing-impaired students during the COVID-19 pandemic. *Romanian Journal for Multidimensional Education, 12*, 1–8.

Azubuike, O. B., Adegboye, O., & Quadri, H. (2021). Who gets to learn in a pandemic? Exploring the digital divide in remote learning during the COVID-19 pandemic in Nigeria. *International Journal of Educational Research Open, 2*, 100022. https://doi.org/10.1016/j.ijedro.2020.100022

Baber, H. (2022). Social interaction and effectiveness of the online learning – A Moderating role of maintaining social distance during the pandemic COVID-19. *Asian Education and Development Studies, 11*(1), 159–171. https://doi.org/10.1108/AEDS-09-2020-0209/FULL/XML

Bott, A., & Saunders, G. (2021). A Scoping review of studies investigating hearing loss, social isolation and/or loneliness in adults. *International Journal of Audiology, 60*(S2), 30–46. https://doi.org/10.1080/14992027.2021.1915506/SUPPL_FILE/IIJA_A_1915506_SM3242.PDF

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

CAST. (2022a). *UDL guidelines*. https://udlguidelines.cast.org

CAST. (2022b). *Universal design for learning.* https://udlguidelines.cast.org/?utm_source=castsite&lutm_medium=web&utm_campaign=none&utm_content=aboutudl

Corcoran, T., Whitburn, B., & Knight, E. (2021). Inherent requirements in higher education: Locating You in Us. *Perspectives: Policy and Practice in Higher Education, 26*(2), 69–75. https://doi.org/10.1080/13603108.2021.1986166

Crethar, H. C., & Winterowd, C. L. (2012). Values and social justice in counseling. *Counseling and Values, 57*(1), 3–9. https://doi.org/10.1002/J.2161-007X.2012.00001.X

DeMartino, L., & Weiser, G. S. (2021). Striving for equity in pandemic times: The Administrator's role in the shift to online education in K-12 and higher education spaces. In *Handbook of Research on Inequities in Online Education During Global Crises* (pp. 199–223). https://doi.org/10.4018/978-1-7998-6533-9.ch011

Dijk, J. Van. (2005). *The deepening divide: Inequality in the information society*. Sage Publication.

Gao, N., & Hayes, J. (2021). The digital divide in education. *Public Policy Institute of California*. https://www.ppic.org/publication/the-digital-divide-in-education/.

Goodall, K. T., & Ward, P. R. (2010). Older migrants finding what they need to know for everyday living Lareen Newman Tapping into Serenity. *Quality in Primary Care, 18*, 27–32. https://www.researchgate.net/publication/276921318

Gorski, P. (2005). Education equity and the digital divide. *AACE Journal, 13*(1), 3–45.

Høigaard, R., Kovač, V. B. Ø., Cecilie, N., & Haugen, T. (2015). Academic self-efficacy mediates the effects of school psychological climate on academic achievement. *School Psychology Quarterly, 30*(1), 64–74.

Jacobson, T. E., & Mackey, T. P. (2013). Proposing a metaliteracy model to redefine information literacy. *Communications in Information Literacy, 7*(2), 84–91.

Jones, S., Johnson-Yale, C., Millermaier, S., & Pérez, F. S. (2009). U.S. College students' internet use: Race, gender and digital divides. *Journal of Computer-Mediated Communication, 14*(2), 244–264. https://doi.org/10.1111/J.1083-6101.2009.01439.X

Lambert, S., & Czerniewicz, L. (2020). Approaches to open education and social justice research. *Journal of Interactive Media in Education, 2020*(1), 1–6. https://doi.org/10.5334/jime.584

Leithwood, K., Harris, A., & Hopkins, D. (2019). Seven strong claims about successful school leadership revisited. Https://Doi.Org/10.1080/13632434.2019.1596077, 40(1), 5–22. https://doi.org/10.1080/13632434.2019.1596077

Long, M. R., & Kowalske, G. M. (2022). Understanding STEM Instructors' Experiences with and Perceptions of Deaf and Hard-of-Hearing Students: The First Step toward Increasing Access and Inclusivity. *Journal of Chemical Education, 9*9(1), 274–282. https://doi.org/10.1021/acs.jchemed.1c00409

Lynn, M. A., Templeton, D. C., Ross, A. D., Gehret, A. U., Bida, M., Sanger, T. J., & Pagano, T. (2020). Successes and challenges in teaching chemistry to deaf and hard-of-hearing students in the time of COVID-19. *Journal of Chemical Education, 97*(9), 3322–3326. https://doi.org/10.1021/acs.jchemed.0c00602

Marinoni, G., Land, H. V., & Jensen, T. (2020). The Impact of Covid-19 on higher education around the world. *IAU Global Survey Report*, 23. https://www.uniss.it/sites/default/files/news/iau_covid19_and_he_survey_report_final_may_2020.pdf

McKeown, C., & McKeown, J. (2019). Accessibility in online courses: Understanding the deaf learner. *TechTrends*, *63*(5), 506–513. https://doi.org/10.1007/s11528-019-00385-3

Moldavan, A. M., Capraro, R. M., & Capraro, M. M. (2021). Navigating (and disrupting) the digital divide: Urban teachers' perspectives on secondary mathematics instruction during COVID-19. *Urban Review*, *54*, 277–302. https://doi.org/10.1007/s11256-021-00611-4

Díez, A., Gavira, R., & Molina, V. M. (2015). Students with disabilities in higher education: A Biographical-narrative approach to the role of lecturers. *Higher Education Research & Development, 34(*1), 147–159. https://doi.org/10.1080/07294360.2014.934329

The National Association for Multicultural Education (NAME). (2020). *Definitions of multicultural education - national association for multicultural education*. https://www.nameorg.org/definitions_of_multicultural_e.php

Resta, P., & Laferrière, T. (2008). Issues and challenges related to digital equity. In *International Handbook of Information Technology in Primary and Secondary Education* (pp. 765–778). https://doi.org/10.1007/978-0-387-73315-9_44

Rice, R. E., & Haythornthwatte, C. (2006). Perspectives on internet use: Access, involvement and interaction. In *Handbook of New Media* (pp. 92–113). https://www.hoepli.it/editore/hoepli_file/download_pub/978-88-203-3344-7_Chapter4.pdf

Sabeghi, H., Rezvani, M., Bahrami, M., & Kavi, E. (2021). Exposure of medical education system to digital divide due to COVID-19. *Journal of Medical Education, 20*(2), 8–10. https://doi.org/10.5812/jme.114569

Sakız, H., Özdaş, F., Göksu, İ., & Ekinci, A. (2021). A Longitudinal analysis of academic achievement and its correlates in higher education. *SAGE Open, 11*(1). https://doi.org/10.1177/21582440211003085

SciVal. (2022). *SciVal - Home*. https://www.scival.com/home?dgcid=ScopusHeader

Shaw, A. (2021). Inclusion of disabled Higher Education students: Why are we not there yet? *International Journal of Inclusive Education*. https://doi.org/10.1080/13603116.2021.1968514

Shin, H. (2020). Technological innovation in public education in the era of COVID-19: Focusing on distance education policy in South Korea. *Asian Journal of Innovation and Policy, 9*(2), 207–222. https://doi.org/10.7545/ajip.2020.9.2.207

Song, S. J., Tan, K. H., & Awang, M. M. (2021). Generic digital equity model in education: Mobile-assisted personalized learning (MAPL) through e-modules. *Sustainability (Switzerland), 13*(19). https://doi.org/10.3390/su131911115

Tigwell, G. W., Peiris, R. L., Watson, S., Garavuso, G. M., & Miller, H. (2020). Student and teacher perspectives of learning ASL in an online setting. *Proceedings of ASSETS 2020 - 22nd International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 1–7). https://doi.org/10.1145/3373625.3417298

United Arab Emirates Government (UAE). (2020). *Federal Law (No.) 29 of 2006: Concerning the rights of people of determination*. https://www.tamm.abudhabi/en/aspects-of-life/peopleofdetermination/peopleofdeterminationrights/LawsandRegulations/federal-law-no-29-of-2006-concerning-the-rights-of-people-of-determination

