GCCCE 2023

第 27 届全球华人计算机教育应用大会

THE 27TH GLOBAL CHINESE CONFERENCE ON COMPUTERS IN EDUCATION

北京師範大學
BEIJING NORMAL UNIVERSITY

GCCCE2023

2023 年 5 月 27 日 -5 月 31 日

MAY 27 - MAY 31, 2023

中国 北京 | 北京师范大学（昌平校区）

BEIJING, CHINA | BEIJING NORMAL UNIVERSITY (CHANGPING CAMPUS)

Main Conference Proceedings (English Paper)

大会论文集（英文论文）

ISBN: 978-626-974-780-1

ISSN: 3005-3226

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

Publisher: Global Chinese Society on Computers in Education
Title: Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023).
Editors: Editors: Ju-Ling Shih, Bo Jiang, Minhsien Lee, Chengjiu Yin, Daner Sun, Yu Lu
Edition date:2023,06
Edition: princeps
ISBN: 978-626-974-780-1 (PDF)
ISSN: 3005-3226

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

第 27 届全球华人计算机教育应用大会

The 27th Global Chinese Conference on Computers in Education

GCCCE 2023 大会论文集（英文论文）

GCCCE 2023 Main Conference Proceedings (English Paper)

主编 Editors

Ju-Ling Shih, Taiwan Central University

Bo Jiang, East China Normal University

Minhsien Lee, Taiwan Normal University

Chengjiu Yin, Kobe University, Japan

Daner Sun, The Education University of Hong Kong

Yu Lu, Beijing Normal University

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

Table of Contents

Main Conference English Paper Track

1. Message from the Organizer

Global Chinese Conference on Computers in Education (GCCCE) is an annual international academic conference organized by the Global Chinese Society for Computers in Education (GCSCE). The 24th GCCCE was held at Northwest Normal University, Lanzhou in September 2020. The 25th GCCCE was held by The Education University of Hong Kong, Beijing Normal University, Taiwan Normal University and National Institute of Education, Nanyang Technological University in blended mode on 11 – 15 September 2021. The conference has been developed as a premier academic event for researchers, practitioners and policy makers in the Chinese communities for the worldwide dissemination and sharing of

ideas for research in the field of Computers in Education. The 26th Global Chinese Conference on Computers in Education (GCCCE 2022) will be organized online by Taiwan Tsing Hua University, The Education University of Hong Kong and East China Normal University on 28 May – 1 June 2022.

The 27th Global Chinese Conference on Computers in Education (GCCCE 2023) will be organized online by Beijing Normal University, on 27 May – 31 May 2023. The conference program will comprise keynotes, paper presentations, workshops, forums, Doctoral Student Forum and Teacher Forum.

It is worth noting that, this is the third year of conference since the English Paper Track (EPT) has been found, inviting both ethnic Chinese and non-ethnic Chinese leading international scholars to form a Program Committee in attempt to attract papers from non-Chinese authors around the world. The EPT comes with an individual English-only sharing session, welcoming all conference participants to attend and interact with international scholars. All papers accepted by EPT will be independently edited and published as the GCCCE2022 English Paper Track Proceedings. In addition, like last year, two English keynote speeches will be delivered by two leading scholars.

GCCCE 2023 will be a mixed online and offline meeting format. The offline conference was set up in Changping Campus of Beijing Normal University, and the online conference channel was Zoom conference system. Participants who are unable to attend can participate online through the Zoom conference system

Apart from the inaugurated EPT, nine theme-based sub-conferences are featured in this GCCCE as usual, namely,

C1: Learning Sciences & Computer-Supported Collaborative Learning

C2: Mobile, Ubiquitous & Contextual Learning

C3: Joyful Learning, Educational Games & Digital Toys

C4: Technology in Higher Education & Adult Learning, and Teachers' Professional Development

C5: Technology-Enhanced Language and Humanities Learning

C6: Artificial Intelligence in Education & Smart Learning Environments

C7: Learning Analytics & Assessments

C8: STEM & Maker Education

C9: Educational Technology: Innovations, Policies & Practice

Within EPT and each sub-conference, an Executive Chair, Co-Chairs and Program Committee (PC) Members were appointed to shoulder the review and programming process. Each sub-conference was also set up with additional evaluation. GCCCE 2023 calls for papers from scholars around the world (not limited to ethnic Chinese).

This year, four academic experts and scholars are invited to be the keynote speakers. These keynotes are,

Keynote 1: Time to Wake Up from Our Innovative Learning Dreams and Make Smarter Learning a Reality
Speaker: Curtis J. Bonk, School of Education, Indiana University

Keynote 2: 性格、行為、策略：複合式學習遊戲的跨域性與動態性
Speaker : 施如齡, 網路學習科技研究所, 國立中央大學

Keynote 3: AI+X to boost interdisciplinary research and foster talents
Speaker : 吴飞， 计算机科学与技术学院，浙江大学

Keynote 4 : Integrating Human Knowledge and Machine Intelligence in Education
Speaker : Minhong (Maggie) Wang, Faculty of Education, The University of Hong Kong

The four keynotes, the nine sub-conferences, the EPT and two topical discussion panels form the main conference of GCCCE 2023. Other than that, like previous years, there will be three pre-conference events, including K-12 Teachers' Forum, workshops and Doctoral Student Forum.

A total of 8 workshops on various research topics were featured this year, which accepted 68 workshop papers in total, namely,

Workshop 1：迎接"元宇宙"的世代，如何融入新科技于教学工作坊
Workshop 2：知识建构与教育数字化转型工作坊
Workshop 3：学习投入与学习行为建模工作坊
Workshop 4：創新互動回饋科技提升學習動機工作坊
Workshop 5：学习科学与游戏化学习工作坊
Workshop 6：计算机支持的个性化和协作学习工作坊
Workshop 7：新兴技术支持的协作学习设计与评价
Workshop 8：第五屆「親身體驗，好就用」: 遊戲式 / 遊戲化與教育玩具工作坊

Besides, Doctoral Student Forum will be established in the conference and 6 doctoral candidates will participate. A total of 5 experts and scholars will be invited to make comments and review.

We would like to express our deepest gratitude to all the chairs, co-chairs, committee members and volunteers of the sub-conferences, EPT, workshops, Teachers' Forum, Doctoral Student Forum, and the Local Organizing Committee. We thank them for their contributions and assistance to the conference.

We are running GCCCE2023 at unprecedented times. Yet we sincerely hope that our conference will bring inspiration and a magnificent experience to all the online or physically attending participants. Together, we shall build a stronger, resilient and more

internationalized GCCCE community, and continue to relay the GCCCE torch to successive hosts and new generations of scholars in the coming years.

Ju-Ling Shih, Taiwan Central University
Conference Chair

Bo Jiang, East China Normal University
International Program Coordination Chair

Minhsien Lee, Taiwan Normal University
Chengjiu Yin, Kobe University, Japan
Daner Sun, The Education University of Hong Kong
International Program Coordination Co-Chair

Yu Lu, Beijing Normal University
Local Organizing Committee Co-Chair

2. Conference Organization

Organiser :

Global Chinese Society for Computers in Education (GCSCE)

Hosts :

Beijing Normal University, Beijing

Advanced Innovation Center for Future Education(AICFE)

Conference Chair :

Ju-Ling Shih, Central University

International Program Coordination Chair :

Bo Jiang, East China Normal University

International Program Coordination Co-Chair:

Minhsien Lee, Taiwan Normal University

Chengjiu Yin, Kobe University, Japan

Daner Sun, The Education University of Hong Kong

Local Organizing Committee Co-Chair :

Yu Lu, Beijing Normal University

English Paper Track Programme Committee:

Executive Chair:

Juan Zhou, Tokyo Institute of Technology (Japan)

3. Keynotes

主旨演讲 1

- 2023 年 5 月 29 日（星期一）11:00-12:00
- 线下会场：会堂大报告厅
- Zoom 会议号：841 996 11626，入会密码：647982

Speech title：AI+X to boost interdisciplinary research and foster talents

吴飞 教授
浙江大学计算机科学与技术学院
浙江大学上海高等教育研究院常务副院长

Speech Abstract: Artificial intelligence (AI) has the potential to enhance every technology as it resembles enabling technologies like the combustion engine or electricity. However, contemporary AI systems are good at specific predefined tasks and are unable to learn by themselves from data or from experience, intuitive reasoning, and adaptation. From the perspective of overcoming the limitations of existing AI, interdisciplinary scientific efforts are necessary to boost future research in this field. In this talk, I will review the AI history and the outline the trend of AI, and discuss how AI+X boost interdisciplinary research and foster talents.

Speaker Bio: Fei Wu received his B.Sc., M.Sc. and Ph.D. degrees in computer science from Lanzhou University, University of Macau and Zhejiang University in 1996, 1999 and 2002 respectively. From October, 2009 to August 2010, Fei Wu was a visiting scholar at Prof. Bin Yu's group, University of California, Berkeley. Currently, He is a Qiushi distinguished professor of Zhejiang University at the college of computer science. He is the deputy dean of Shanghai Institute for Advanced Study of Zhejiang University, and the director of Institute of Artificial Intelligence of Zhejiang University. He is currently the Section Executive Editors-in-Chief of Engineering, editorial members of Frontiers of Information Technology & Electronic Engineering. He has won various honors such as the Award of National Science Fund for Distinguished Young Scholars of China (2016). His research interests mainly include Artificial Intelligence, Multimedia Analysis and Retrieval and Machine Learning.

主旨演讲 2

- 2023 年 5 月 30 日（星期二）8:30-9:30
- 线下会场：会堂大报告厅
- Zoom 会议号：816 792 89940，入会密码：869128

Speech Title: Time to Wake Up from Our Innovative Learning Dreams and Make Smarter Learning a Reality

Professor Curtis J. Bonk
School of Education, Indiana University, USA

Speech Abstract: For a half century, educators, psychologists, and researchers have been predicting that highly intensive, innovative, and individualized learning formats are only a few years away. Learners of all ages would enter enticing microworlds, highly engaging learning experience holodecks, fully immersive hands-on scenarios, high fidelity simulations and games, AI-based adaptive microlearning snippets, and completely free and open educational resources and courses on any topic. Massive open online classes were promised one day and then on demand microlearning snippets were delivered in the next. The learning related dreams we had in past decades were quickly forgotten as the next wave of learning technology came along. But all those dreams will prove pointless if they fail to address true problems or issues that some aspect of society is struggling with. It is time to wake up from such dreams of a glistening technological future and have our dream machines help us envision a world filled with open, informal, adaptive, nontraditional, and self-directed learning opportunities. When that happens, we will truly have arrived in the age of smarter and more innovative forms of learning where the learner is finally in charge of the dreams.

Speaker Bio: Curtis J. Bonk is Professor in the School of Education at Indiana University (IU) teaching psychology and technology courses and Adjunct in the School of Informatics at IU. He is a former software entrepreneur, certified public accountant, corporate controller, and educational psychologist who presently is an educational technologist, award-winning writer, highly published researcher, statewide and national awardee in innovative teaching with technology, and internationally acclaimed presenter. In 2020, he was awarded the IU President's Award for Excellence in Teaching and Learning Technology — and in 2021, received the David H. Jonassen Excellence in Research Award. In April 2022, the American Educational Research Association named him a 2022 AERA Fellow for his exceptional contributions to, and excellence in, education research and the following week he was honored with the International Engagement award from the IU School of Education. Curt is the author of nearly 400 publications and has many widely used technology and learning related books. He can be contacted at cjbonk@indiana.edu and his homepage is http://curtbonk.com/.

主旨演讲 3

- 2023 年 5 月 30 日（星期二）13:30-14:30
- 线下会场：会堂大报告厅
- Zoom 会议号：816 792 89940，入会密码：869128

Speech Title: 性格、行為、策略：複合式學習遊戲的跨域性與動態性

施如齡 教授
台湾中央大學網路學習科技研究所,

Speech Abstract:

複合式遊戲主要整合各種科技型態，包括數位系統、AR、機器人、IOT 等融入於桌遊中並產生各種教學模式、策略、機制與型態的變形。在複合式遊戲中，以建構主義與社會認知為基礎，建構倣真情境，讓學生在歷史情境中，瞭解在每一片土地上的人、事、物。透過目標任務與衝突事件，讓學習者以不同角色去看到差異立場與多元觀點，提高學生歷史思維與歷史感，解開遊戲的隱藏傳遞訊息。此演講主要分為兩個部分：第一部分為「遊戲設計與科技應用」，說明遊戲設計的核心概念與多元變化；學習者、大地圖、卡牌、物件、機構、機器人、電腦系統等，相互之間的動態連結與連帶反應，使學習者探究、分析、使用、觀察遊戲參數的變化。同時，這種讓所有學生共同參與的情境、多方參與式的 IOT 物件連動複雜數值，以及經由協商、決策與互動所產生的動態結局，即是複合式遊戲的精髓與科技創新的應用發想。第二部分為「動態行為與多模分析」，說明多方動態競合的策略遊戲，讓學習者於衝突事件中進行個人選擇、集體決策、跨群溝通與協商，為課室帶來嶄新的學習樣態。學生的性格與其他相關因子如何影響其動態行為、社會關係、解難策略，都是此演講探討的議題。

Speaker Bio:

施教授為美國哥倫比亞大學師範學院教育博士，傳播與教育科技雙碩士學位。專長於遊戲式學習、跨域教學設計、質性研究、多模行為分析、人文社會研究、廣電製作與傳播、STEAM 與機器人等。於 2012 年發表全球首創 3D RPG 台灣史詩數位學習遊戲，2015 年開發探索教育型跨平台數位諮商遊戲，2018 年發表機器人融入議題導向跨域學習遊戲。近年發表一系列以數位人文為基礎之情境議題遊戲，以機器人、仿生獸、物聯網等融入跨域學習之競合策略遊戲，融合歷史思維、運算思維、設計思考等多元整合的創新教學模式。其學術論文發表於 CHB, C&E, BJET, ETS 等，擔任 IJSG 與 RPTEL 學術期刊編輯、GCSCE 學會主席，擁有實體化場景移動遊戲之海峽兩岸專利。曾榮獲台灣科技部 2011 年吳大猷先生紀念獎，以及 2012-2019 連續兩屆優秀年輕學者計畫，與中央大學教學傑出獎。

主旨演讲 4

- 2023 年 5 月 31 日（星期三）9:30-10:30
- 线下会场：会堂大报告厅
- Zoom 会议号：838 578 30298，入会密码：828286

Speech Title: Integrating Human Knowledge and Machine Intelligence in Education

Professor Minhong (Maggie) Wang,
Faculty of Education,
The University of Hong Kong

Speech Abstract: Artificial intelligence (AI) has been increasingly utilized to support education in various aspects such as empowering smart learning content, offering immediate feedback, providing intelligent support, and assisting in teachers in instructional decision making. AI beats human intelligence in terms of computing power and memory in processing a huge amount of data and information, which enables tasks to be executed at a higher speed and more accurately. However, machine intelligence is dependent on human intellect, in particular human knowledge. Also, human brains outperform machines in many aspects such as comprehending abstract terms, working with complex thoughts, creating knowledge or innovative ideas, and processing emotions. To achieve the full potential of AI in education, it is important to integrate human knowledge and the power of AI systems. This talk will discuss how recent advances in AI may impact on education in various aspects and how human knowledge and machine intelligence can work together to shape the future of education.

Speaker Bio: Dr. Minhong (Maggie) Wang is Professor and Director of the Laboratory for Knowledge Management & E-Learning in the Faculty of Education, The University of Hong Kong. She is also Eastern Scholar Chair Professor at East China Normal University and Visiting Research Professor at the Advanced Innovation Center for Future Education of Beijing Normal University. She is the Editor-in-Chief of Knowledge Management & E-Learning (indexed in Scopus & ESCI). Her research focus is on learning technologies for cognitive development, creative thinking and complex problem solving, knowledge management and visualization, and artificial intelligence applications. She has published more than 200 items including one monograph and 115 journal articles (73 in SSCI/SCI indexed journals; 48 in Q1 and 18 in Q2 journals) among others. She is recognized as ESI Top 1% Scholar in (a) Social Sciences, General, and (b) Economics & Business. More details can be found at http://web.edu.hku.hk/staff/academic/magwang.

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

Exploration and implementation of AI general course in Guangzhou Information Technology Vocational School

Jingyi Feng
Guangzhou Information Technology Vocational School
* 286839625@qq.com

Abstract: *With the fast development of artificial intelligence technology, the popularization of artificial intelligence education is one of the significant contents of China's education development strategy. Therefore, vocational schools should take their responsibilities to meet the society's demands for "AI+X" compound talents. The research analyzes the problems existing in the current artificial intelligence general courses in secondary vocational schools. For example, the insufficient research on teaching methodology, and the existing curriculum doesn't match the learning foundation of secondary vocational students. Therefore, combining the characteristics of artificial intelligence majors and the learning conditions of secondary vocational students, which provide theoretical reference and practical guidance for the general education course of artificial intelligence in secondary vocational schools.*

Keywords: vocational school, artificial intelligence curriculum, general education, task-driven teaching method

1. Introduction

In recent years, the ubiquitous big data and intelligence technology have made "Artificial Intelligence +" extraordinary in various fields. The country's policy documents also clearly propose to provide relevant guarantees for the development of the new generation of artificial intelligence, adopting a new model of "artificial intelligence + X" compound professional training to promote talents development (Ministry of Education, 2017). However, most of the secondary vocational students in our country is receiving simple skills training. What's more, there is currently no unified curriculum standard, syllabus and complete teaching material system in the country in the field of artificial intelligence. As a result, the components of artificial intelligence literacy of secondary vocational students are not clear, and the content of learning is also out of line with the development of the times and the needs of students (Yunfei, 2022). Under the background that the government, schools and teachers lack theoretical support and practical experience in the definition and design of artificial intelligence courses and for the purpose of cultivating technical talents who adapt to the thinking mode of the new era (Youjia, 2022). We try to use the task-driven teaching method to open up a new teaching path of artificial intelligence general courses for the cultivation of talents for secondary vocational students. This paper takes Guangzhou Information Technology Vocational School as an example, combined with the actual situation of secondary vocational teachers and students for analyzing.

2. The concept and current situation of artificial intelligence general courses

General education revolves around the aims of what kind of people to develop, how to develop people and for whom to develop people to adapt to the needs of talents in a rapidly changing world (Jian, 2022). The artificial intelligence course has the uniqueness of teaching content and form different from other courses (Zehui, 2022). In general, the course content about artificial intelligence mainly focuses on

cultivating students' innovative thinking, mastering relevant knowledge and skills, and establishing correct attitudes and concepts. There are still some problems in the development of artificial intelligence courses. First, The lack of research on artificial intelligence talents cultivation in existing secondary vocational education. China's 2020 secondary vocational school information curriculum standard clearly puts forward "artificial intelligence preliminary". In order to take into account both student employment and economic development, it is necessary to highlight the characteristics of vocational education of learning by doing and teaching by doing (Jian, 2020). Second, Existing artificial intelligence courses do not match the learning foundation of secondary vocational students. Existing research has found that the artificial intelligence courses offered by most schools can enrich the course content to the greatest extent and choose the teaching method according to the optimal principle (Zhixin, 2022). However, the analysis of existing courses shows that most of the content students learn involves a variety of algorithms, and students even need to have a certain basic knowledge of programming to cope with it. There are some problems exiting in the classroom learning of secondary vocational students, such as the theoretical foundation is not solid (Xiaomin, 2022). In response to the above problems, Guangzhou Information Technology Vocational School, has developed a general course of artificial intelligence with practical application as the core. It focuses on the integration of disciplines in the attributes of artificial intelligence.

3. The setting of artificial intelligence general course mode in secondary vocational schools

With the accumulation of big data in the era of artificial intelligence, the high-level application scenarios for solving specific practical problems are constantly updated. Moreover, the deep integration of scenarios with all walks of life in the society has also enabled the relevant artificial intelligence technologies to be applied to the greatest extent (Xinhua Publishing House, 2017). According to the research of Zhengli (2021). How can these highly specialized learning contents be transformed into learning tasks that are comparable to the learning level and ability characteristics of secondary vocational students, being understood and accepted by them? We need to realize the shift from teaching-centered to learning-centered in the process of classroom learning in information technology teaching. Students can complete the corresponding course tasks in the process of active knowledge and exploration. First, Clarify the main tasks of the course. Curriculum has always been the foundation of talent cultivation. One of the core issues that needs to be clarified in curriculum theory and teaching theory is whether the basis of curriculum development is object, content, process or task (Dayuan, 2017). Second， build the curriculum framework. The frame structure is shown in Fig. 1.

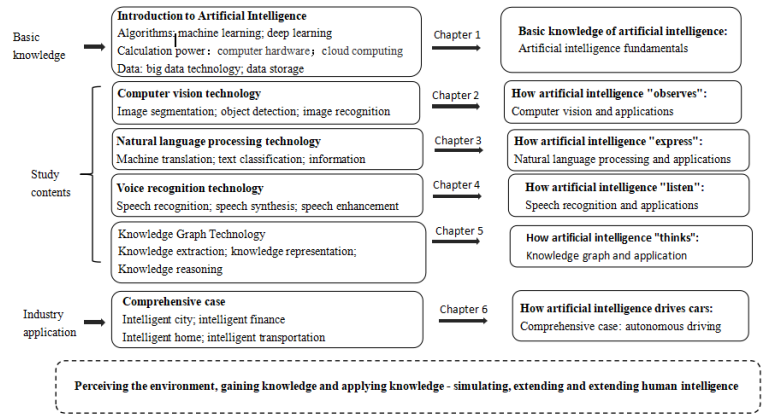


Figure 1. Artificial intelligence general curriculum framework structure.

Third, specify instructional design path. Based on the concept of task-driven teaching mode, it's significant to build courses that meet the requirements of the artificial intelligence era and highlight the academic conditions of secondary vocational students. The teaching content adopts modular closed-loop teaching. Students are more systematic in the process of learning knowledge. The basic module belongs to the course introduction period. The task of the application module is to introduce new developments in artificial intelligence related application scenarios based on the update of course content and knowledge. Programming languages were invented by engineers to tell computers what to do. Experimental module attaches importance to the bilateral situational experience of teachers and students in the form of interactive communication. The teaching design of artificial intelligence course is shown in Figure 2. The above modules are not carried out mechanically in one direction. In the teaching process, teaching can be interspersed according to the mastery of students and technical updates, which are mutual influence and mutual penetration.

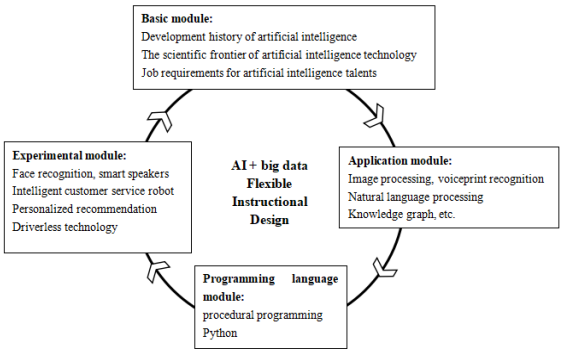


Figure 2. Four modules of artificial intelligence course teaching design.

4. Demonstration example of case teaching method of artificial intelligence course

The task-driven teaching mode is a type of education that meets the development requirements of vocational school students. According to the basic links of the task-driven teaching method, the development direction of artificial intelligence majors and industries, and the actual needs of talents, this paper is based on the second section of the above curriculum framework of How to observe with artificial intelligence in Observing Color: Facial Expression Recognition for case teaching demonstration. Students are having task-driven learning under the guidance of teachers (see Fig. 3).

Figure 3. Artificial intelligence course teaching.

First, create contextual situation and propose a task; Second, analyze tasks and collect knowledge; Third, task implementation and hands-on practice; Forth, evaluation, reflection, consolidation and improvement. The above case link of the "Facial Expression Recognition" course based on task-driven teaching is shown in Figure 4. The students who practiced the AI course also learned through the "SenseTime Education Platform". According to the statistics, from February 2022 to January 2023, the overall usage of the platform has gradually improved, with a total of 79,275 visits. A total of 960 students

were divided into 21 classes, and learned the AI course under the guidance of 19 teachers. Many students believe that the establishment of the AI course is very meaningful. First, the AI course can expand students' knowledge, and secondly, learning related courses can help students' future development and employment in the field of technology. In addition, the AI course focuses on cultivating students' innovation and practical abilities. Of course, there are also some challenges and issues mentioned in the students' evaluations. Some students think that the course requires more time and effort to learn and master. Moreover, the course usually requires a certain foundation in programming and mathematics. At the same time, some students also mentioned that the content of the AI course is relatively abstract and theoretical, and more practical cases and application scenarios are needed to deepen their understanding. Overall, the students' evaluation of the AI course is positive, and they believe that this is a forward-looking and practical course.

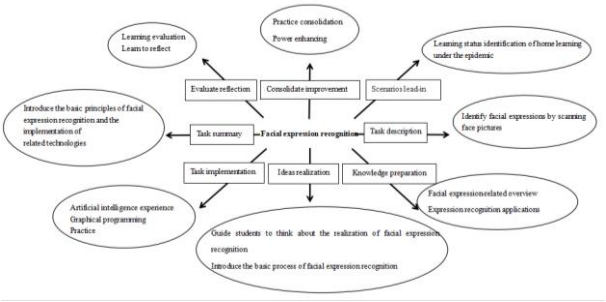


Figure 4. Course case of facial expression recognition.

5. Conclusion

The advent of artificial intelligence era has accelerated the frequency of scientific and technological updates and the speed of job demand adjustment. The general education of artificial intelligence in secondary vocational schools is imperative. This paper is based on the task-driven teaching method and the knowledge base of secondary vocational students. The details of the AI course, such as practical application cases, interactive laboratories, group discussions, guest speeches, and course assignments, can effectively stimulate students' curiosity increase their interest and enthusiasm for AI courses, and help them better understand and apply AI technology. According to the construction of the textbook General Knowledge of Artificial Intelligence developed with school-enterprise collaborative education model, a series of curriculum-related modes, curriculum ideas and content cases are discussed in this paper, which provides multi-dimensional and deep-level training support for the technical talents cultivation of artificial intelligence in secondary vocational schools. Future research can continue to deepen and improve artificial intelligence courses in secondary vocational schools based on the samples obtained from practical teaching and data analysis. With the student-oriented concept, the goal of effectively transforming curriculum teaching into vocational planning for secondary vocational students will be achieved.

Acknowledgements

This work was supported by the Research Projects of Guangdong Province Secondary Vocational Education Teaching Reform under Grant [GDZZJG2029]; and the National Industrial and Information Technology Vocational Education and Teaching Steering Committee project under Grant [GXHZWZ20507].

References

Ministry of Education. (2017). *Information Technology Curriculum Standards for Secondary Vocational Schools.* Retrieved from: http:// www.gov.cn /zhengce/content/2017/07/20/content_5211996.htm, 2017-07-20.

Yunfei, G., & Gang, L. (2022). Analysis on the Path of High-quality Publishing of Artificial Intelligence Course Textbooks. *Science and Technology and Publishing*, *5*, 149-153.