United Arab Emirates Government (UAE). (2021). *Protecting the rights of people of determination (persons with disabilities) - The Official Portal of the UAE Government*. https://u.ae/en/about-the-uae/human-rights-in-the-uae/protecting-the-rights-of-people-of-determination

Vaismoradi, M., Turunen, H., & Bondas, T. (2013). Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nursing & Health Sciences, 15*(3), 398–405. https://doi.org/10.1111/nhs.12048

Vos, T., Barber, R. M., Bell, B., Bertozzi-Villa, A., Biryukov, S., Bolliger, I., Charlson, F., Davis, A., Degenhardt, L., Dicker, D., Duan, L., Erskine, H., Feigin, V. L., Ferrari, A. J., Fitzmaurice, C., Fleming, T., Graetz, N., Guinovart, C., Haagsma, J., … Murray, C. J. L. (2015). Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: A Systematic analysis for the Global Burden of Disease Study 2013. *The Lancet, 386*(9995), 743–800. https://doi.org/10.1016/S0140-6736(15)60692-4

Wang, M. Te, & Holcombe, R. (2010). Adolescents' perceptions of school environment, engagement, and academic achievement in middle school. *American Educational Research Journal, 47*(3), 633–662. https://doi.org/10.3102/0002831209361209

World Health Organization (WHO). (2020). *Listings of WHO's response to COVID-19*. https://www.who.int/news/item/29-06-2020-covidtimeline

Williams, T. K., McIntosh, R. W., & Russell, W. B. (2021). Equity in distance education during COVID-19. *Research in Social Sciences and Technology, 6*(1), 1–24. https://doi.org/10.46303/ressat.2021.1

Woo, K., Gosper, M., McNeill, M., Preston, G., Green, D., & Phillips, R. (2016). Web-based lecture technologies: Blurring the boundaries between face-to-face and distance learning. *Research in Learning Technology, 16*(2), 81–93. https://doi.org/10.1080/09687760802315895

Wu, T. F., Chen, C. M., Lo, H. S., Yeh, Y. M., & Chen, M. C. (2018). Factors related to ICT competencies for students with learning disabilities. *Educational Technology & Society, 21*(4), 76–88.

Xiao, J. (2021). Decoding new normal in education for the post-COVID-19 world: Beyond the digital solution. *Asian Journal of Distance Education, 16*(1), 2021. http://asianjde.com/ojs/index.php/AsianJDE/article/view/558/341

Zhao, L., Liu, Y., & Su, Y.-S. (2023). Personality Traits' Prediction of the Digital Skills Divide between Urban and Rural College Students: A Longitudinal and Cross-Sectional Analysis of Online Learning During the COVID-19 Pandemic. *Educational Technology & Society*, *26*(4), 150-162. https://doi.org/10.30191/ETS.202310_26(4).0011

# Personality Traits' Prediction of the Digital Skills Divide between Urban and Rural College Students: A Longitudinal and Cross-Sectional Analysis of Online Learning During the COVID-19 Pandemic

**Li Zhao[1], Yue Liu[1] and Yu-Sheng Su[2,3*]**
[1]School of Education Science, Nanjing Normal University, Nanjing, China // [2]Department of Computer Science and Information Engineering, National Chung Cheng University, Taiwan // [3]Department of Computer Science and Engineering, National Taiwan Ocean University, Taiwan // li.zhao@njnu.edu.cn // 200602149@njnu.edu.cn // ccucssu@gmail.com // ntouaddisonsu@gmail.com
[*]Corresponding author

**ABSTRACT:** The digital skills divide has been raised as a serious issue during the COVID-19. However, few studies explored the predictive influence of personality traits on college students' digital skills in online learning. To address this gap, this study took the second-level digital divide as the focus to conduct a two-round survey of college students over nearly 2 years based on cross-sectional and longitudinal methods to explore whether there is a digital skill divide or personality trait differences between urban and rural college students while learning online, and whether college students' personality traits can predict their digital skills. The results confirmed the rural and urban college students' digital skills divide. There were significant differences in all dimensions of their digital skills except for mobile skills. In addition, the digital skills divide of these college students persisted for nearly 2 years. Specifically, this study further confirmed that there were significant differences in the urban and rural college students' extraversion, neuroticism, and agreeableness, but there was no significant difference in their openness and conscientiousness. Additionally, there was no change in the urban and rural college students' personality trait differences in nearly 2 years. Personality traits could positively predict college students' digital skills. This study provides evidence for bridging the second-level digital divide of the rural and urban college students from the perspective of personality traits.

**Keywords:** Digital divide, College students, Cross-sectional analysis, Longitudinal analysis, Personality traits, Digital skills

## 1. Introduction

During the COVID-19, online learning has been globally adopted as an effective response measure to ensure orderly and uninterrupted teaching (Dhawan, 2020). However, a series of educational inequalities have also been found in the large-scale online learning practices in the past 2 years. For example, the online learning inequality between students in rural and urban areas has become more pronounced (Zhao et al., 2022). The inequality of online learning includes the inequality of Internet access and ICT technology use (Nguyen et al., 2021). The former can be defined as the first-level digital divide, while the latter is the second-level digital divide (Riggins & Dewan, 2005). Access to the Internet has been widely realized in many universities (Anderson, 2001). However, digital skills are needed for students to bridge the second-level digital divide (Cotten & Jelenewicz, 2006). By using ICT, rural students could reduce network homophily, increase the number of weak ties to others, and narrow the gap between their digital skills and those of urban students (Robinson et al., 2015). Thus, this study aimed to explain the reasoning of rural and urban college students' digital skills divide by exploring the determinants of digital skills.

Some scholars have proposed that digital divide as an issue of society deserves urgent attention during the period of COVID-19 and post-COVID-19 era (Azubuike et al., 2021; Dhawan, 2020). Prior studies have stated that social and cultural capital leads to the appearance of first-level digital divide (Azubuike et al., 2021; Ren et al., 2022). As for the exploration in factors of the second-level digital divide, the research was limited to objective personal factors, while there have been few studies on internal factors of individuals. For example, through a comparative study of teenagers in 39 countries, Drabowic (2014) concluded that boys use the Internet more frequently than girls for the purpose of education. Liao et al. (2016) also indicated that rural and urban students' parental educational levels had a great impact on their digital inequality. While findings such as those above contribute to a long-term understanding of the causes of differences in digital technology use (the second-level digital divide), they do not provide more practical measures to address inequalities in education. Specifically, demographic variables such as age, gender, and socioeconomic factors (e.g., family income or education level) cannot change or eliminate differences in a short period of time. Therefore, empirical studies need to determine the effect of individual psychological factors influencing the use of digital skills by urban and rural college students in online learning.

Based on the studies on social psychology and digital skills, individual personality traits will affect their technology use behavior and intentions (Digman, 1990; Tang et al., 2016). However, there are still few studies in this area, and especially, lack of empirical studies on the influence of personality traits on ICT technology use in the context of large-scale online learning. To bridge the above gap, several research questions were proposed: (1) Does the second-level digital divide exist for rural and urban college students during the process of COVID-19 online learning? (2) Additionally, are there differences in the personal traits of rural and urban college students? (3) Further, do personality traits contribute to college students' digital skills?

## 2. Literature review

### 2.1. The digital skills divide found between urban and rural college students

The gap existing in the information-rich and the information-poor can be defined as digital divide (Tien & Fu, 2006). The concept of digital divide is multidimensional and complicated and has been divided in three levels (Chipeva et al., 2018). The first-level digital divide refers to the digital access divide (Friedman, 2001), representing inequality in terms of acquirement to information technology (IT). As digital technologies become readily available to us, the situation of inequal digital access is gradually weakening. Then, second-level digital divide was proposed which referes to inequality in the ability to use information technology caused by the inequal access to digital equipment and some background elements (Riggins & Dewan, 2005). The biggest difference between the first and second-level digital divide is that the latter goes beyond access and focuses more on users' psychological, social, and cultural backgrounds (van Dijk, 2006). Subsequently, the digital divide was extended to the third level in Wei's et al. (2011) study, which refers to the unequal outcomings of using IT and investments.