Youjia, L., Yihang, L., Yonggu, W., Yi, W., Jiarong, M., & Mengfang, H. (2022). Artificial Intelligence Basic Education Course Global Map: Development Framework and Implementation Strategy——Analysis of UNESCO's "K-12 Artificial Intelligence Curriculum Map" report. *Journal of Distance Education*, *40*(3), 3-15.

Jian, W. (2022). Exploration and thinking on general education in the new era. *Chinese University Teaching*, *4*, 9-13.

Zehui, X., Jiajing, Y., Qianyi, W., & Binggang, H. (2022). Design and application of performance evaluation in artificial intelligence courses. *Modern Educational Technology*, *32* (5), 32-41.

Jian, W. (2020). Cultivate the core literacy of information technology to create high-quality professional talents that meet the needs of the development of the times - secondary vocational schools Reflections on the Implementation of Information Technology Curriculum Standards. *China Vocational and Technical Education*, *17*, 19-23.

Zhixin, Z., Hui, D., Lu, G., & Kai, G. (2020). Artificial Intelligence in Primary and Secondary Schools in Developed Areas Current Situation, Problems and Countermeasures of Curriculum Construction——A Case Study of a "New First-tier" City. *China Electronic Education*, *9*, 40-49.

Xiaomin, L., & Ying, D. (2022). A phenomenological study on the learning experience of nursing students in secondary vocational colleges. *Zhonghua Nursing Education*, *19*(4), 293-297.

The writing group of Artificial Intelligence Concise Knowledge Reader. (2017). *Artificial Intelligence Concise Knowledge Book*. Beijing: Xinhua Publishing House.

Zhengli, X. (2021). Research on AI job demand analysis based on big data technology. *Guangxi Science*.*28*(3),321-329.

Dayuan, D.(2017). Structural Logic of Work Process Systematic Curriculum. *Education and Vocation*, *13*, 5-12.

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

Exploring the Role of Online EFL Learners' Academic Grit in Their Learning Engagement: A Structural Equation Model

Yuhan Guan[1], Yanqing Yi[1], Jinjin Liu[1], Feiyu Teng[1], Lin Luan[1*]
[1] Beijing University of Posts and Telecommunications
[*] luanlin@bupt.edu.cn

Abstract: *Online learning under the COVID-19 pandemic is gradually becoming an option that most universities use for emergencies. This study presents a structural relationship model that integrates English language learners' academic grit with their online learning engagement. Two questionnaires, Academic Grit (AG) and Online English Learning Engagement (OLLE), were developed and administered to 453 Chinese university students. The path analysis revealed that academic grit played a positive role in the different aspects of their online learning engagement. Nevertheless, none of the sub-dimensions of academic grit could predict emotional engagement. This study also provided implications for the research on AG and OLLE.*

Keywords: academic grit; online learning engagement; EFL learner; COVID-19; structural equation modelling (SEM)

1. Introduction

The COVID-19 pandemic has affected almost every industry. Education has also been greatly impacted by the pandemic. In response to the crisis, educational institutions have been forced to adopt virtual learning, so digital transformation is no longer a luxury, but a necessity (Leo, 2021). During the COVID-19 pandemic lockdown, the Chinese government has issued a policy known as *Suspending Classes without Stopping Learning* in order to ensure that teaching and learning will continue without disruptions. Therefore, it is vital to optimize online EFL learning during this pandemic lockdown (Luan et al., 2020).

2. Literature Review

2.1. Academic Grit

The notion of grit was originally conceptualized by Duckworth et al. (2007) as perseverance and passion for achieving long-term and higher-order goals. They state that grit consists of two basic constructs: consistency of interest and perseverance of effort. The former refers to an individual's consistent enthusiasm for achieving their higher-order goals despite failures, challenges, and obstacles, while the latter refers to a person's tendency to invest energy and effort over an extended period of time (Zarrinabadi, 2022).

2.2. Online English Learning Engagement

Student engagement is a multidimensional construct that involves behavioral, emotional, and cognitive dimensions (Fredricks et al., 2004). Engagement refers to a state of heightened attention and

involvement where participation is not only in the cognitive dimension, but also in social, behavioral and emotional dimensions (Philp and Duchesne, 2016). In this research, the four-component model proposed by Philp and Duchesne (2016), consisting of cognitive engagement, behavioral engagement, emotional engagement and social engagement, is adopted to represent online English learning engagement.

2.3. Academic Grit and Online Learning Engagement

Strong associations have been found between academic grit and learning engagement. Katherine et al. (2017) indicated that academic grit is positively associated with behavioral engagement among college school students. Although previous scholars have conducted extensive research on academic grit and learning engagement of L2 learners, the impact of academic grit on learning engagement has not been well studied, especially considering these two constructs in online learning environments. Therefore, this study aims to explore the intricate relationship between online EFL learners' academic grit and learning engagement.

3. Methodology

3.1. Research Context

The present study was conducted in an English course at the first author's university during the semester of the academic year of 2020-2021. During the pandemic lockdown, instructors taught online through course delivery tools (e.g., Tencent Classroom), video conferencing platforms (e.g., Tencent Meeting), and other social media (e.g., WeChat). A random sample of 453 (74.4% male) students, aged 18 to 23 years, were enrolled. As the participants were college students majoring in science and technology in a technological university with telecommunications as its focus, the male students outnumbered the female students in this study.

3.2. Instruments

In this study, we used two instruments: academic Grit (AG) and online learning engagement (OLE). A five-point Likert scale was used in all questionnaires from 1 (Strongly Disagreed) to 5 (Strongly Agreed), presented in Chinese, students' native language. Each dimension consists of three to five items. The AG survey was based on the instrument developed by Benishek et al. (2005). The survey included four dimensions: commitment; control of effort; control of affect; challenge. To measure students' online English learning engagement, 16 items were from the revised learning engagement scale developed by Luan et al. (2020). Four factors were included: cognitive engagement, behavioral engagement, emotional engagement and social engagement.

3.3. Research Procedure

First, this study conceptualized the two main research constructs with clearly-defined factors based on the precious research frameworks, and then proposed a hypothesized structural model concerning the relationship among all factors. Then, a structural equation modeling approach has been adopted to test the hypothetical model through confirmatory factor analysis (CFA) and path analysis. Finally, the complex inter-relations among all the factors of the two constructs were investigated. The SPSS 22.0 and AMOS 22.0 were employed to conduct the validity and reliability tests of the two instruments.

4. Results

4.1. CFA Analysis of the Academic Grit Survey and Online English Learning Engagement Survey

In order to verify the construct of the academic grit (AG) survey, confirmatory factors analysis was conducted. The results showed that all factor loadings were higher than the cut-off value of 0.50. All Average Variance Extracted values (AVE) had exceeded 0.60. The Composite Reliability values (CR) ranged from 0.86 to 0.92. Moreover, all alpha values were above 0.7 and the overall Cronbach's value was 0.92. Therefore, the reliability of the questionnaire was established. In addition, its fit statistics were as follows: χ^2/df =1.75, RMR=0.49, GFI=0.91, NFI=0.92, IFI=0.97, CFI=0.97, based on the Chi-square criterion and the fitting statistics of structural equation model, this survey had good structural validity.

Similar research method was applied to the measurement of online English learning engagement (OELE) survey. The results showed that all Average Variance Extracted values (AVE) of components of OELE had exceeded 0.60, the Composite Reliability values (CR) ranged from 0.86 to 0.90, all alpha values were above 0.7 and the overall Cronbach's value was 0.91. Moreover, χ^2/df =1.55, RMSEA=0.50, GFI=0.91, NFI=0.93, IFI=0.98, CFI=0.98. Statistics all indicated that OELE survey had a good reliability and structural validity.

4.2. Path Analysis

The path analysis was conducted to explore the relationship between learners' academic grit and their online English learning engagement. The final structure model is displayed in Fig. 1.

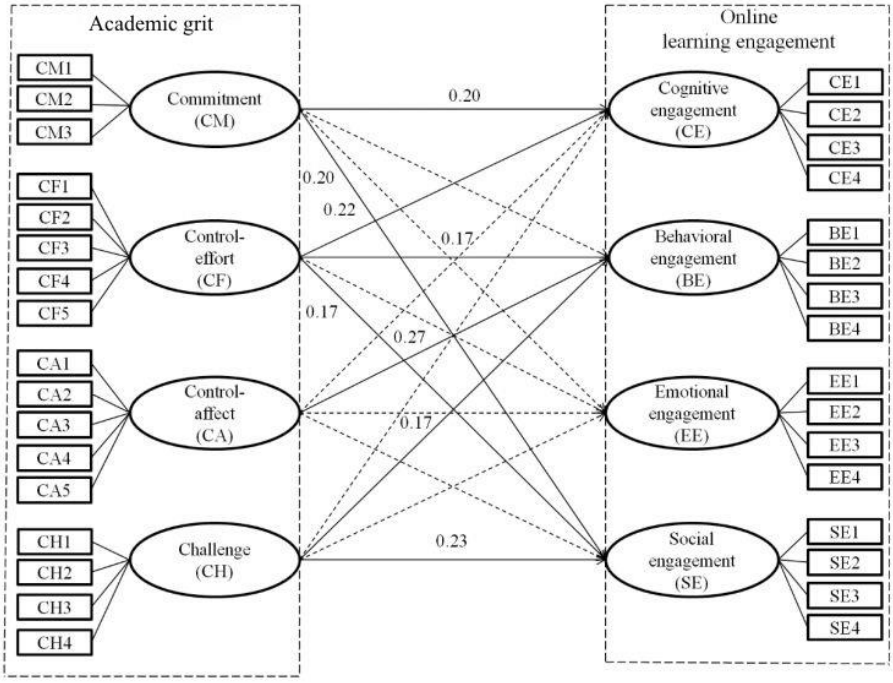


Figure 1. The final model of the structural relations between the AG and OLE.

First of all, the results of the path model testing revealed a good model fit with acceptable fitting indices (χ^2/df=1.81; CFI=0.90; TLI=0.89 IFI=0.90; RMSEA=0.06). Then a summary of the standardized path coefficients was analyzed, and the associated significance was indicated in solid lines. As shown in Fig.1, the factor "control-effort" is the most positive factor which significantly predicates three factors of online English learning engagement, with path coefficients ranging from 0.17 to 0.27, all the estimates

were statistically significant at $p<0.001$. Learners' challenge can also positively explain the variations in behavioral engagement ($\beta=0.17$, $p<0.001$) and social engagement ($\beta=0.23$, $p<0.001$). Meanwhile, the factor "commitment" has the positive relationships with cognitive engagement and social engagement. Control-affect is a significant factor for behavioral engagement. Surprisingly, academic grit failed to predict learners' emotional engagement.

5. Discussion and Conclusion

In this study, a proposed model of learners' academic grit and learning engagement in English courses was explored in the context of technology-enhanced environment, indicating the positive effects of academic grit to engagement and academic outcomes for university students, thus supporting the findings of Hodge et al. (2018). The following conclusions were drawn from this study: First, the results suggested that students' cognitive engagement can be promoted by commitment and control-effort. Many students need to learn to be gritty in order to persevere through the challenges and difficulties they face and achieve their long-term goals, and once they do, they will do better in school (Katherine, 2017). Nevertheless, none of the sub-dimensions of academic grit could predict emotional engagement. Second, the results indicated that control-effort, control-affect, and challenge are positive indicators of online behavioral engagement. A possible reason for this finding is that emotional engagement is more affected by contextual factors, such as teacher support and peer support (Luan et al., 2020). These results suggest that the higher the grit, the deeper online English learning engagement will be. Third, commitment, control-effort and challenge all contribute to online social engagement. Students can maintain their concentration and enthusiasm by doing correct attributions and working through the temptations and uncertainties of the learning process.

This study still has several limitations. First, the participants of this study were college students majored in science and technology at one university, so the representativeness of the results should be further verified with a larger sample within more diversified contexts. Second, as all data were collected from participants' self-reported survey, another reasonable step would be to employ multiple data analysis methods, such as learning analytics methods using data retrieved from the learning management system (Luan et al., 2020).

Acknowledgements

This research was supported by Graduate Student Teaching Reform Project supported by Beijing University of Posts and Telecommunications (2022Y013) and 2022 World Language and Culture Research Project (WYZL2022BJ0015).

References

Benishek, L. A., Feldman, J. M., Shipon, R. W., & Mecham, S. D. (2005). Development and evaluation of the revised academic hardiness scale. *Journal of Career Assessment*, 13(1), 59-76.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: perseverance and passion for long-term goals. *Journal of personality and social psychology*, *92*(6), 1087.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. Review of Educational Research, 74(1), 59–109.

Hodge & Wright & Bennett. (2018). The role of grit in determining engagement and academic outcomes for university students. *Research in Higher Education*, 59, 448-460.

Katherine, M., Allan, W., & Jiseung Y. (2017). How true is grit? Assessing its relations to high school and college students' personality characteristics, self-regulation, engagement, and achievement[J]. *Journal of Educational Psychology*, 109(5), 599-620.

Leo, S., Alsharari, N. M., Abbas, J., & Alshurideh, M. T. (2021). From offline to online learning: A qualitative study of challenges and opportunities as a response to the COVID-19 pandemic in the UAE higher education context. *The Effect of Coronavirus Disease (COVID-19) on Business Intelligence*, 203-217.

Liu, H., & Han, X. (2022). Exploring senior high school students' English academic resilience in the Chinese context. *Chinese Journal of Applied Linguistics*, *45*(1), 49-68.

Luan, L., Hong, J. C., Cao, M., Dong, Y., & Hou, X. (2020). Exploring the role of online EFL learners' perceived social support in their learning engagement: a structural equation model. *Interactive Learning Environments*, 1-12.

Philp, J., & Duchesne, S. (2016). Exploring engagement in tasks in the language classroom. *Annual Review of Applied Linguistics*, *36*, 50-72.