During the period of COVID-19, inequalities in digital education have been well documented based on urban-rural and socioeconomic differences (Nerse, 2020). Although educational coherence was guaranteed to some extent because of the development of online learning during COVID-19, there remain unequal learning outcomes (Zhao et al., 2022). Thus, this study focused on the digital skill divided, considering that the first-level digital divide will gradually close due to the popularity of Internet equipment (Hargittai & Hsieh, 2012). As the symbol of the second-level digital divide, digital skills are believed to be survival skills and assets (Eshet, 2004; van Deursen & van Dijk, 2011), and can have an important impact on the effect of online learning (Cabello-Hutt et al., 2018; Livingstone & Helsper, 2010). Although the significance of digital skills has been emphasized again and again theoretically, it is practically often overlooked (Mabila et al., 2013). Therefore, the current study explored college students' digital skills and paid attention to their dynamic increase and decrease with the development of online learning.

When students enter university, the level of their digital skills varies from person to person (Correa, 2010; DiMaggio et al., 2004), and this difference is particularly prominent between urban and rural students (Stern et al., 2009). Li et al. (2022) suggested that the regional inequality in distance higher education is enlarged by online tertiary education to a greater extent in China. Welser et al. (2019) believed that online learning communities can expand teaching opportunities, improve digital skills, and offset the inequality caused by this inherent attribute. However, Wang (2020) argued that with the large-scale application of online learning and thus the change in teaching environment, urban and rural students would show digital differences owing to different physical and mental status. Some studies have shown that determinants including college students' sociodemographic and socioeconomic status have largely contributed to the digital divide (Scheerder et al., 2017). Goode (2010) mentioned that low-income students were those most underprepared for the digital college environment. Azionya and Nhedzi (2021) found that students from rural or low-income families faced more serious digital inequality in COVID-19. In view of the disputes between the two sides, the exploration of the differences and changes in the digital skills of urban and rural college students against the background of online learning was necessary.

### 2.2. The difference between rural and urban college students' personality traits and the possible effect on their digital skills

Walczuch and Lundgren (2004) defined personality traits as an individual's unique internal characteristics. At present, many psychologists accept Digman's (1990) classification. The domain of personality comprises five main constructs (Digman, 1990), known as the Big Five, dividing the personality traits into respectively agreeableness, extraversion, conscientiousness, neuroticism, and openness. Studies on social psychology have found that personality traits can determine one's beliefs and behavior in various aspects of one's life (Digman, 1990), such as the socio-cultural, family, social, and geographical aspects (Ebru et al., 2019), and influence one's feelings, thoughts, and actions (McCrae & John, 1992). For example, college students from rural and low-income families were found to have lower degrees of openness (Huang et al., 2007), and the participation of students with different personality characteristics in the classroom differed. Students with higher openness and agreeableness were found to be more efficient in terms of actively participating in class activities (Zhang & Renshaw, 2020). This difference in students' class participation will in turn result in different learning effects (Ko et al., 2016).

Some studies have used personality to explain the differences in technical mastery and have studied the relationship between personality and technology use. For example, it was found that it is often easier to teach digital skills to students with high agreeability (Palczynska, 2021). Landers and Lounsbury (2006) examined how students' personality traits and their Internet use were related, and found that those students who were more introverted, less agreeable, and less conscientious demonstrated higher levels of Internet usage. Mäkikangas et al. (2013) found a constant association between higher work engagement and the big five factors of emotional stability, conscientiousness, openness to experience, agreeableness, and extraversion. Meanwhile, Ahmed and Rasheed (2020) pointed out that university librarians exhibited a significant relationship between their personality traits and their digital literacy skills.

Furthermore, personality traits have been proved to significantly influence college students' online learning (Chuang, 2008). Individuals with high extraversion, conscientiousness, and openness showed strong learning goal orientation (Payne et al., 2007), and were more willing to invest more learning energy to achieve goals. Meanwhile, college students with strong personality traits such as agreeableness, conscientiousness, and openness demonstrated better performance in online learning (Yu, 2021). The big five model can be used to identify if there are correlations between learners' personalities and their perceptions of the online learning method (Arispe & Blake, 2012). Outside of the school's learning atmosphere, students' personality traits have a stronger influence on their learning (Singh et al., 2021), when online learning is mandatory, as during the COVID-19 pandemic, rather than merely being an alternative form of learning (Lockee, 2021). Consequently, the relationship of personality traits and digital skills in the context of online learning were analyzed in this study.

### 2.3. Purpose and research questions

Differing from existing studies which focused on demographic and socioeconomic factors, this study took personality traits as a factor of the digital skill divide. This study discusses the digital skills divide and personality traits during the period of online learning based on the method of cross-sectional and longitudinal research over nearly 2 years. The cross-sectional and longitudinal approaches (Anstey & Hofer, 2004) have been adopted in many studies on changes during a period. Saito et al. (2018) applied the longitudinal and cross-sectional method to explore the high school students' change in multiple learning-related variables from December 2016 to March 2017. In the current study, we first cross-sectionally explored how the urban and rural college students differed in their digital skills divide and personality traits over a period of nearly 2 years. Then, how personality traits predicted their digital skills was examined for each of the tests, and we longitudinally compared the change in the correlation over the period of nearly 2 years. Thus, this study proposed three research questions as follows.

RQ1: Is there a digital skill divide between the urban and rural college students during online learning in nearly 2 years?
RQ2: Are there differences in the personality traits of the urban and rural college students during nearly 2 years?
RQ3: Is there a correlation between personality traits and digital skills and would the relationship change over the period of nearly 2 years?

## 3. Methodology

### 3.1. Participants

To ensure the validity and reliability of the questionnaire, we distributed online questionnaires to college students from China by random sampling method twice, two years apart. The first questionnaire was issued in 2020 and the second two years later. The participants were informed of the purpose of the questionnaire in advance and volunteered to anonymously participate in the survey. In June 2020, 716 valid samples were collected, and in January 2022, 778 valid samples were collected. Table 1 shows the demographic information and scores of dimensions of these samples. The two samples were roughly the same in terms of gender, grade and urban-rural ratios. Some dimensions of digital skills scored slightly higher in 2022 than in 2020.

*Table 1.* The distribution of demographic information and scores of dimensions

| | Dimension | 2020 ($N$ = 716) | | 2022 ($N$ = 778) | |
|---|---|---|---|---|---|
| | | Frequency/$M$ | Percentage/$SD$ | Frequency/$M$ | Percentage/$SD$ |
| Gender | Male | 328 | 45.8% | 346 | 45% |
| | Female | 388 | 54.2% | 432 | 55% |
| Grade | Freshman | 393 | 54.9% | | |
| | Sophomore | 323 | 45.1% | | |
| | Junior | | | 381 | 49.5% |
| | Senior | | | 388 | 50.5% |
| Place of domicile | Rural | 141 | 19.7% | 197 | 25.6% |
| | Town | 159 | 22.2% | 165 | 21.5% |
| | County | 214 | 29.9% | 190 | 24.8% |
| | City | 202 | 28.2% | 217 | 28.2% |
| Digital skills | Operational skills | 3.28 | .89 | 3.66 | 1.03 |
| | Mobile skills | 4.01 | .71 | 4.32 | .74 |
| | Creative skills | 3.04 | .85 | 3.36 | .98 |
| | Social skills | 2.04 | .79 | 3.66 | .88 |
| | Safety skills | 2.55 | 1.17 | 3.23 | .95 |
| Personality traits | Extraversion | 2.51 | 1.04 | 2.53 | .94 |
| | Neuroticism | 2.97 | .98 | 2.88 | .85 |
| | Agreeableness | 2.53 | 1.09 | 2.82 | 1.13 |
| | Openness | 2.70 | .80 | 2.70 | .80 |
| | Conscientiousness | 3.04 | .85 | 3.01 | .84 |

### 3.2. Instruments

The questionnaire comprised three parts. The first part consisted of questions about demographic information, including participants' gender, grade, and place of domicile. In this study, urban and rural distribution was divided into four locations, namely rural, county, town, and city, to better explore their differences. In China, a county is usually made up of several towns and is an administrative division unit one level higher than a town and one level lower than a city (Yu et al., 2018). "Rural" and "county" refer more to rural areas, while "town" and "city" refer more to urban areas. The second part was the measurement of ICT skills which reflect the digital divide between college students. The last part of the questionnaire was to investigate college students' personality traits with responses ranging from 1 which means *strongly disagree* to 5 which means *strongly agree*. Three experts majoring in Educational Technology were invited to examine all questionnaire items. Experts' comments were accepted to revise the related items.