Zarrinabadi, N., Rezazadeh, M., & Mohammadzadeh Mohammadabadi, A. (2022). L2 grit and language mindsets as predictors of EFL learners' attitudes toward effectiveness and value of CALL. *Computer Assisted Language Learning*, 1-18.

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

Virtual Reality in Language Learning: A Review of Selected Journal Publications from 2018-2022 and Implications for Future Research

Yanqing Yi[1], Yuhan Guan[1], Feiyu Teng[1], Jinjin Liu[1], Lin Luan[1*]
[1] Beijing University of Posts and Telecommunications
[*] luanlin@bupt.edu.cn

Abstract: *The increasing popularity of Virtual Reality (VR) has provoked researchers' and educators' interest to investigate the impact of virtual environments on various fields of learning. However, scholars and educators lack a recent review of VR application in language learning from a research perspective. Thus this study reviews the scholarly literature on VR application in language teaching and learning published on high-impact journals between 2018 to 2022. The selected articles are analyzed under the following perspectives including the use of VR technologies, the research context and participants, target language, target language skills and research methods. Moreover, this study analyses VR as an emerging educational tool entailing both benefits and drawbacks and provides some profound suggestions and implications for future researchers and practitioners.*
Keywords: virtual reality, VR, language learning, literature review, VR-assisted language learning

1. Introduction

Virtual reality (VR) as an emerging and promising technology has attracted substantial attention in learning (Parmaxi, 2020). The incorporation of VR in the instruction allows learners to access simulated, immersive and authentic environments (Tai et al., 2020) that would be otherwise inaccessible in the real world. Most of the existing literature reviews have systematically described the VR application in language learning(Qiu et al., 2021), yet few papers reviewed the trends in VR technology-supporting language learning from a research perspective (Parmaxi, 2020) and some papers merely concerned parts of research issues (Wang et al., 2020). In order to better understand how VR is supporting language learning from research perspective in accordance with its rapid advancement, this paper will systematically capture the current state of VR studies in language learning published from 2018 to 2022 in top-ranked journals. The research questions addressed in this study are as follows:

(1) What the use of technologies, the research context and participants, target language, target language skills and research methods have been investigated in the selected articles?

(2) What are the benefits and limitations of using VR as an educational tool in language learning?

2. Methodology

Following the suggestion in the review studies by Fu and Hwang (2018) and Parmaxi and Zaphiris (2017), this research focuses on four high-impact journals in the field of technology-enhanced language learning, including *Computer Assisted Language Learning*, *ReCALL*, *Languages*, and *Language Learning and Technology*. The keywords used in the searching of the targeted four journal include: ("virtual reality" OR "virtual environment" OR "virtual world" OR "immersive environment" OR "virtual reality learning environment") AND ("language learning" OR "technology-enhanced language

learning" OR "language education" OR "language acquisition" OR "foreign language"), with a 5-year time-frame (2018-2022). The initial search yielded 32 manuscripts. By reviewing the articles in details, studies were eligible for inclusion in this research if they qualify the following criteria: a) focused on the application of VR in a language course; b) empirical studies; c) journal articles. Finally, the content of 16 articles was retained and analyzed here. To further explain the research dynamics of VR-assisted language learning, this study puts an emphasis on the key information of the research including the use of technologies, the research context and participants, target language, target language skills and research methods.

3. Results

As shown in Table 1, the use of VR technologies, the research context and participants, and target language are displayed to illustrate the basic information of the selected articles. The VR technologies used by the studies vary, in which Second Life is the most used VR device and other studies employ customised virtual environments or VR tools developed by Google EduVenture such as Google Expeditions and Google Cardboard. The language learning settings in which the VR experiment was conducted again varied. With regard to the participants' educational institution, the majority of studies focused on university students followed by middle school students and high school students, whilst limited studies involved adult users. In addition, most of the studies had English as their target language. Less interest has been received for Spanish, French and Japanese as a target language. Moreover, Figure 1 illustrates the target language skills being investigated by the selected articles, in which speaking, writing and motivation have attracted the most attention followed by vocabulary and overall linguistic competence, while less studies dealt with listening skill.

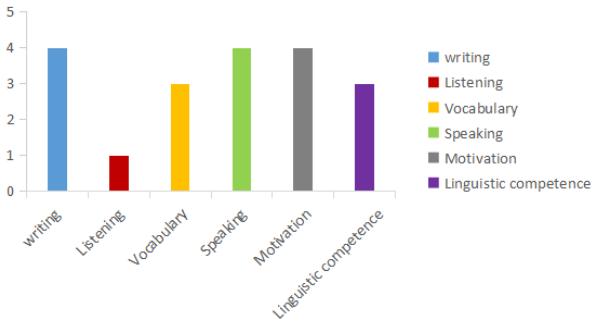


Figure 1. Target language skills.

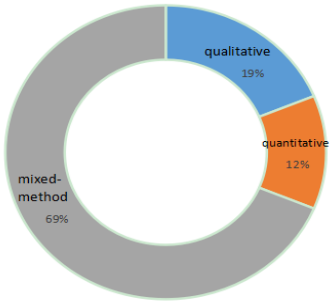


Figure 2. Research methods.

Table 1. Overview of research articles.

Author(s)	Learning device(s)	Context and participants	Target language
Chen (2018)	Second Life	10 adult EFL learners	English
Melchor-Couto (2018)	Second Life	18 university students	Spanish
Park (2018)	Second Life	5 enlisted soldiers	English
Yamazaki (2018)	Meet-Me (MM)	11 university students	Japanese
Legault et al. (2019)	VR headsets	64 undergraduate students	English
Nobrega & Rozenfeld (2019)	Google Cardboard/VR headsets	13 public school students	French
Chen et al. (2020)	Google Earth Virtual Reality	22 middle school students	English-Spanish
Ebadi & Ebadijalal (2020)	Google Expeditions	20 EFL learners	English
Tai et al. (2020)	Mondly VR app/Samsung Gear VR	49 seventh-graders	English
Lai & Chen (2021)	Oculus Go	30 high school students	English
Lin et al. (2021)	EduVenture VR	38 undergraduates	English
Mendes de Oliveira et al. (2021)	Babbel's virtual-classroom	52 adult users	English
Xie et al. (2021)	Google Cardboard and Google Expeditions	12 college students	English/Chinese
Ebadijalal & Yousofi (2022)	Google Expeditions	42 EFL learners	English
Hoang & Hoang (2022)	Google Docs	24 high school students	English
Tai (2022)	VR headsets	49 seventh-graders	English

As illustrated in Figure 2, among the selected empirical studies, mixed methods design is the mostly utilized. For example, some studies (e.g., Ebadijalal & Yousofi, 2022; Lai & Chen, 2021;Park, 2018; Tai et al., 2020) employed both tests, questionnaires and interviews to obtain a comprehensive understanding of the impact of VR-assisted language learning. Moreover, the qualitative methods (e.g., Ebadi & Ebadijalal, 2020; Nobrega & Rozenfeld, 2019) and quantitative methods (e.g., Chen, 2018; Legault et al., 2019) were also employed to assess and analyze the VR application in language learning.

Table 2 indicates that VR, an emerging educational tool in language learning, comes with benefits and drawbacks. Previous empirical studies have claimed that virtual environment is beneficial for language learning (Legault et al., 2019), but it still fails to reach for a large application due to its high-cost and less accessibility (Qiu et al., 2021).

Table 2. VR-supporting language learning approach.

Strengths	**Weakness**
• It can provide multimodal-supporting and near-real-life language learning environments. • It supports highly interactive learning scenarios with visual, auditory and tactile sensations. • It helps to transfer language acquisition from virtual to reality. • It motivates learners' self-expression and encourages them to communicate with virtual characters.	• VR's technical cost problem poses challenges to its application in the classroom. • Long-time use of VR devices may affect learners' physical and mental health.
Opportunities	**Threats**
• VR can be used to enhance the overall linguistic competences such as speaking and listening skills. • The continuous educational reform requires emerging technologies such as VR to improve the learning environments. • It can record learners' learning behaviors and learning process for future analysis and research. • It can provide learners non-verbal communication especially when learning is conducted online.	• Task designs remains as a great challenge in the process of VR-assisted language learning. • Higher demand for network stability and higher standard for VR devices. • Learners may be addictive to the virtual environments and fail to have serious learning attitude.

4. Discussion and Conclusion

This study reviews the recent VR-supported language learning research published on four top-tier language journals between 2018 to 2022 from multiple dimensions. By analyzing the use of VR

technologies, the research context and participants, target language, target language skills and research methods, we can affirm VR's high application potential and value in language learning (Qiu et al., 2021). According to the research findings, VR studies in language learning mainly focused on writing and speaking skills and motivation in recent years, leaving the potential of listening and vocabulary to be insufficiently investigated. In addition, the results of analyzing the research methods indicated that the majority of studies employed mixed methods to collect both quantitative and qualitative data, in which tests, questionnaires and interviews were most frequently used for examining the impact of the application of VR in language learning. This study also has several limitations as well as important pedagogical implications. First, as most of the selected studies focused on general language skills such as writing and speaking, future VR research should embrace more minor language aspects. Moreover, future VR studies can also employ multimodal methods especially the quantitative methods to trace the long-term impact of VR-assisted language learning.

Acknowledgements

This research was supported by Graduate Student Teaching Reform Project supported by Beijing University of Posts and Telecommunications (2022Y013) and Teaching Reform Project of Intercultural Communication Course supported by School of Humanities, Beijing University of Posts and Telecommunications.

References

Chen, J. C. (2018). The interplay of tasks, strategies and negotiations in Second Life. *Computer Assisted Language Learning*, *31*(8), 960-986.

Chen, Y., Smith, T. J., York, C. S., & Mayall, H. J. (2020). Google Earth Virtual Reality and expository writing for young English Learners from a Funds of Knowledge perspective. *Computer Assisted Language Learning*, *33*(1-2), 1-25.

Chen, B., Wang, Y., & Wang, L. (2022). The effects of virtual reality-assisted language learning: A meta-analysis. *Sustainability*, *14*(6), 3147.

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.

Ebadijalal, M., & Yousofi, N. (2022). 'Take me to a virtual trip if you want me to write better!': the impact of Google Expeditions on EFL learners' writing motivation and performance. *Computer Assisted Language Learning*, 1-23.

Fu, Q. K., & Hwang, G. J. (2018). Trends in mobile technology-supported collaborative learning: A systematic review of journal publications from 2007 to 2016. *Computers & Education*, *119*, 129-143.

Hoang, D. T. N., & Hoang, T. (2022). Enhancing EFL students' academic writing skills in online learning via Google Docs-based collaboration: A mixed-methods study. *Computer Assisted Language Learning*, 1-23.

Lai, K. W. K., & Chen, H. J. H. (2021). A comparative study on the effects of a VR and PC visual novel game on vocabulary learning. *Computer Assisted Language Learning*, 1-34.

Legault, J., Zhao, J., Chi, Y. A., Chen, W., Klippel, A., & Li, P. (2019). Immersive virtual reality as an effective tool for second language vocabulary learning. *Languages*, *4*(1), 13.

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.

Melchor-Couto, S. (2018). Virtual world anonymity and foreign language oral interaction. *ReCALL*, *30*(2), 232-249.

Mendes de Oliveira, M., Sporn, Z., Kliemann, L., Borschke, A., & Meyering, M. (2021). Online language learning and workplace communication: a study on Babbel's virtual-classroom solution. *Computer Assisted Language Learning*, 1-26.

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.

Park, M. (2018). Innovative assessment of aviation English in a virtual world: Windows into cognitive and metacognitive strategies. *ReCALL*, *30*(2), 196-213.

Parmaxi, A., & Zaphiris, P. (2017). Web 2.0 in computer-assisted language learning: A research synthesis and implications for instructional design and educational practice. *Interactive Learning Environments*, *25*(6), 704–716.

Qiu, X. Y., Chiu, C. K., Zhao, L. L., Sun, C. F., & Chen, S. J. (2021). Trends in VR/AR technology-supporting language learning from 2008 to 2019: A research perspective. *Interactive Learning Environments*, 1-24.

Tai, T. Y. (2022). Impact of mobile virtual reality on EFL learners' listening comprehension. *Language Learning & Technology*, *26*(1).

Tai, T. Y., Chen, H. H. J., & Todd, G. (2022). The impact of a virtual reality app on adolescent EFL learners' vocabulary learning. *Computer Assisted Language Learning*, *35*(4), 892-917.

Wang, C. P., Lan, Y. J., Tseng, W. T., Lin, Y. T. R., & Gupta, K. C. L. (2020). On the effects of 3D virtual worlds in language learning–a meta-analysis. *Computer Assisted Language Learning*, *33*(8), 891-915.

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.

Yamazaki, K. (2018). Computer-assisted learning of communication (CALC): A case study of Japanese learning in a 3D virtual world. *ReCALL*, *30*(2), 214.