#### *3.2.1. Digital skills scale (Digital divide)*

As the digital divide is a multidimensional and complex concept, different researchers have emphasized different issues when evaluating it. For example, Chen et al. (2010) studied the difference in users' acceptance of technology based on TPB theory, which evaluated the first-level digital divide. The instrument proposed by Li and Hu (2020) was applied in this study. Digital skills are an individual's ability to retrieve, filter, evaluate, create, and exchange information through the safe use of digital devices. The content and measurement of digital skills are evolving with the times, providing more support for new ICT. The scale includes the five dimensions of Operational skills with 5 items, Mobile skills with 3 items, Creative skills with 5 items, Social skills with 4 items, and Safety skills with 5 items, for example: "I know how to play games with others online." In previous studies, the scale was found to be reliable. The Cronbach's α value in this study was .879.

#### *3.2.2. Personality traits scales*

The personality questionnaire is based on the Big Five model, which is the best-known measure of personality structure (Golbeck et al., 2011). The scale comprises the five dimensions of Extraversion with 4 items, Neuroticism with 4 items, Agreeableness with 4 items, Openness with 4 items, and Exteriority with 4 items, for example: "I sympathize with others' feelings," and "I am skilled at handling social situations." In previous studies, the scale was found to be reliable. The Cronbach's α value in this study was .786.

### *3.2.3. Instrument reliability and validity analysis*

To examine the internal structural validity of the scale, factor analysis was applied. The value of KMO was .855. The *p* value was less than 0.001, which showed that it could be subjected to factor analysis (Kaiser, 1974). Table 2 shows the rotation factor load for each factor. After deleting the two items: Operational Skills 5 and Safety Skills 4 (factor loading less than 0.5), 10 factors were extracted, accounting for 80.4% of the total variance. The factor loading of each construct was above 0.5, showing great convergence validity (Chin, 1998).

*Table 2.* Principal component factor analysis with variance rotation ($N$ = 1,494)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Operational skills1 | | .897 | | | | | | | | |
| Operational skills2 | | .896 | | | | | | | | |
| Operational skills3 | | .853 | | | | | | | | |
| Operational skills4 | | .850 | | | | | | | | |
| Mobile skills1 | | | | | | | | | | .913 |
| Mobile skills2 | | | | | | | | | | .926 |
| Mobile skills3 | | | | | | | | | | .786 |
| Creative skills1 | | | .799 | | | | | | | |
| Creative skills2 | | | .786 | | | | | | | |
| Creative skills3 | | | .790 | | | | | | | |
| Creative skills4 | | | .861 | | | | | | | |
| Creative skills5 | | | .816 | | | | | | | |
| Social skills1 | .866 | | | | | | | | | |
| Social skills2 | .910 | | | | | | | | | |
| Social skills3 | .912 | | | | | | | | | |
| Social skills4 | .892 | | | | | | | | | |
| Safety skills1 | | | | | .894 | | | | | |
| Safety skills2 | | | | | .903 | | | | | |
| Safety skills3 | | | | | .894 | | | | | |
| Safety skills5 | | | | | .758 | | | | | |
| Extraversion1 | | | | .860 | | | | | | |
| Extraversion2 | | | | .900 | | | | | | |
| Extraversion3 | | | | .915 | | | | | | |
| Extraversion4 | | | | .867 | | | | | | |
| Neuroticism1 | | | | | | | .881 | | | |
| Neuroticism2 | | | | | | | .874 | | | |
| Neuroticism3 | | | | | | | .854 | | | |
| Neuroticism4 | | | | | | | .856 | | | |
| Agreeableness1 | | | | | | .793 | | | | |
| Agreeableness2 | | | | | | .900 | | | | |
| Agreeableness3 | | | | | | .908 | | | | |
| Agreeableness4 | | | | | | .893 | | | | |
| Openness1 | | | | | | | | | .856 | |
| Openness2 | | | | | | | | | .880 | |
| Openness3 | | | | | | | | | .853 | |
| Openness4 | | | | | | | | | .859 | |
| Conscientiousness1 | | | | | | | | .868 | | |
| Conscientiousness2 | | | | | | | | .866 | | |
| Conscientiousness3 | | | | | | | | .856 | | |
| Conscientiousness4 | | | | | | | | .806 | | |

Meanwhile, Cronbach's α, average variance extracted (AVE), and complex reliability (CR) were tested to evaluate the reliability, convergence validity, and discriminant validity of the structure. CR and the Cronbach's α value of each construct were greater than 0.7 (as seen in Table 3), indicating that the reliability of the constructs was good (Hair et al., 2019). Meanwhile, the AVE of each construct was greater than 0.5, showing the good convergence validity of the scale (Fornell & Larcker, 1981).

*Table 3.* Dimension reliability and validity analysis ($N$ = 1,494)

| Latent variable | FL | CR | AVE | Cronbach's α |
|---|---|---|---|---|
| Operational skills | .850~.896 | .9284 | .7644 | .923 |
| Mobile skills | .786~.926 | .9088 | .7696 | .892 |
| Creative skills | .786~.861 | .9055 | .6575 | .888 |
| Social skills | .866~.912 | .9416 | .8014 | .953 |
| Safety skills | .758~.903 | .9216 | .7471 | .923 |
| Extraversion | .860~.915 | .9357 | .7846 | .935 |
| Neuroticism | .854~.881 | .9233 | .7505 | .910 |
| Agreeableness | .793~.908 | .9286 | .7652 | .901 |
| Openness | .853~.880 | .9205 | .7432 | .888 |
| Conscientiousness | .806~.868 | .9119 | .7214 | .873 |

Finally, we conducted a discriminant validity test on the scale. The square root of AVE of each dimension exceeded the absolute value of Pearson's correlation coefficient between the two dimensions (see Table 4). This indicated that the scale had good discriminant validity (Schumacker & Lomax, 2016).

*Table 4.* Dimension Discriminant Validity Analysis ($N$ = 1,494).

| Construct | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Operational skills | **.874** | | | | | | | | | |
| Mobile skills | .020 | **.877** | | | | | | | | |
| Creative skills | .191** | .061* | **.811** | | | | | | | |
| Social skills | .277** | .279** | .214** | **.895** | | | | | | |
| Safety skills | .301** | .221** | .142** | .400** | **.864** | | | | | |
| Extraversion | .162** | .189** | .253** | .131** | .169** | **.886** | | | | |
| Neuroticism | .173** | .145** | .200** | .072** | .176** | .330** | **.866** | | | |
| Agreeableness | .159** | .212** | .171** | .241** | .076** | .168** | .021 | **.875** | | |
| Openness | .028 | -.099** | .034 | .003 | .001 | .022 | .003 | .118** | **.862** | |
| Conscientiousness | -.012 | .012 | -.043 | .004 | -.098** | -.078** | -.026 | .074** | -.051 | **.849** |

*Note.* $^{**}p < .01$; $^{*}p < .05$. Figures on the diagonal are the square roots of AVE while others are the inter-construct correlation values.

### 3.3. Data analysis

There were three stages of data analysis in this study. The first stage involved the two rounds of data being used for factor analysis to assure the internal structural validity of the scale; the reliability and convergence validity of the scale were also tested.

In the second stage, using the two rounds of data, one-way ANOVA variance analysis was performed to examine the differences in the college students' place of domicile, digital skills, and personality traits, and compared the similarities and differences of the two sets of data.