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

Virtual Reality for English Vocabulary Learning in Higher Education: An Exploratory Study

Yanqing Yi[1], Yuhan Guan[1], Feiyu Teng[1], Jinjin Liu[1], Lin Luan[1*]
[1] Beijing University of Posts and Telecommunications
[*] luanlin@bupt.edu.cn

Abstract: *Virtual Reality (VR) technology provides learners with authentic, immersive and interactive learning environments and emerges as a promising tool for language education. Previous research preferred to employ VR to investigate language competence, leaving the potential of integrating VR into college EFL learners' vocabulary learning yet to be insufficiently investigated. Therefore, this study conducted a quasi-experiment to examine the impact of a self-developed 3D Virtual Environment (3DVE) on EFL learners' vocabulary learning. Sixty-four college students participated in the study and were randomly assigned to either an experimental (VR players) or control (video watchers) group. The results showed that the VR players' vocabulary performance was significantly better than the video watchers'. The 3DVE contextualized vocabulary learning by allowing students to access various simulated real life experiences, comprehensible input and meaningful interaction through multimodal support, and the VR players' learning motivation still needed to be enhanced. Most VR players felt that VR-assisted vocabulary learning was beneficial and enjoyable, but the platform still needs to be optimized in terms of task design and technical issues.*

Keywords: virtual reality, EFL vocabulary learning, college student, VR-assisted language learning, situated learning

1. Introduction

Due to the outbreak of Covid-19, a growing number of learning activities are shifting to online settings, which presents a formidable challenge to many disciplines especially language education (Song et al., 2022). English, an international language, is a compulsory course in China. When learning a language, vocabulary is paramount and is regarded as the core of overall language competence (Yu & Trainin, 2022). However, Chinese English as a foreign language (EFL) learners have often been struggled with vocabulary learning (Alfadil, 2020) for the lack of authentic learning contexts. With the advancement of technology, VR as one of the emerging technologies, has gained popularity in language learning (Tai et al., 2022) for its ability to build immersive and authentic environments. However, previous research had conflicting results regarding the impact of VR-assisted vocabulary learning (Cheng et al., 2017). Therefore, the present study aims to explore whether VR technology is beneficial to facilitating students' vocabulary learning.

5. Literature Review

2.1. VR in EFL Education

VR is "a highly interactive, computer-based, multimedia environment in which the user becomes a participant with the computer in a 'virtually real' world" (Pantelidis, 1993). Integrating VR into language learning has increasingly attracted researchers' and educators' attention in recent years. Its immersive language learning environment allows learners to assume a first-person perspective in 3D settings (Song et al., 2022). VR also provides highly interactive learning scenarios with auditory, visual, smell, taste and tactile sensations, in which learners can communicate with virtual characters using the targeted language (Yeh et al., 2020). The aforementioned studies claimed that VR is a beneficial tool to facilitate language learning in various aspects (Tai et al., 2020).

2.2. VR-assisted Vocabulary Learning

Vocabulary is vitally important in acquiring a foreign or second language (Yu & Trainin, 2022). Previous research has found that VR could be employed to facilitate learners' vocabulary learning(Tai et al., 2020). For example, Alfadil (2020) explored the impact of VR mediation on vocabulary learning. The results showed that the VR players outperformed their counterparts in vocabulary learning, and have a positive attitude towards the VR-assisted language learning approach. However, most VR studies on vocabulary learning have used highly immersive VR through head-mounted displays (HMDs) and the use of desktop VR, a more accessible and cost-effective (Song et al., 2022) VR solution compared with HMDs, remains in its infancy. This being the case, this study aims to investigate the impact of desktop VR on Chinese EFL learners' vocabulary learning and their perceptions towards desktop VR-assisted vocabulary learning. The research questions addressed in this study are:

(1) Does the VR environment significantly improve the EFL learners' vocabulary learning?

(2) Does the VR environment provide different benefits to the EFL learners' vocabulary learning? If so, what VR features contribute to the effectiveness of the vocabulary learning?

(3) What are the EFL learners' perceptions towards the VR environment for their vocabulary learning?

3. Method

3.1. Research Context and Participants

This study was conducted in a College English course at a comprehensive university in Chinese mainland. A total of 64 second-year undergraduate students (38 males and 26 females) participated in the study and were randomly assigned to either the experimental (VR players) or the control (video watchers) group, and their English ability were relatively the same with each other. The VR players did learning activities in the VR platform, while the video watchers watched the same instructional videos on the personal computer.

3.2. Instruments

3.2.1. Situational English in Virtual Reality (SEVR)

This study was conducted based on a self-developed *Situational English in Virtual Reality* platform *(SEVR)*. By simulating major scenarios at international airports, the platform was embedded with seven task modes in order to satisfy learners' different learning demands.

3.2.2. Vocabulary Tests

Vocabulary tests consisted of two sections, one section with 10 questions was the compound dictation and the other section with 20 questions was the translation from Chinese to English, and the

content of vocabulary pretest and post-test were the same. The total score of the vocabulary test was 30 points with each correct answer awarded 1 point.

3.2.3. VR-assisted Vocabulary Learning Questionnaire

The VR-assisted vocabulary learning (VRVL) questionnaire was adapted from Tai et al. (2020) including six dimensions, with a total of 12 close-ended items. All the questionnaire items were presented in participants' native language, Chinese, on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).

3.2.4. Semi-structured Interview

Semi-structured interviews were conducted with VR players to collect qualitative data about their perceptions towards VR-assisted vocabulary learning, and t***he framework of the interview was adapted from*** Chien et al. (2020).

3.3. Research procedure

The experiment lasted for six weeks, with two 45-minute classes each week. A background information questionnaire and a vocabulary pretest were administered to all participants in the first week. Then, all participants were trained to the use of cloud platform, moreover the VR players were also introduced to the operation of VR platform. From week three to week five, the VR players operated on the VR platform, while the video watchers watched the same instructional videos without the VR elements. In week six, all participants were required to finish a vocabulary post-test and a questionnaire. Moreover, semi-structured interviews were conducted to the VR players.

3.4. Data Analysis

Firstly, ANCOVA analysis was performed to investigate the impact of two different learning styles on learners' vocabulary performance. Secondly, descriptive data were calculated to illustrate the VR players' responses to the VRVL questionnaire.Thirdly, a qualitative analysis was conducted concerning the interview data.

4. Results

4.1. Vocabulary tests

One-way analysis of covariance (ANCOVA) was performed on data from students' pretest and post-test scores of vocabulary learning questionnaires to explore the differences between VR players and video watchers. As illustrated in Table1, the results suggested a significant difference in vocabulary scores between the two groups, demonstrating that the VR players (M = 20.875) outperformed the video watchers (M = 14.094) in vocabulary performance.

Table 1. Results of ANCOVA for the VR players and video watchers' vocabulary tests.

Group	*N*	*M*	*SD*	*F*	*p*
VR	32	20.875	4.412	45.842	0.000
Video	32	14.094	4.268		

4.2. VRVL questionnaire results

To investigate how the VR experiment facilitated learners' vocabulary learning, this study conducted a vocabulary post-test questionnaire among the VR players. In general, the majority of the VR players

positively perceived the VR equipment for vocabulary learning. Specifically, the VR players had the highest scores in the simulation dimension (*M* = 4.141; *SD* = 0.795), as they thought that the 3D environment was beneficial to learning vocabulary. However, they had the lowest scores in the motivation dimension (*M* = 3.594; *SD* = 0.901), which still needed to be enhanced.

Table 2. Results of the VRVL questionnaire

Dimensions, descriptive data and items (selected)	*M*	*SD*
Simulaton (SI)： *M* = 4.141； *SD* = 0.795		
I feel the virtual environment is a “physical ” world.	4.063	0.840
Immersion (IM)： *M* = 3.922； *SD* = 0.645		
The immersive environment is beneficial for me to focus on vocabulary learning.	4.031	0.647
Interactivity (IN)： *M* = 4.016； *SD* = 0.634		
Interacting with virtual characters helps me learn vocabulary.	4.031	0.595
Presence (PR)： *M* = 4.032； *SD* = 0.621		
Interacting with 3D objects from multiple angles really helps me learn vocabulary.	4.063	0.619
Experience (EX)： *M* = 3.969； *SD* = 0.736		
I would like to navigate in the virtual world.	3.844	0.884
Motivation (MO)： *M* = 3.594； *SD* = 0.901		
I am interested in the virtual experiment.	3.875	0.751

4.3. Results of Interview

The most prominent finding regarding the VR players’ perceptions towards the desktop VR was that the immersive and interactive virtual reality learning environments facilitated their vocabulary learning. For instance, student A02 claimed, “I immersed myself in the virtual environment and appreciated to interact with virtual airport officials.” Moreover, they also have some suggestions that the tasks should be upgraded to meet individualized academic needs, and some minor technical problems such as network disability need to be addressed to avoid experimental data losses.

5. Discussion and Conclusion

This study investigated the impact of desktop VR on college EFL learners’ vocabulary learning. It supported the findings of Nicolaidou et al.’s (2021) study, indicating the positive impact of immersive, authentic and interactive learning environments provided by VR technology on college students’ vocabulary learning. The research also corroborated with Alfadil’s (2020) study demonstrating the beneficial effects of VR game on EFL learners’ vocabulary acquisition. Moreover, students enjoyed the VR approach but their learning motivation still needed to be enhanced.

This study also has several limitations. Firstly, this study combined a small sample size, therefore future studies are recommended to include larger sample in diverse contexts. Secondly, the task design

and the VR platform should be upgraded and optimized to provide players with better VR-assisted language learning experience.

Acknowledgements

This research was supported by Graduate Student Teaching Reform Project supported by Beijing University of Posts and Telecommunications (2022Y013).

References

Alfadil, M. (2020). Effectiveness of virtual reality game in foreign language vocabulary acquisition. *Computers & Education*, *153*, 103893.

Cheng, A., Yang, L., & Andersen, E. (2017, May). Teaching language and culture with a virtual reality game. In *Proceedings of the 2017 CHI conference on human factors in computing systems* (pp. 541-549).

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.

Nicolaidou, I., Pissas, P., & Boglou, D. (2021). Comparing immersive Virtual Reality to mobile applications in foreign language learning in higher education: a quasi-experiment. *Interactive Learning Environments*, 1-15.

Pantelidis, V. S. (1993). Virtual reality in the classroom. *Educational technology*, *33*(4), 23-27.

Song, Y., Wen, Y., Yang, Y., & Cao, J. (2022). Developing a 'Virtual Go mode' on a mobile app to enhance primary students' vocabulary learning engagement: an exploratory study. *Innovation in Language Learning and Teaching*, 1-10.

Tai, T. Y., Chen, H. H. J., & Todd, G. (2020). The impact of a virtual reality app on adolescent EFL learners' vocabulary learning. *Computer Assisted Language Learning*, 1-26.

Yu, A., & Trainin, G. (2022). A meta-analysis examining technology-assisted L2 vocabulary learning. *ReCALL*, *34*(2), 235-252.

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

A Pedagogical Innovation of Using Scratch Animations for Learning Chinese Language Creative Writing and Learning Coding Skills for Computational Thinking Development in Primary Schools

Siu Cheung KONG [1 *], Wai Ying KWOK [2]

[1] Department of Mathematics and Information Technology, The Education University of Hong Kong, Hong Kong

[2] Centre for Learning, Teaching and Technology, The Education University of Hong Kong, Hong Kong

[*] sckong@eduhk.hk

Abstract: *This study innovated the pedagogical use of Scratch animations to support senior primary students in Hong Kong to learn and compose creative writings in Chinese Language classrooms. The pedagogy "To Play, To Learn, To Reflect, To Create, To Code" was designed to expose students to the process of first watching the Scratch animation in the theme of "Natural World"; then learning the building blocks and writing skills for describing plants, animals and landscape; then reflecting on the how-to-write and how-to-organize knowledge for creative writing in the descriptive writing genre; and finally creating a thematic descriptive essay and coding Scratch narration recording for an authentic application of the how-to-write and how-to-organize knowledge to describe signature characteristics of and express personal feelings toward the selected scenery. The pedagogical innovation can contribute to the design of technology-supported pedagogy which nurtures observant, reflective, and expressive learners who are able to articulately communicate their own ideas and feelings through Chinese Language writing compositions; and at the same time motivates students to learn coding for computational thinking development in Chinese Language writing classrooms.*

Keywords: Scratch, animation, Chinese Language, creative writing, primary schools

1. Introduction and Background of Study

Digital technology has deeply penetrated in our daily lives in the 21st century. School education is putting efforts to innovate technology-supported pedagogies to nurture students to be reflective learners who are active creators of technology (Education Bureau, 2015; Sharples, 2019). In Hong Kong, "Creative Writing" in Chinese Language subject is a major curriculum component for nurturing students as young in primary schools to become observant, reflective, and expressive to communicate their own ideas and feelings (Chu & Chow, 2017; Curriculum Development Council, 2017). This study innovated a pedagogy to use and code Scratch animations for learning Chinese Language creative writing.

Sharples (2019) illustrates 40 practical technology-supported pedagogies, and advocates to engage students in reflective learning through learning from animations, of which students watch and interact with short animations related to subject learning for exploring the subject topics and drawing on experience to gain deeper understanding. Animations – typically in a short cartoon video form – are multimedia resources commonly used in school education, especially for primary school students who find difficult to use short-term memory to process rich information in the learning process (Hanif, 2020;

Ploetzner & Lowe, 2012). The educational use of animations mainly serves to introduce key concepts and stimulate learning interest in a specific issue related to the target subject topic (Ploetzner & Lowe, 2012; Sharples, 2019).

In primary education, Scratch animations are popularly used in digital storytelling activities in language subjects – Scratch is a block-based programming environment with an intuitive interface-design for children to make simple actions to drag, drop, and combine code-blocks for an easy creation of various programs like animations, with an immediate observation of programming outcomes. Scratch is thus popularly used in primary schools to support young students to learn coding and develop computational thinking (Parsazadeh et al., 2021; Sarasa-Cabezuelo, 2019). It is naturally fit to integrate the use of Scratch animations in creative writing classrooms, in which students link the process of developing scripts and staging sprites in Scratch animation-creation with the process of describing characters and making narrations in creative writing-composition (Bull et al., 2017; Parsazadeh et al., 2021).