In the third stage, according to the analysis results of the second stage, we used the two rounds of data to build a model of the five dimensions of personality traits that were significantly different from the place of domicile of the college students and their digital skills. Confirmatory factor analysis (CFA) aimed to check the reliability and validity of the model, to analyze the establishment of each path, and to compare the similarities and differences between the two models.

## 4. Results

In this study, factor analysis and one-way ANOVA variance analysis were conducted using SPSS 24.0, whereas structural model verification was performed with Amos 24.0. The data analysis results are reported as follows.

### 4.1. Analysis of the digital skills divide between the urban and rural college students

As shown in Table 5, one-way ANOVA variance analysis was used in this study to test differences in the digital skills of the rural and urban college students. The analysis of the sample data in 2020 found significant differences in all dimensions of digital skills of the rural and urban college students (Operational skills: $F$ = 4.280, $p < .01$; Creative skills: $F$ = 82.110, $p < .001$; Social skills: $F$ = 6.939, $p < .001$; Safety skills: $F$ = 3.013, $p < .05$) except for mobile skills ($F$ = 1.384, $p > .05$). The results were broadly consistent with the data analyzed nearly 2 years later. Partial eta squared (*η2*) was applied to represent effect size (small effect ≥ 0.01 and < 0.06, moderate effect ≥ 0.06 and < 0.14, large effect ≥ 0.14) (Dörnyei, 2007). It can be found that, except for the dimension of creative skills, the effect sizes of the other three dimensions were all greater in 2022 than in 2020; Operational skills: $\eta^2$ = .018 (2020) < .049 (2022); Social skills: $\eta^2$ = .028 (2020) < .182 (2022); and Safety skills: $\eta^2$ = .013 (2020) < .061 (2022), indicating that the development of online education will increase the digital divide of rural and urban college students. Therefore, RQ1 was answered.

*Table 5.* One-way ANOVA variance analysis of urban and rural college students' digital skills

| Dependent variable | 2020 ($N$ = 716) | | | 2022 ($N$ = 778) | | |
|---|---|---|---|---|---|---|
| | $F$ | $p$ | Partial Eta squared | $F$ | $p$ | Partial Eta squared |
| Operational skills | 4.280 | .005$^{**}$ | .018 | 13.211 | .000$^{***}$ | .049 |
| Mobile skills | 1.384 | .246 | | .998 | .393 | |
| Creative skills | 82.110 | .000$^{***}$ | .257 | 88.217 | .000$^{***}$ | .257 |
| Social skills | 6.939 | .000$^{***}$ | .028 | 56.732 | .000$^{***}$ | .182 |
| Safety skills | 3.013 | .029$^{*}$ | .013 | 16.519 | .000$^{***}$ | .061 |

*Note.* $^{***}p < .001$; $^{**}p < .01$; $^{*}p < .05$.

### 4.2. Analysis of the differences in the personality traits of the urban and rural college students

This study applied one-way ANOVA variance analysis to check the differences in personality traits of the urban and rural college students (see Table 6). The analysis of the sample data in 2020 showed significant differences in all dimensions of personality traits of the urban and rural college students (Extraversion: $F$ = 7.200, $p < .001$; Neuroticism: $F$ = 6.287, $p < .001$; Agreeableness $F$ = 5.772, $p < .01$) except for openness ($F$ = 1.389, $p > .05$) and conscientiousness ($F$ = 1.198, $p > .05$). According to the mean value (shown in Table 7), rural college students had lower extraversion, neuroticism and agreeableness than urban college students in 2020. The results were broadly consistent with the data analyzed nearly 2 years later. According to the effect size, the three dimensions in 2020 were slightly larger than those in 2022, but they all remained in the same effect dimension: Extraversion: $\eta^2$ = .029 (2020) > .023 (2022); Neuroticism: $\eta^2$ = .026 (2020) > .017 (2022); Agreeableness: $\eta^2$ = .024 (2020) > .012 (2022), indicating that there was no change in urban and rural college students' personality traits differences in nearly 2 years. As shown in Table 7, the level of extraversion, neuroticism and agreeableness among rural college students was also lower than that of urban college students in 2022. Therefore, RQ2 was answered.

*Table 6.* One-way ANOVA variance analysis of urban and rural college students' personality traits

| Dependent variable | 2020 ($N$ = 716) | | | 2022 ($N$ = 778) | | |
|---|---|---|---|---|---|---|
| | $F$ | $p$ | Partial Eta Squared | $F$ | $p$ | Partial Eta Squared |
| Extraversion | 7.200 | .000$^{***}$ | .029 | 5.893 | .001$^{**}$ | .023 |
| Neuroticism | 6.287 | .000$^{***}$ | .026 | 4.458 | .004$^{**}$ | .017 |
| Agreeableness | 5.772 | .001$^{**}$ | .024 | 3.128 | .025$^{*}$ | .012 |
| Openness | 1.389 | .245 | | 1.987 | .114 | |
| Conscientiousness | 1.198 | .310 | | .261 | .854 | |

*Note.* $^{***}p < .001$; $^{**}p < .01$; $^{*}p < .05$.

*Table 7.* The mean value of personal traits among college students in different regions

| Dependent variable | 2020 ($N$ = 716) | | | | 2022 ($N$ = 769) | | | |
|---|---|---|---|---|---|---|---|---|
| | Rural | Town | County | City | Rural | Town | County | City |
| Extraversion | 8.94 | 9.53 | 10.03 | 10.66 | 9.35 | 10.14 | 10.67 | 10.60 |
| Neuroticism | 11.04 | 11.37 | 11.50 | 12.06 | 10.92 | 11.24 | 11.98 | 11.90 |
| Agreeableness | 10.40 | 10.72 | 11.04 | 11.34 | 10.51 | 11.15 | 11.56 | 11.77 |
| Openness | 10.49 | 10.58 | 10.86 | 10.87 | 10.48 | 10.56 | 11.02 | 11.12 |
| Conscientiousness | 11.95 | 12.12 | 12.34 | 12.12 | 11.89 | 12.00 | 12.14 | 12.14 |

### 4.3. Analysis of the correlation between personality traits and digital skills

To further explore the relationship of personality traits and digital skills, this study tested the relationship between three personalities with significant differences in the urban and rural college students' digital skills. Since there were significant differences in all dimensions of the college students' personality traits except for openness and conscientiousness, the correlations between extraversion and the five dimensions of digital skills, neuroticism and the five dimensions of digital skills, and agreeableness and the five dimensions of digital skills should be tested.

The fit indices of the two sets of data for 2020 and 2022 were within an acceptable range (RMSEA = .058 and .066 < 0.08; CFI = .947 and .923 > 0.9), indicating that the model had good reliability and validity (Hair et al., 2019). Therefore, a further path coefficient test was performed. As shown in Table 8, all the paths were significant with the exception of the path of agreeableness to safety skills ($\beta$ = .104, S.E. = .038, $t$ = .354, $p$ > .05). In other words, extraversion, neuroticism, and the five dimensions of digital skills were all positively correlated. This phenomenon did not change in nearly 2 years; that is, the data for 2022 confirmed the conclusion of 2020. The path of agreeableness to safety skills was still not supported ($\beta$ = .010, S.E. = .021, $t$ = .509, $p$ > .05). Therefore, RQ3 was answered. The results showed that personality traits are directly related to the existence of urban and rural college students' digital divide.