2. The Scratch-supported Pedagogy for Chinese Language Creative Writing Classrooms

The topic "Creative Writing" is a major component in primary Chinese Language curriculum in Hong Kong. With the goal of preparing students to become observant and self-expressive in day-to-day communication, this topic sets to enable students to learn and apply the building blocks – including vocabularies (e.g., coloring words, onomatopoeia, etc.) and transitional sentences; as well as the writing skills – including description methods (e.g., overall and partial description, sensory description, describe in chronological order, walking description, etc.) and rhetorical devices (e.g., simile, metaphor, metonymy, personification, parallelism, etc.) for describing scenery and objects in descriptive writing (Chu & Chow, 2017; Curriculum Development Council, 2017). This study took the core elements of the "To Play, To Learn, To Code" pedagogy for coding education (Kong, Lai, & Sun, 2020) to design a five-step pedagogy "To Play, To Learn, To Reflect, To Create, To Code" for Chinese Language creative writing classrooms. Supplementary with a Scratch animation and ten guided activity worksheets, this pedagogy guides students to "play" the Scratch animation to motivate them to write; then to "learn" descriptive writing by working with the details of the Scratch animation; then to "reflect" on the how-to-write and how-to-organize knowledge to pave the way to "create" descriptive writing; as well as to "code" Scratch animation as a bonus of learning coding in the writing classrooms. Computational thinking development in this pedagogy focuses on four concepts of "sequence", "repetition", "conditionals" and "parallelism" that are commonly used in coding.

This study – unlike the past research which briefly highlights the pedagogical needs and issues – sets to detail the steps, aims and activities of the pedagogical design which innovatively arranges students to first learn subject knowledge through Scratch animation and finally develop computational thinking through Scratch coding in general classrooms. Students at the beginning get familiar with learning with Scratch animation; and get curious about how to produce such the Scratch animation for learning. This arrangement thus exposes students to a subject learning process in line with the rationale of coding education – to first understand the problem (in this case, to learn with Scratch animation) and then think of and learn about coding solutions (in this case, to produce Scratch animation for learning) to solve the problem.

2.1. "To Play" through Watching Scratch Animation with Guided Activity Worksheets

The first "To Play" step aims to stimulate students' motivation in descriptive writing. Students watch the Scratch animation (with two guided activity worksheets) in the theme of "Natural World" to observe the images of plants, animals, and scenery while listening to the narration and sound in the Scratch animation (see Figure 1) for an initial sense of the building blocks of coloring words and onomatopoeia; and the writing skills such as overall and partial description.

Figure 1. Screens of the Scratch animation for "To Play" step to introduce vocabularies and description methods.

2.2. "To Learn" Building Blocks and Writing Skills from Scratch Animation and Guided Activity Worksheets

The second "To Learn" step is important for equipping students with the how-to-write and how-to-organize knowledge for descriptive writing. With the support from the Scratch animation and two guided activity worksheets, students in this step learn how to describe flowers using suitable coloring words and rhetorical devices; how to describe sounds of animals using suitable onomatopoeia; and how to write transitional sentences for structuring the connection between paragraphs describing the view-zooming from a mountain to particular trees inside (see Figure 2).

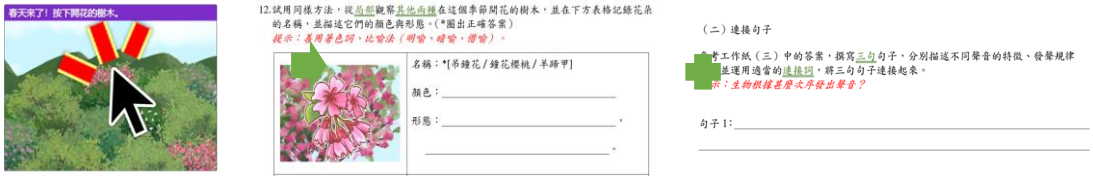


Figure 2. A worksheet activity in "To Learn" step for describing flowers in the Scratch animation.

2.3. "To Reflect" on Building Blocks and Writing Skills for Descriptive Writing

The third "To Reflect" step – supplementary with two guided activity worksheets – promotes students to reflect on the linkage between the experience and knowledge gained from the "To Play" and "To Learn" steps. Students in this step are provided with a summary of the building blocks and writing skills for their revision purposes (see Figure 3).

Figure 3. A worksheet activity in "To Reflect" step for reflecting on the use of description methods and rhetorical devices.

2.4. "To Create" a Descriptive Essay after Essay Appreciation

The fourth "To Create" step asks students to complete two guided activity worksheets for essay appreciation and creative writing (see Figure 4). For essay appreciation, students read and analyze a

sample essay “Falling in love with my housing estate” and a classic descriptive article “Spring” by Zhu Ziqing to consolidate their building blocks and writing skills. For creative writing, students think of any unforgettable scene (such as seaside of Sai Kung District) and apply all knowledge gained from the “To Play”, “To Learn” and “To Reflect” steps to create a descriptive essay.

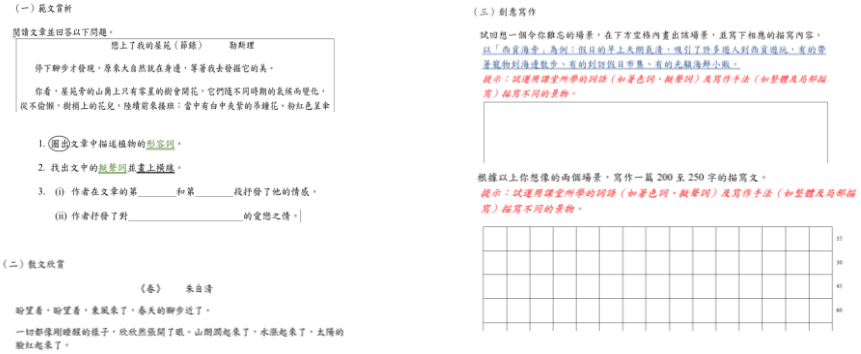


Figure 4. Worksheet activities for essay appreciation and creative writing in the “To Create” step.

2.5. “To Code” for Sprite Movement and Narration Recording for Scratch Animations

The fifth “To Code” step asks students to code a Scratch template, with two guided activity worksheets, to control the sprite movement and add the narration recording in the thematic Scratch animation (see Figure 5). Students code the given “Grey Bird” sprite for enabling the sprite to repeat the movement of flapping wings as well as change location and costumes. Students also add an audio-recording of their own narrations to the thematic Scratch animation under the Events code-blocks for sending-and-receiving messages for audio broadcast. The reason for this step is that students get familiar with Scratch animation after the process of learning descriptive writing; and get curious to know how to design such Scratch animation. Students are thus ready to learn skills to code Scratch animation, such as using code-blocks of “when I receive message” and “broadcast message” for coding a Scratch Event to make different bird-sprites sing simultaneously. In this coding process, students also learn computational thinking concepts such as sequence of codes, repetition of statements, conditionals for costume-change, and parallelism of sound-broadcasting for different sprites.

Figure 5. A worksheet activity in the “To Code” step for coding sprite movement and adding narration recording.

3. Conclusion and Future Work

The pedagogical innovation designed in this study uses Scratch animations to support Hong Kong senior primary students to learn and compose Chinese Language creative writing. The “To Play, To Learn, To Reflect, To Create, To Code” pedagogy exposes students to first watch the Scratch animation in the theme of “Natural World”; then learn about and reflect on the building blocks and writing skills for the thematic creative writing; and finally create a descriptive essay and code a Scratch template for an authentic application of knowledge about Chinese Language creative writing.

The pedagogical innovation will be trialed in Grade 5 Chinese Language creative writing classrooms in the selected Hong Kong primary schools, in the form of a 15-lesson teaching lasted for 525 minutes.

A quasi-experimental research will be conducted for a mixed-method evaluation of the pedagogical innovation through attainment tests, questionnaire surveys and focus group interviews with students. The trial teaching experience and outcome will be reported later for the contributions, from the subject-specific perspective, to the implementation and evaluation of an effective pedagogy that naturally integrates the use of Scratch animations into Chinese Language creative writing lessons in primary schools.

References

Bull, G., Schmidt-Crawford, D. A., McKenna, M. C., & Cohoon, J. (2017). Storymaking: Combining making and storytelling in a school makerspace. *Theory into Practice*, *56*(4), 271-281.

Chu, W. L. C., & Chow, K. C. (2017). Using video dubbing activities in primary Chinese classroom for creative writing: A case study of students with multilingual background. *L1-Educational Studies in Language and Literature*, *17*(3), 1-32.

Curriculum Development Council. (2017). *Chinese Language Education Key Learning Area Curriculum Guide (Primary 1 - Secondary 6)*. Hong Kong: Curriculum Development Council.

Education Bureau. (2015). *Realising IT potential ・Unleashing learning power: A holistic approach*. Hong Kong: Education Bureau.

Hanif, M. (2020). The development and effectiveness of motion graphic animation videos to improve primary school students' sciences learning outcomes. *International Journal of Instruction*, *13*(4), 247-266.

Kong, S. C., Lai, M., & Sun, D. (2020). Teacher development in computational thinking: Design and learning outcomes of programming concepts, practices and pedagogy. *Computers and Education*, *151*, 103872.

Ploetzner, R., & Lowe, R. (2012). A systematic characterisation of expository animations. *Computers in Human Behavior*, *28*, 781-794.

Parsazadeh, N., Cheng, P.-Y., Wu, T.-T., & Huang, Y.-M. (2021). Integrating computational thinking concept into digital storytelling to improve learners' motivation and performance. *Journal of Educational Computing Research*, *59*(3), 470-495.

Sarasa-Cabezuelo, A. (2019). Use of Scratch for the teaching of second languages. *International Journal of Emerging Technologies in Learning*, *14*(21), 80-95.

Sharples, M. (2019). *Practical pedagogy: 40 new ways to teach and learn*. Abingdon, Oxon; New York, NY: Routledge.

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

How to Measure Graduate Education Students' Dispositions on Statistical Literacy in Online Learning: A Pilot Study

Stephanie Cho[1], Jingshun Zhang[1], Shahid Rasool[1], Charles Xiaoxue Wang[1], Clarisse Halpern[1], and Carla Huck[1]

[1]Florida Gulf Coast University, USA

scho0435@eagle.fgcu.edu

Abstract: *Many graduate education students face challenges in mastering statistical literacy when they take statistics courses online. To develop effectively facilitate learning of statistics courses online and use appropriate teaching strategies, it is necessary to consider students' dispositional elements, such as motivation, anxiety, attitudes, and beliefs toward statistics, as well as knowledge elements. However, research has not been sufficiently conducted to investigate the relationships between statistical literacy and various dispositional variables. Literature also indicates that instruments that comprehensively evaluate these relationships are lacking. Therefore, the researchers developed a survey to measure graduate education students' dispositions on statistical literacy in online learning to investigate their relationships. This paper reports the process of instrument design and the results of the pilot study. Findings will help faculty to better understand students' dispositions related to mastering statistical literacy and to inform faculty of effective teaching strategies.*

Keywords: instrument design, statistical literacy, dispositions, online learning, graduate education students

1. Introduction

Graduate education students are expected to master statistical literacy in their program of studies and to pursue their careers as education leaders and researchers (Ritzhaupt et al., 2020; Thompson et al., 2019). However, graduate education students experience challenges when taking statistics courses and have even more difficulties when these courses are taught online (DeVaney, 2016; Ritzhaupt et al., 2020). Students' dispositions, such as motivation, anxiety, attitudes, and beliefs, are important factors influencing teaching and learning strategies that improve students' learning experience and satisfaction. However, there have not been sufficient studies on the effects of graduate education students' dispositions on their statistical literacy when they take statistics courses online. To fill the gap in the literature, this study is set to investigate the relationships between graduate education students' dispositions (motivation, anxiety, attitudes, and beliefs) and statistical literacy in online learning. The research aims to answer the following question: What are the relationships between graduate education students' dispositions (motivation, anxiety, attitudes, and beliefs) and statistical literacy in online learning?

2. Literature Review

Over the past decade, online learning has become increasingly important in higher education (Dumford & Miller, 2018; Seaman et al., 2018). Although online learning has many advantages, such as being accessible, flexible, and student-centered, it has several disadvantages, such as giving many

students a sense of isolation due to the lack of interaction, lack of community, demotivation, and technical difficulties (Bates, 2019; Holmes, & Reid, 2017). Online courses previously tended to be offered as an alternative to traditional courses, but with the COVID-19 pandemic, they have become a necessity.

Statistical literacy is the ability to read, understand, interpret, and communicate statistical information (Gal, 2002; Ziegler, 2014). In a data-driven society, statistical literacy is essential for critical thinking and decision-making and is a core competency in education (Gal, 2002; Johannssen et al., 2021). Knowledge and disposition are two critical factors in mastering statistical literacy (Gal, 2002). Therefore, it is necessary to consider students' dispositional elements, such as motivation, anxiety, attitudes, and beliefs, as well as knowledge elements to develop effective teaching strategies for statistics courses (Gal, 2002; Sharma, 2017).

Many graduate programs in education offer statistics courses in both traditional and online formats, and many graduate students experience challenges in completing these courses (Thompson et al., 2019). In particular, when statistics courses are taught online, graduate education students experience significantly more challenges than when provided in a face-to-face format. Nevertheless, research on how to teach graduate education students in the online format is lacking (Ritzhaupt et al., 2020). Online learning's disadvantages could become potential problems that affect students' performance and satisfaction, which ultimately impacts their graduation rate (DeVaney, 2016; Onwuegbuzie, 2004).

Many instruments have been developed to measure students' statistical comprehension and literacy in statistics courses (Sabbag et al., 2018). Similarly, multiple instruments are available to measure students' anxiety or attitudes toward statistics (Cruise et al., 1985; Schau et al., 1995; Wise, 1985). However, based on the broad systematic literature review, studies comprehensively evaluated the relationships between statistical literacy and various dispositional variables and validated instruments are rarely found.