*Table 8*. The path analysis

| | 2020 ($N$ = 716) | | | | 2022 ($N$ = 769) | | | |
|---|---|---|---|---|---|---|---|---|
| Causal Factors | Estimate | S.E. | $t$ | $p$ | Estimate | S.E. | $t$ | $p$ |
| EXT→ OPE | .085 | .035 | 2.403 | .016$^{*}$ | .154 | .051 | 3.021 | .003$^{**}$ |
| EXT→ MOB | .059 | .022 | 2.637 | .008$^{**}$ | .121 | .032 | 3.847 | .000$^{***}$ |
| EXT→ CRE | .121 | .031 | 3.901 | .000$^{***}$ | .372 | .048 | 7.788 | .000$^{***}$ |
| EXT→ SOC | .101 | .030 | 3.412 | .000$^{***}$ | .136 | .043 | 3.168 | .002$^{**}$ |
| EXT→ SAF | .094 | .039 | 2.406 | .016$^{*}$ | .104 | .030 | 3.431 | .000$^{***}$ |
| NEU→ OPE | .157 | .035 | 4.456 | .000$^{***}$ | .188 | .049 | 3.835 | .000$^{***}$ |
| NEU→ MOB | .078 | .022 | 3.484 | .000$^{***}$ | .071 | .030 | 2.395 | .017$^{*}$ |
| NEU→ CRE | .183 | .031 | 5.863 | 000$^{***}$ | .242 | .045 | 5.353 | .000$^{***}$ |
| NEU→ SOC | .105 | .030 | 3.559 | 000$^{***}$ | .078 | .041 | 1.911 | .056 |
| NEU→ SAF | .243 | .040 | 6.116 | 000$^{***}$ | .065 | .029 | 2.255 | .024$^{*}$ |
| AGR→ OPE | .147 | .035 | 4.225 | 000$^{***}$ | .097 | .036 | 2.708 | .007$^{**}$ |
| AGR→ MOB | .104 | .022 | 4.652 | 000$^{***}$ | .102 | .022 | 4.613 | .000$^{***}$ |
| AGR→ CRE | .144 | .031 | 4.669 | 000$^{***}$ | .103 | .033 | 3.137 | .002$^{**}$ |
| AGR→ SOC | .120 | .029 | 4.134 | 000$^{***}$ | .194 | .030 | 6.409 | .000$^{***}$ |
| AGR→ SAF | .014 | .037 | .362 | .717 | .010 | .021 | .503 | .615 |

*Note.* $^{***}p$ < .001; $^{**}p$ < .01; $^{*}p$ < .05.

## 5. Discussion

### 5.1. The urban-rural divide exacerbates the digital skills divide between college students

The results revealed that a digital skills divide does indeed exist between the urban and rural college students. It showed that the digital skill divide persists between them, with the balance of digital skills equality tipped in favor of urban college students. This finding is in line with Wang's (2020) study, which revealed that differences in learning investment, learning resources, and physical and mental status resulted in a digital skill divide between the urban and rural college students. College students from rural areas lack basic learning equipment and suffer from more serious digital inequality (Azionya & Nhedzi, 2021) which led to the second-level digital divide (Lebenicnik & Starcic, 2020). Urban students of higher socioeconomic status were more in control of ICT-related activities than rural students (Hohlfeld et al., 2017). Li and Ranieri (2013) found that there were gaps between rural and urban students in the indicators of Internet inequality which led to the disadvantaged position of rural students in Internet usage. Specifically, this study found that there were significant differences in all dimensions of digital skills of the urban and rural college students, except for mobile skills. To bridge the digital divide, it is necessary to classify different types of digital skills. Mobile skills, including the ability to retrieve, choose, and evaluate information on the Internet, are basic skills for our daily life. Inequality did exist for social skills which are regarded as the ability to understand and exchange meaning via online communication and interactions and to develop one's social capital, creative skills which help individuals create, publish or share a variety of quality content online, and safety skills (Scheerder et al., 2017). In addition, this study further indicated that the digital skills divide changed significantly in nearly 2 years during the COVID-19. The massive usage of digital technologies largely magnified the digital skills divide (Beaunoyer et al., 2020). The digital skills divide of urban and rural students in the online learning has become more prominent in COVID-19. Therefore, RQ1 was answered.

#### 5.2. The urban-rural divide significantly affects the differences in college students' personality traits

A significant difference in the personality traits of urban and rural college students was also found in this study. Although personality traits are a stable variable that do not change significantly in the short term, personality trait differences of college students are significantly influenced by the urban-rural divide. This finding is consistent with Ebru's et al. (2019) conclusion that personality traits are affected by social, cultural, family, and geographical factors. Students from rural or low-income families tend to have low degrees of openness (Huang et al., 2007), which also affects their personality traits. The social capital gap between urban and rural students could be a significant factor explaining their personality differences. Seevers et al. (2015) found that agreeableness and extroversion were able to predict, respectively, the quality of ties and tie strength, while Baay et al. (2014) found a relationship between extroversion and openness to experiences and students' social capital. Specifically, this study further confirmed that there were significant differences between urban and rural college students in terms of their extraversion, neuroticism, and agreeableness, but there was no significant difference in their openness and conscientiousness. And there was no change in urban and rural college students' personality trait differences in nearly 2 years. RQ2 was thus answered.

#### 5.3. Bridging the digital skills divide between urban and rural college students by optimizing their personality traits

By investing the relation between personality traits and digital skill divide, we found that personality traits play a positive role both in the application of the Internet and the contraction of the digital skill divide. This finding could be partly observed in previous studies in which students with high agreeableness had a stronger ability to learn digital skills (Palczynska, 2021). Landers and Lounsbury's (2006) study also confirmed this point further. They showed that introverted and unpopular college students were more inclined to use the Internet. In the prior study, personality traits were regarded as having an influence on the way people interact with digital tools (Parmaksiz, 2022), which may affect their acquisition of digital skills. However, this study found that except for agreeableness and safety skills, extraversion, neuroticism, and the five dimensions of digital skills were all positively correlated through the two tests. Before individuals can adopt a new technology, they first must accept the new idea and develop an attitude towards it. Individual knowledge and characteristics influence technology adoption and dissemination. It means that the optimization of personality traits promotes the growth of digital skills and is conducive to reducing the digital skills divide. In addition, the difference in personality traits leads to digital skills divide between college students from different places of residence. This phenomenon did not change in nearly 2 years. Therefore, RQ3 was answered.

## 6. Conclusion and limitations

#### 6.1. Conclusions

To understand and address this important issue, this study, based on longitudinal research, explored the digital skills divide of urban and rural college students, the influence of personality traits on digital skills, and the changes in differences and correlation during nearly 2 years of online learning. According to the results, the significant difference in the digital skills of rural and urban college students was observed in 2020 when online learning was launched on a widespread scale, and this phenomenon was also pronounced in 2022. As for differences in personality traits and their correlation with digital skills, the results showed that urban and rural college students showed differences in three dimensions of personality traits (Extraversion, Neuroticism, Agreeableness). More importantly, the personality traits including extraversion and neuroticism were positively correlated to the five dimensions of digital skills.

### 6.2. Implications

There are some theoretical implications of this study. Firstly, the results confirmed differences in the digital skills of urban and rural college students in different dimensions, contributing to the expansion of digital divide studies on online learning. Secondly, the study has explored the influence of personality traits on the second-level digital divide, which differs from previous studies which focused on demographic, socioeconomic, and regional influences on the digital divide (e.g., Drabowicz, 2014; Gameel & Wilkins, 2019). Therefore, this study has expanded the relevant literature on determinants of the digital divide.

This study also has some practical implications, namely that the results can be utilized to enhance key stakeholders' awareness of the digital divide of urban and rural college students while learning online. More attention should be paid to the development of the digital skills of college students from rural areas. In addition, the findings can be used to encourage rural college students to develop their own personalities and make more use of ICT on campus to mitigate the digital skills divide caused by socio-economic factors. Furthermore, the digital divide does not only exist in China but is an international phenomenon. The digital divide is present at a global level, and the specific solutions identified in this study should also be applied at a global level.

### 6.3. Limitations and future study

There are several limitations in this study that needs consideration in future research. Firstly, the findings only apply to college students who can freely access the Internet on campus. For college students at other educational levels or from areas where ICTs are not well developed, the phenomena in the digital access divide should also be discussed. Secondly, the online learning outcomes and the digital outcoming divide among college students also deserve our attention. Thirdly, this study only discusses the influence of personality traits on the digital divide; other personal psychological factors (e.g., self-efficacy and motivation) were not included. Therefore, the influence of multiple personal factors on the digital divide should be discussed in future research. Furthermore, the personality differences between urban and rural college students are caused by social and cultural capital to some extent; whether such personal characteristics further aggravate the digital skills gap between urban and rural college students deserves to be explored in the future. Finally, the longitudinal research method was used in this study; it did not take the age or a period effect on the difference between the two surveys into consideration. In the future, a more rigorous research design is expected.

## Acknowledgement

This study was supported by the National Science and Technology Council of Taiwan under grants NSTC 111-2410-H-025 -039, NSTC 111-2410-H-019-006-MY3, NSTC 111-2622-H-019-001, and NSTC 111-2423-H-153-001-MY3.

## Conflict of interest

All the authors have declared that they have no conflict of interest in this study.

## References

Ahmed, S., & Rasheed, T. (2020). Relationship between personality traits and digital literacy skills: A Study of university librarians. *Digital Library Perspectives*, *36*(2), 191-206. https://doi.org/10.1108/DLP-02-2020-0005

Anderson, K. J. (2001). Internet use among college students: An Exploratory study. *Journal of American College Health, 50*(1), 21-26.

Anstey, K. J., & Hofer, S. M. (2004). Longitudinal designs, methods and analysis in psychiatric research. *Australian and New Zealand Journal of Psychiatry, 38*(3), 93-104. https://doi.org/10.1080/j.1440-1614.2004.01343.x

Arispe, K., & Blake, R. J. (2012). Individual factors and successful learning in a hybrid course. *System*, *40*(4), 449-465. https://doi.org/10.1016/j.syste m.2012.10.013

Azionya, C. M., & Nhedzi, A. (2021). The Digital divide and higher education challenge with emergency online learning: Analysis of tweets in the wake of the COVID-19 lockdown. *Turkish Online Journal of Distance Education, 22*(4), 164-182. https://doi.org/10.17718/tojde.1002822

Azubuike, O. B., Adegboye, O., & Quadri, H. (2021). Who gets to learn in a pandemic? Exploring the digital divide in remote learning during the COVID-19 pandemic in Nigeria. *International Journal of Educational Research Ope*n, 2-2, 100022. https://doi.org/10.1016/j.ijedro.2020.100022

Baay, P. E., van Aken, M. A. G., de Ridder, D. T. D., & van der Lippe, T. (2014). Understanding the role of social capital in adolescents' Big Five personality effects on school-to-work transitions. *Journal of Adolescence,37*(5), 739-748. https://doi.org/10.1016/j.adolescence.2014.04.015

Beaunoyer, E., Dupéré, S. & Guitton, M. (2020). COVID-19 and digital inequalities: reciprocal impacts and mitigation strategies. *Computers in Human Behavior, 111*, 106424. https://doi.org/10.1016/j.chb.2020.106424

Cabello-Hutt, T., Cabello, P., & Claro, M. (2018). Online opportunities and risks for children and adolescents: The Role of digital skills, age, gender and parental mediation in Brazil. *New Media & Society, 20*(7), 2411-2431. https://doi.org/10.1177/1461444817724168

Chen, D., Lin, Z., & Lai, F. (2010). Crossing the chasm-understanding China's rural digital divide. *Journal of Global Information Technology Management*, *13*(2), 4-36. https://doi.org/10.1080/1097198X.2010.10856513

Chin, W. W. (1998). Commentary: Issues and opinion on structural equation modeling. *MIS Quarterly, 22*(1), vii-xvi.

Chipeva, P., Cruz-Jesus, F., Oliveira, T., & Irani, Z. (2018). Digital divide at individual level: Evidence for Eastern and Western European countries. *Government Information Quarterly, 3*5(3), 460-479. https://doi.org/10.1016/j.giq.2018.06.003

Chuang, C. M. (2008). *The Study of college students' learning styles, big five personality traits, learning motivation and their on-line learning satisfaction*. (Unpublished master's thesis). Hsuan Chuang University, Taiwan.

Correa, T. (2010). The Participation divide among "online experts": Experience, skills and psychological factors as predictors of college students' web content creation. *Journal of Computer-Mediated Communication, 16*(1), 71-92. https://doi.org/10.1111/j.1083-6101.2010.01532.x

Cotten, S. R., & Jelenewicz, S. M. (2006). A Disappearing digital divide among college students? Peeling away the layers of the digital divide. *Social Science Computer Review, 24*(4), 497-506. https://doi.org/10.1177/0894439306286852

Dhawan, S. (2020). Online learning: A Panacea in the time of COVID-19 crisis. *Journal of Educational Technology Systems, 49*(1), 5-22. https://doi.org/10.1177/0047239520934018

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41*, 414-440.

DiMaggio, P., Hargittai, E., Celeste, C., & Shafer, S. (2004). Digital inequality: From unequal access to differentiated use. In K. M. Neckerman (Ed.), *Social Inequality* (pp. 355-400). Russell Sage Foundation.

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press.

Drabowicz, T. (2014). Gender and digital usage inequality among adolescents: A Comparative study of 39 countries. *Computers & Education, 74*, 98-111. https://doi.org/10.1016/j.compedu.2014.01.016

Ebru, A., Hande, K., & Gül, K. (2019). The Effect of university students' individual innovation and lifelong learning trends on entrepreneurship orientation. *Sustainability, 11*(22), 6201. https://doi.org/10.3390/su11226201

Eshet, Y. (2004). Digital literacy: A Conceptual framework for survival skills in the digital era. *Journal of Educational Multimedia and Hypermedia, 13*(1), 93-106.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39-50. https://doi.org/10.1177/002224378101800312

Friedman, W. H. (2001). The Digital divide. In *Proceedings of the 7th Americas' Conference on Information Systems (AMCIS)* (pp. 2081-2086). http://aisel.aisnet.org/amcis2001/401

Gameel, B. G., & Wilkins, K. G. (2019). When it comes to MOOCs, where you are from makes a difference. *Computers & Education, 136*, 49-60. https://doi.org/10.1016/j.compedu.2019.02.014

Golbeck, J., Robles, C., & Turner, K. (2011). Predicting personality with social media. *Association for Computing Machinery. 11*, 253-262. https://doi.org/10.1145/1979742.1979614

Goode, J. (2010). The Digital identity divide: how technology knowledge impacts college students. *New Media & Society, 12*(3), 497-513. https://doi.org/10.1177/1461444809343560

Hargittai, E., & Hsieh, Y. P. (2012). Succinct survey measures of web-use skills. *Social Science Computer Review, 30*(1), 95-107. https://doi.org/10.1177/0894439310397146

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage.

Hohlfeld, T. N., Ritzhaupt, A. D., Dawson, K., & Wilson, M. L. (2017). An Examination of seven years of technology integration in Florida schools: Through the lens of the levels of digital divide in schools. *Computers & Education, 113*, 135-161. https://doi.org/10.1016/j.compedu.2017.05.017

Huang, X., Ling, H., Yang, B., & Dou, G. (2007). Screening of personality disorders among Chinese college students by personality diagnostic questionnaire-4+. *Journal of Personality Disorders, 21*(4), 448-454. https://doi.org/10.1521/pedi.2007.21.4.448

Kaiser, H. F. (1974). An Index of factorial simplicity. *Psychometrika, 39*(1), 31-36. https://doi.org/10.1007/BF02291575.