3. Research Methods

This study aims to investigate graduate education students' dispositions on statistical literacy in online learning. The researchers will explore the relationships between various factors in students' dispositions to statistics (e.g., motivation, anxiety, attitudes, and beliefs) and students' learning achievement (statistical literacy) in online statistics courses. The research design included the development of survey items and survey delivery method, data collection and analysis, and further improvement of survey items and the instrument. The research team used Qualtrics to design a survey instrument to measure students' dispositions and their statistical literacy in online statistics learning. The survey instrument was designed with references to several existing instruments: the Academic Motivation Scale (Vallerand et al., 1992), the Statistical Anxiety Rating Scale (Cruise et al., 1985), the Survey of Attitudes Toward Statistics (Schau et al., 1995), the Attitudes Toward Statistics Scale (Wise, 1985), and the Basic Literacy in Statistics assessment (Ziegler, 2014), with the addition of a demographic information section and open-ended questions. The target population is graduate students majoring in education who are taking online statistics courses. The researchers used online modes of data collection by administering the survey instrument via e-mail and Canvas (a university's online learning management system) to save costs and achieve a high response rate (Fraenkel et al., 2014). The data was collected for instrument improvement based on feedback on clarity, errors, and impartiality and for testing the type and format of survey items.

The instrument includes five sections: (1) items for demographic information; (2) items for dispositions (motivation, anxiety, attitudes, and beliefs); (3) items for statistical literacy; (4) items for

students' experience in online learning; and (5) items of five open-ended questions. Table 1 shows examples of survey items of survey sections and information categories to reflect on the structure and items of the instrument.

Table 1. The structure and items of the instrument

Section	Categories	Item #	Examples
I- Demographic Information	I1- Demographics	1-3	What gender do you identify yourself with?
	I2- Educational Background	4-7	What is your educational background?
	I3- Work Experience	8	Please rate the importance of statistics in our job.
II- Dispositions (One matrix table)	II1- Motivation	9-16	I am happy to learn new statistics knowledge.
	II2- Anxiety	17-25	I am anxious about discussing or communicating statistical information.
	II3- Attitudes	26-33	I enjoy taking statistics courses.
	II4- Beliefs	34-41	Statistics courses will help me to improve my research skills for the paper.
III- Statistical Literacy (One matrix table)	III1- Self-assessment	42-47	I have basic interpretation skills to describe what the statistics results mean in the context of the problem.
	III2- Assessment	48-59	I can understand the difference between a sample and a population.
IV- Online Learning (One matrix table)	IV1- Learning Experience	60-68	I prefer to take statistics courses online.
	IV2- Learning Preference and satisfaction	69-71	What are the reasons for taking online courses?
V- Open-ended questions	V1- Statistics Learning	72	Share your experience with learning statistics.
	V2- Online Statistics Learning	73-74	What are the advantages of learning statistics online?

4. Pilot Study

In order to make sure that the participants clearly understood the direction of the survey, a pilot study was conducted for four weeks with this newly developed instrument to obtain feedback on the clarity, errors, and impartiality of survey items. In addition, the pilot study will identify any ethical and practical issues that may halt the main research study in the future (Doody & Doody, 2015). The researchers used multiple approaches to analyze data to improve the survey and to have responsiveness and applicability. This pilot survey was electronically administered to 62 graduate education students and faculty members that included 40 master's and 15 doctoral degree-seeking students who were taking online statistics courses in the College of Education at two public universities in Florida. Seven faculty members were recruited to review the instrument to increase its content validity through their expertise. The guiding questions were as follows: How long does it take to complete the survey? Are there any confusing or unclear items? Are there any item structures that need to be adjusted? Are there any items with uncomfortable language? What additional suggestions should be made to improve this survey? After the research team collected participants' feedback and suggestions, the researchers performed data analysis for each item. In addition, the researchers ran a focus group to integrate the results and analyze the feedback to improve the instrument items.

5. Preliminary Results and Discussion

The survey was administered electronically via Qualtrics to collect data regarding demographic information and items pertaining to measuring graduate education students' dispositions on statistics literacy in online learning. Among the participants who were invited to the survey, 55 participants responded to all survey items with an 89% response rate. The overall survey had high reliability with Cronbach's alpha coefficient of .944, and Cronbach's alpha for motivation, attitudes, beliefs, statistical literacy, and online learning scales were .859, .757, .873, .950, and .886, respectively. The alpha value of the anxiety scale was increased from .463 to .799 after removing two items with a low corrected item-total correlation. The analysis of the response patterns for this pilot study provided useful feedback to enhance the quality of the instrument for future research projects. As a result of the pilot study, a few irrelevant items were removed under the demographic section, and a few were revised under the dispositions section. In addition, two survey items were reworded for clarity, jargon was removed if needed, and the questions were shortened to facilitate easy survey completion. For the open-ended question section, an overlapped item was removed, and one item was rephrased to eliminate biases. Additionally, the researchers decided to activate "the back button" to provide an option in case the participants desire to modify any answers. A brief summary of the findings and revised survey items are shown below in Table 2.

Table 2. Preliminary findings and revised survey items

Types of Issues	Categories	Original Items	Suggested Actions	Revised Items
Need clarification	I. Demographic Information IV. Online Learning	Not clearly explained what online learning means.	Clarify information.	In this study, online courses include virtual & hybrid courses and synchronous & asynchronous learning.
	IV. Online Learning	Please rate the following statements according to your experience with online learning.	Revised the question.	Please rate the following statements according to your experience with statistics online learning.
Rewording	II. Dispositions	Statistics courses will help me to improve my research skills for the paper.	Revised the question.	Statistics courses will help me to improve my research skills for the research paper.
Missing answer choice	I. Demographic Information	Please rate the importance of statistics in your current or future job.	Consider those who do not have a job.	Added "N/A" in the answer choice.
Similar Items	V. Open-ended	In your opinion, what are the most significant advantages and disadvantages of learning statistics online?	There is an overlap with the question asking for challenges.	What are the most significant advantages of learning statistics online?

6. Limitations

This is a pilot study that was conducted to test the implementation feasibility and content validity of this newly developed instrument. The pilot study was conducted in November, near the end of the fall semester, which may have affected the quality of student responses. A better timing needs to be considered when administering the survey in future studies. The researchers used the convenience sampling technique to collect data. Therefore, the respondents may not be true representatives of other graduate programs in education across the USA.

7. Significant Contributions

This instrument was developed as a first step to improve the survey, which will be used to conduct a major research project. This study is expected to emphasize the importance of students' dispositions in mastering statistical literacy in graduate education programs to prepare students to become educational leaders, teacher educators, and scholars after graduation. The findings of the study will help faculty who teach statistics courses better understand the relationships between graduate education students' dispositions and statistical literacy in online learning. Furthermore, the results will inform faculty of effective teaching strategies to improve statistics courses in online learning. The findings will help graduate education students in doctoral research and dissertations grow as researchers, become leaders in the education world, and increase their statistical literacy for their future professional development.

8. Key References (Other references will be available upon request)

Cruise, R. J., Cash, R. W., & Bolton, D. L. (1985). Development and validation of an instrument to measure statistical anxiety. *American Statistical Association Proceedings of the Section on Statistical Education, 4*(3), 92–97.

DeVaney, T. A. (2016). Confirmatory factor analysis of the statistical anxiety rating scale with online graduate students. *Psychological Reports*, *118*(2), 565–586. https://doi.org/10.1177/0033294116644093

Sharma, S. (2017). Definitions and models of statistical literacy: A literature review. *Open Review of Educational Research, 4*(1), 118–133. https://doi.org/10.1080/23265507.2017.1354313

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

Examining Learners' Emotions in Collaborative Language Learning: A Review of Selected Journal Publications from 2018-2022

Haizhen Guo [1], You Su[1*]
[1] Beijing University of Posts and Telecommunications
[*] suyou@bupt.edu.cn

Abstract: *Students' emotions have been increasingly considered an important variable affecting learners' performance in computer-supported collaborative learning. However, there have been few systematic review studies investigating how emotions were assessed and analyzed in collaborative language learning settings. To fill this gap, this study reviews empirical research on language learners' emotions in collaborative language learning published from 2018 to 2022 in top-tier journals specializing in technology-enhanced language learning. The results indicate that: a) the analysis of emotions only accounts for a limited part in most selected studies, in particular, emotion regulation as a hot topic in emotion research, receives little attention in the area of collaborative language learning. b) the high proportion of studies focus on enjoyment, motivation, anxiety, self-efficacy or self-confidence, and general attitudes of learners, whereas nuanced analysis of other micro emotions like boredom, frustration, and hope is often neglected in the frameworks for analyzing emotions. c) the measurements of emotions are mainly based on questionnaires and interviews in most reviewed studies, and future research is suggested to adopt multi-modal assessments to triangulate the findings. Overall, this study presents a holistic view of emotion research in collaborative language learning, and the findings are expected to be beneficial for both researchers and language teaching practitioners.*
Keywords: collaborative language learning, emotion, emotion regulation

1. Introduction

Students' emotions and emotion regulation strategies have been increasingly considered important variables affecting learners' performance in computer-supported collaborative learning. There has been a concomitant rise of needs and interest in emotion research in technology-enhanced collaborative learning environments, and a systematic review of emotion research can arouse broadened knowledge and may yield insights into this promising research area. Prior review studies have synthesized the research on technology-enhanced collaborative learning from different perspectives. For example, Li (2018) summarized 21 empirical studies on computer-mediated collaborative writing and provided insights for future studies in this area. Zhang et al. (2022) conducted a review of 42 studies on technology-enhanced collaborative writing (TECW) in second and foreign language learning by identifying their implementation in terms of practices, technology, and challenges. Gerbeth et al. (2022) reviewed the positive relations between emotional competence and team learning behaviors by examining 32 studies in different domains ranging from health care to the military. However, in the pool of systematic reviews of collaborative learning, there is a lack of review studies probing into learners' emotional status in collaborative language learning. To fill this gap, we conduct our review based on the following three research questions.

(1) What research contexts have been investigated regarding school levels, target languages, and language skills?

(2) What emotions have been examined in empirical research on collaborative language learning?

(3) What frameworks and instruments have been adopted for analyzing emotions in collaborative language learning?

2. Methodology

The present study focuses on four top-tier journals in technology-enhanced language learning, namely, *Computer Assisted Language Learning*, *ReCALL*, *System*, *and Language Learning & Technology*, which are the only four CALL journals included in the Social Sciences Citation Index. The search process can be divided into three steps. First, we searched the websites of these four journals with the keywords include "emotion", "affection", or "emotion regulation" with the publication years set from 2018 to 2022. The initial literature search yielded 454 articles. Second, based on titles and abstracts, the author screened all the articles and narrowed them down according to the following criteria: a) empirical studies; b) addressed collaborative language learning settings; c) analyzing learners' emotion or emotion regulation in one part at least of the results and discussion. Third, we screened the full texts of the remaining 27 ones to determine whether they were conducted in collaborative language learning. Consequently, with 7 ineligible studies removed, 20 studies were retained for the review and analysis in this study. Adopting a holistic approach, the author summarized the 20 articles by exhibiting the key information from the facets of context and participants, a framework for analyzing emotions, frequencies of different types of emotions, and data collection methods.

3. Results

3.1. Participants and contexts of reviewed studies

Table 1. Participants and contexts of reviewed studies

Authors	Number of participants	School level	Target language	Language skills
Azkarai and Kopinska (2020)	62	Elementary school	English	Listening , grammar
B árk ányi (2021)	35644	Uncertain	Spanish	Speaking
Bielak. and Mystkowska-Wiertelak (2020)	142	Undergraduate	English	Comprehensive skill
Canals (2020)	18	Undergraduate	English, Spanish	Speaking
Chen and Yu (2019)	30	Undergraduate	English	Writing
Cho and Castaneda (2019)	82	Undergraduate	Spanish	Grammar
Dao (2020)	56	Middle school	English	Speaking
Guo et al.(2020)	39	Undergraduate	English	Speaking
Jiang and Dewaele (2019)	564	Undergraduate	English	Listening, speaking
Li et al.(2018)	2078	High school	Chinese	Comprehensive Skill
Qiu and Lee (2019)	24	Undergraduate	English	Writing
Saeedakhtar et al.(2021)	48	High school	English	Listening
Selcuk et al.(2019)	6	High school	English	Writing
Su et al.(2019)	285	Undergraduate	English	Reading, writing
Villarreal and Lazaro-Ibarrola (2022)	26	Elementary school	English	Writing
Wu et al.(2019)	48	Undergraduate	English	Writing
Zabihi and Ghahramanzadeh (2022)	54	Undergraduate	English	Writing
Zhang and Tsung (2020)	216	Undergraduate	Chinese	Comprehensive skill
Zhang et al.(2021)	6	Undergraduate	English	Writing
Zheng et al.(2021)	51	Undergraduate	English	Speaking

In the selected studies, participants' school levels ranged from elementary school to undergraduate, with college students gaining more attention (65%) than any other levels. The majority of research focused on English(80%), whereas few studies have explored other languages like Spanish and Chinese. Regarding language skills, it was found that writing and speaking were paid more attention to students' collaborative language learning.