Ko, J., Park, S., Yu, H., Kim, S., & Kim, D. (2016). The Structural relationship between student engagement and learning outcomes in Korea. *Asia-Pacific Education Researcher, 25*(1), 147-157. https://doi.org/10.1007/s40299-015-0245-2

Landers, R. N., & Lounsbury, J. W. (2006). An Investigation of Big Five and narrow personality traits in relation to Internet usage. *Computers in Human Behavior, 22*(2), 283-293. https://doi.org/10.1016/j.chb.2004.06.001

Lebenicnik, M., & Starcic, A. I. (2020). Examining the contemporary digital divide of university students with specific reference to students with special educational needs. *British Journal of Educational Technology, 51*(6),2422-2441. https://doi.org/10.1111/bjet.12920

Li, F., Wu, F. & Wang, Z. (2022). Distance higher education and regional equality in China. *European Journal of Education, 57*, 273– 288. https://doi.org/10.1111/ejed.12502

Li, X., & Hu, R., J. (2020). Developing and validating the digital skills scale for school children (DSS-SC). *Information, Communication & Society, 25*(10), 1365-1382. https://doi.org/10.1080/1369118X.2020.1864002

Li, Y., & Ranieri, M. (2013). Educational and social correlates of the digital divide for rural and urban children: A Study on primary school students in a provincial city of China. *Computers & Education, 60*(1), 197-209. https://doi.org/10.1016/j.compedu.2012.08.001

Liao, P. A., Chang, H.-H., Wang, J.-H., & Sun, L.-C. (2016). What are the determinants of rural-urban digital inequality among schoolchildren in Taiwan? Insights from blinder-oaxaca decomposition. *Computers & Education, 95*, 123-133. https://doi.org/10.1016/j.compedu.2016.01.002

Livingstone, S., & Helsper, E. (2010). Balancing opportunities and risks in teenagers' use of the internet: The Role of online skills and internet self-efficacy. *New Media & Society, 12*(2), 309-329. https://doi.org/10.1177/1461444809342697

Lockee, B. B. (2021). Online education in the post-COVID era. *Nature Electronics, 4*(1), 5-6. https://doi.org/10.1038/s41928-020-00534-0

Mabila, J., Ssemugabi, S., & Gelderblom, H. (2013). Does assessing e-skills competence at an open distance learning. *Higher Education Institution Matter*, 539-544.

Mäkikangas, A., Feldt, T., Kinnunen, U., & Mauno, S. (2013). Does personality matter? A Review of individual differences in occupational well-being. *Advances in positive organizational psychology, 1*, 107-143. https://doi.org/10.1108/S2046-410X(2013)0000001008

McCrae, R. R., & John, O. P. (1992). An Introduction to the five-factor model and its applications. *Journal of Personality, 60*(2), 175-215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

Nerse, S. (2020). Inequalities in digital education: Rural-urban and socioeconomic differences. *Insan & Toplum, 10* (4), 413-444. https://doi.org/10.12658/M0548

Nguyen, M. H., Hargittai, E., & Marler, W. (2021). Digital inequality in communication during a time of physical distancing: The case of COVID-19. *Computers in Human Behavior, 120*, 106717. https://doi.org/10.1016/j.chb.2021.106717

Palczynska, M. (2021). Overeducation and wages: the role of cognitive skills and personality traits. *Baltic Journal of Economics, 21*(1), 85-111. https://doi.org/10.1080/1406099X.2021.1950388

Parmaksiz, İ. (2022). The Mediating role of personality traits on the relationship between academic self-efficacy and digital addiction. *Education and Information Technologies, 27*(6), 8883-8902. https://doi.org/10.1007/s10639-022-10996-8.

Payne, S. C., Youngcourt, S. S., & Beaubien, J. M. (2007). A Meta-analytic examination of the goal orientation nomological net. *Journal of Applied Psychology, 92*(1), 128-150. https://doi.org/10.1037/0021-9010.92.1.128

Ren, W., Zhu, X., & Yang, J. (2022). The SES-based difference of adolescents' digital skills and usages: An explanation from family cultural capital. *Computers & Education, 177*, 104382. https://doi.org/10.1016/j.compedu.2021.104382

Riggins, F., & Dewan, S. (2005). The Digital divide: Current and future research directions. J*ournal of the Association for Information Systems, 6*(12), 298-337. https://doi.org/10.17705/1jais.00074

Robinson, L., Cotten, S. R., Ono, H., Quan-Haase, A., Mesch, G., Chen, W., Schulz, J., Hale, T. M. & Stern, M. J. (2015). Digital inequalities and why they matter. *Information, Communication & Society, 18(5)*, 569-582. https://doi.org/10.1080/1369118X.2015.1012532

Saito, K., Dewaele, J., Abe, M., & In'nami, Y. (2018). Motivation, emotion, learning experience, and second language comprehensibility development in classroom settings: A Cross-sectional and longitudinal study. *A Journal of Research in Language Studies, 68*(3), 709-743. https://doi.org/10.1111/lang.12297

Seevers, M. T., Johnson, B. R., & Darnold, T. C. (2015). Social networks in the classroom: Personality factors as antecedents of student social capital. *American Journal of Business Education,8(3)*, 193-206. https://doi.org/10.19030/ajbe.v8i3.9281

Scheerder, A., Van Deursen, A., & Van Dijk, J. (2017). Determinants of internet skills, uses and outcomes. A Systematic review of the second- and third-level digital divide. *Telematics and Informatics, 34*(8), 1607-1624. https://doi.org/10.1016/j.tele.2017.07.007

Schumacker, R. E. & Lomax, R. G. (2016). *A Beginner's guide to structural equation modeling*. Routledge.

Singh, J. B., Sharma, S. K., & Gupta, P. (2021). Physical learning environment challenges in the digital divide: How to design effective instruction during COVID-19? *Communications of the Association for Information Systems, 48*(1), 18. https://doi.org/10.17705/1CAIS.04818

Stern, M. J., Adams, A. E., & Elsasser, S. (2009). Digital inequality and place: The Effects of technological diffusion on internet proficiency and usage across rural, suburban, and urban counties. *Sociological Inquiry, 79*(4), 391-417. https://doi.org/10.1111/j.1475-682X.2009. 00302.x

Tang, J. H., Chen, M. C., Yang, C. Y., Chung, T. Y., & Lee, Y. A. (2016). Personality traits, interpersonal relationships, online social support, and Facebook addiction. *Telematics and Informatics, 33*(1), 102-108. https://doi.org/10.1016/j.tele.2015.06.003

Tien, F. F., & Fu, T. T. (2006). The correlates of the digital divide and their impact on college student learning. *Computers & Education, 50*(1), 421-436. https://doi.org/10.1016/j.compedu.2006.07.005

van Deursen, A., & van Dijk, J. (2011). Internet skills and the digital divide. *New Media & Society, 13*(6), 893-911. https://doi.org/10.1177/1461444810386774

van Dijk, J. A. G. M. (2006). Digital divide research, achievements and shortcomings. *Poetics, 34*(4-5), 221-235. https://doi.org/10.1016/j.poetic.2006.05.004

Walczuch, R., & Lundgren, H. (2004). Psychological antecedents of institution-based consumer trust in e-retailing. *Information and Management, 42*(1), 159-177. https://doi.org/10.1016/j.im.2003.12.009

Wang, W. (2020). The Study on online learning frequency, learning engagement and mental health of college students during COVID-19. In *2020 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)* (pp. 705-707). IEEE. https://doi.org/10.1109/WIIAT50758.2020.00107

Wei, K. K., Teo, H. H., Chan, H. C., & Tan, B. C. Y. (2011). Conceptualizing and testing a social cognitive model of the digital divide. *Information Systems Research, 22*(1), 170-187. https://doi.org/10.1287/isre.1090.0273

Welser, H. T., Khan, M. L., & Dickard, M. (2019). Digital remediation: Social support and online learning communities can help offset rural digital inequality. *Information, Communication & Society, 22*(5), 717-723. https://doi.org/10.1080/1369118X.2019.1566485

Yu, H., Deng, Y., & Xu, S. (2018). Evolutionary pattern and effect of administrative division adjustment during urbanization of China: Empirical analysis on multiple scales. *Chinese Geographical Science, 28*(5), 758-772. https://doi.org/10.1007/s11769-018-0990-2

Yu, Z. (2021). The Effects of gender, educational level, and personality on online learning outcomes during the COVID-19 pandemic. *International Journal of Educational Technology in Higher Education, 18*(1), 14. https://doi.org/10.1186/s41239-021-00252-3

Zhang, D. C., & Renshaw, T. L. (2020). Personality and college student subjective wellbeing: A Domain-specific approach. *Journal of Happiness Studies, 21*(3), 997-1014. https://doi.org/10.1007/s10902-019-00116-8

Zhao, L., Cao, C., Li, Y., & Li, Y. (2022). Determinants of the digital outcome divide in e-learning between rural and urban students: Empirical evidence from the COVID-19 pandemic based on capital theory. *Computers in Human Behavior, 130,* 107177. https://doi.org/10.1016/j.chb.2021.107177