3.2. Representative frameworks adopted for analyzing emotions in collaborative language learning

In general, three strands in the reviewed studies can be identified: positive emotions (e.g. Guo et al., 2020; Jiang & Dewaele, 2019; Qiu & Lee 2020; Selcuk, 2019; Su et al., 2019), negative emotions (e.g. Wu et al., 2019; Zheng et al., 2021), and emotion regulation strategies (Bielak & Mystkowska-Wiertelak, 2020; Zhang et al., 2021). Firstly, Li et al. (2018) validated a Foreign Language Enjoyment (FLE) scale consisting of three dimensions, namely, FLE-private, FLE-teacher, and FLE-atmosphere. This instrument has also been successfully adapted to various learning contexts (Jiang & Dewaele, 2019; Zhang & Tsung, 2020; Zhang et al., 2021). Secondly, one representative study on negative emotions by Zheng et al. (2021) examined English language learners' public speaking anxiety and revealed that video-based formative assessment could relieve learners' speaking anxiety. Thirdly, Bielak and Mystkowska-Wiertelak (2020) explored Polish college English learners' use of emotion regulation strategies using both the Manage Your Emotion (MYE) scenario questionnaire and coding scheme based on the vignette methodology. In addition, with the support of Anion Variable Tester Software and a coding scheme adapted from Hadwin et al. (2018), Zhang et al. (2021) examined language learners' emotion regulation from three aspects: self-regulation, co-regulation, socially-shared regulation, and these strategies were found to be critical in creating a favorable collaborative atmosphere to achieve knowledge construction and improvement.

3.3. Emotions and data collection methods

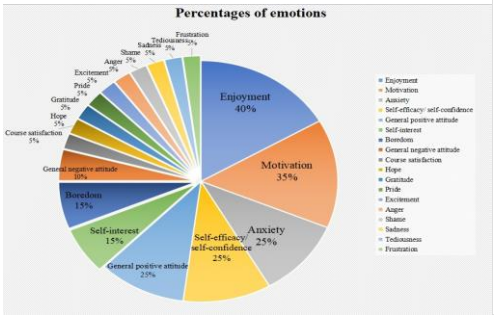


Figure 1. Percentages of emotions

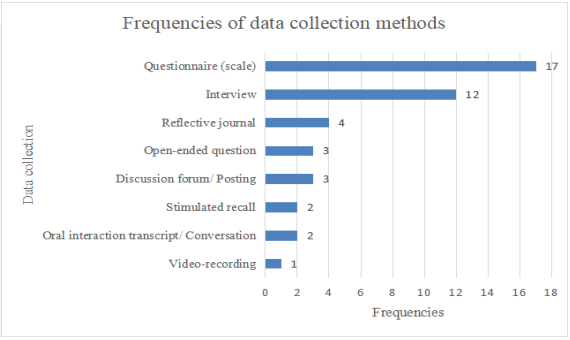


Figure 2. Frequencies of data collection methods

In our pool of studies, there were eighteen types of emotions examined by the researchers, shown in figure 1. Enjoyment (40%) was the most frequently examined emotion in collaborative language learning, followed by motivation (35%) and anxiety (25%), self-efficacy or self-confidence (25%), and general positive attitude (25%). Additionally, emotions like course satisfaction, hope, gratitude, pride, excitement, anger, shame, sadness, tediousness, and frustration were only explored by five percent of the reviewed studies. Figure 2 shows the data collection methods from the twenty studies. Overall, eight types of data collection methods were adopted and most studies reported more than one type of them. Questionnaires (85%) were used the most frequently, followed by interviews (60%) and reflective journals (20%). Nonetheless, only 5 percent of research has adopted video recording for collecting data on the emotions of learners in collaborative language learning.

4. Discussion and conclusion

Through a thorough review of selected studies, the findings and suggestions of this study were identified:

First, the analysis of emotions in collaborative learning accounts for a rather limited part of the whole research, and emotions are neglected in this area. In particular, emotion regulation, as a crucial variable influencing learners' performance, has been paid much less attention by researchers in collaborative language learning. Second, few studies tap into the factors influencing learners' emotions. Many variables may play an important role in learners' emotions in collaborative language learning. For instance, Zhang et al. (2021) indicated the effectiveness of using emojis by peers on promoting students' emotion regulation and create an enjoyable shared atmosphere; group leaders, as emotional engagement supporters, boost group members' motivation and self-confidence in language learning through providing praise and motivational feedback (Selcuk et al., 2019); Dao (2020) demonstrated the usefulness of providing peer interaction strategies instruction on language learners' emotional engagement by fostering more effective interaction between peers as mutually affective supporters. Therefore, more corresponding research should be implemented in the future. Third, the micro emotions of learners should be dug into as well. Not only anxiety and enjoyment, but also other largely unexplored emotions of online learners such as boredom, frustration, and hope require further examination. Fourth, given that emotions are complex, volatile, and individually different, only adopting self-reported data alone is not convincing enough. To further triangulate research findings, multi-modal data collection of the learning process could be a more reliable and persuasive method to capture the levels and changes of their emotions and evaluate emotions experienced moment-by-moment.

References

Bielak, J. & Mystkowska-Wiertelak, A. (2020). Investigating language learners' emotion-regulation strategies with the help of the vignette methodology. *System*, 90, 102208.

Cho, M.-H. & Castañeda, D.A. (2019) . Motivational and affective engagement in learning Spanish with a mobile application. *System*, 81, 90–99.

Dao, P. (2020). Effect of interaction strategy instruction on learner engagement in peer interaction. System, 91, 102244.

Gerbeth, S., Stamouli, E. & Mulder, R.H. (2022). The relationships between emotional competence and team learning behaviours. *Educational Research Review*, 36, 100439.

Guo, Y., Xu, J. & Xu, X. (2020). An investigation into EFL learners' motivational dynamics during a group Communicative Task: A classroom-based case study. *System*, 89, 102214.

Jiang, Y. & Dewaele, J.-M. (2019). How unique is the foreign language classroom enjoyment and anxiety of Chinese EFL learners? *System*, 82, 13–25.

Li, C., Jiang, G. & Dewaele, J.-M. (2018). Understanding Chinese high school students' foreign language enjoyment: Validation of the Chinese version of the foreign language enjoyment scale. *System*, 76, 183–196.

Li, M. (2018). Computer-mediated collaborative writing in L2 contexts: An analysis of empirical research. *Computer Assisted Language Learning*, 31(8), 882–904.

Qiu, X. & Lee, M.-K. (2020). Regulated learning and self-efficacy beliefs in peer collaborative writing: An exploratory study of L2 Learners' written products, task discussions, and self-reports. *System*, 93, 102312.

Selcuk, H., Jones, J. & Vonkova, H. (2019). The emergence and influence of group leaders in web-based collaborative writing: Self-reported accounts of EFL learners. *Computer Assisted Language Learning*, 34(8), 1040–1060.

Su, Y. et al. (2018). Moving literature circles into wiki-based environment: The role of online self-regulation in EFL learners' attitude toward collaborative learning. *Computer Assisted Language Learning*, 32(5-6), 556–586.

Tan, J.S.H. and Chen, W. (2022). Peer feedback to support collaborative knowledge improvement: What kind of feedback feed-forward?, *Computers & Education*, 187, 104467.

Wu, W.-C.V. et al. (2019). Free from demotivation in EFL writing: The use of online flipped writing instruction. *Computer Assisted Language Learning*, 33(4), 353–387.

Zabihi, R. & Ghahramanzadeh, B. (2022). Proficiency pairing and engagement with language in L2 peer interaction. *System*, 105, 102725.

Zhang, R. et al. (2022). Implementing technology-enhanced collaborative writing in second and Foreign Language Learning: A review of practices, technology and challenges, *Education and Information Technologies*, 27(6), 8041–8069.

Zhang, Z., Liu, T. & Lee, C.B. (2021). Language learners' enjoyment and emotion regulation in online collaborative learning. *System*, 98, 102478.

Zheng, C., Wang, L. & Chai, C.S. (2021). Self-assessment first or peer-assessment first: Effects of video-based formative practice on learners' English public speaking anxiety and performance. *Computer Assisted Language Learning*, 1–34.

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

A Comparative Study of English Learning Motivation and Online Self-regulation Between Students at Two Universities in Beijing and Taipei

Jing-yi WANG, Xin LIU, Yu-rui BAI, Ping LU, Chun-ping ZHENG*
Beijing University of Posts and Telecommunications
* zhengchunping@bupt.edu.cn

Abstract**:** *This research compared college students' English learning motivation and online self-regulation in Beijing and Taipei, China. A quantitative study was conducted based on 449 students' responses to two questionnaires at two universities (BUPT and TUST). The results indicated that motivational factors of English language learners in both universities were positively correlated with all factors in their online self-regulation. Meanwhile, "Others' Expectations" played a significantly positive role in predicting most factors of online self-regulated learning. Students at BUPT presented a higher Ideal L2 Self, for their "Cultural Interest" and "Instrumentality-promotion" had stronger predictive power, while students at TUST showed a higher tendency on Ought-to L2 Self. This research facilitated the understanding of English learning motivation and online self-regulation among college students at two universities in Beijing and Taipei. It also provided pedagogical implications for improving the quality of online education.*

Keywords: online self-regulation, L2 motivation, online language learning, EFL

1. Introduction

In recent years, studies targeting online learning have revealed its advantages of convenience, flexibility, student-centered, and low cost, but also pointed out deficiencies such as redundant information and online distractions (e.g., Dhawan, 2020; Simonova et al., 2021). Therefore, online learning requires students to equip themselves with high-level self-regulatory learning abilities, which could be triggered and maintained by their learning motivation (**错误!未找到引用源。**; **错误!未找到引用源。**). Therefore, it's of profound significance to probe L2 learners' motivation and self-regulation in online learning environments. Although previous research has explored the structural relationships between L2 learners' motivation and their self-regulation, the comparison of the two constructs concerning students in different regions remains inconclusive, especially in online environments. This study emphasizes the similarities and differences between students at two universities in Beijing and Taipei, which is beneficial to better understand the current situation of college students' L2 learning in these two cities.

2. Literature Review

Learners' motivation is a psychological process that stimulates, guides, and maintains learning activities (Fetsco & McClure, 2005). Dörnyei (2005) proposed the L2 Motivational Self System (L2MSS), which can be divided into three dimensions, namely, the Ideal L2 Self, the Ought-to L2 Self,

and L2 Learning Experience. Research indicated that L2MSS was applied in different language learning perspectives (e.g., Lopez, 2022; Wong, 2020). Based on this system, an instrument was adapted from You and Dörnyei's questionnaire (2016) to measure English learning motivation in online environments, targeting Chinese students from two universities in Beijing and Taipei.

Self-regulation originated from the concept of social cognitive theory. In recent years, students' self-regulatory abilities in online language learning environments have gained increasing attention (e.g., Lian et al., 2021; Kayaduman et al., 2022). Zheng et al (2016) adapted the Online Self-regulated Learning Questionnaire (OSLQ) (Barnard et al., 2009) for evaluating Chinese students' online English learning. The questionnaire was proven valid and has been employed in many following studies to examine students' online self-regulated English learning (e.g., Li & Zhou, 2020; Zheng & Wang, 2020). In this study, two more dimensions, metacognitive regulation and affection regulation were added to the questionnaire based on Tseng's research (Tseng et al., 2006) to probe the factorial structures of students' online self-regulation.

Motivation is a key factor of the forethought phase in self-regulation and is essential to the subsequent self-regulatory learning behavior (Zimmerman, 2002; Zheng et al., 2018). Previous studies have investigated the relationship between motivation and self-regulation in L2 learning contexts (e.g., Wen & Piao, 2020; Yossatorn et al., 2022). This study aims to explore the relations between L2 motivation and online self-regulation of English language learners from two universities, as well as to compare the features of students from both universities concerning the two variables. It is guided by the following two research questions:

(1) What are the factorial structures of English language learners' motivation and online self-regulation?

(2) What are the relations among the factors of learners' motivation and online self-regulation?

3. Research Method

3.1. Research Context and Participants

The study was conducted at two comprehensive universities, Beijing University of Posts and Telecommunications (BUPT) and Taiwan University of Science and Technology (TUST). Both universities are equipped with elaborate computer-supported learning environments for language learners. The study included 449 participants, with 247 BUPT students recruited from a compulsory course "College English Listening and Speaking" and 202 TUST students selected from different English optional courses. The participants were almost equally distributed between boys and girls and were all aged within the range of 18-24. All participants had rich experience in online English language learning. BUPT students took good advantage of online learning platforms provided by the university, while TUST students were better at discovering online resources autonomously. Meanwhile, most students from TUST studied at home or in libraries, while dormitories or classrooms were preferred by students in BUPT. Different courses and after-class activities ignited higher social interests of TUST students who are more willing to share their ideas with friends or even foreigners than students in BUPT.

3.2. Instruments

Two questionnaires were used in this study and measured with a five-point Likert scale. One is the online language learning motivation (OLLM) questionnaire (You & Dörnyei, 2016), which consisted of five dimensions, namely, Online English Learning Experience (OELE), Cultural Interest (CI),

Instrumentality-promotion (IPO), Instrumentality-prevention (IPR), and Others' Expectations (OE). Among them, IPO reflects learners' desire to achieve certain goals through language learning, such as obtaining a good job, while IPR shows learners' desire to avoid negative learning consequences, such as a failure in communication. The other one, the online self-regulated English learning (OSEL) questionnaire contained eight factors, namely, Time management (TM), Task strategies (TS), Environment structuring (ES), Help-seeking (HS), Self-evaluation (SE), Metacognitive Regulation (MR), and Affective Regulation (AR). Among them, MR refers to the use of metacognitive strategies, such as monitoring and controlling concentration, while AR demonstrates learners' ability to regulate their emotions during language learning.

3.3. Data Collection and Analysis

This study mainly conducted quantitative research methods and the data were collected through the above two questionnaires, both designed in Chinese for reader-friendly reasons. Drawing upon the previous research frameworks, a hypothesized structural model concerning the relationship among factors in the two questionnaires was proposed. Therefore, confirmatory factor analysis (CFA) was conducted to examine the validity and reliability of these two adapted questionnaires. Then, the relationship between the finalized OLLM and OSEL factors was analyzed through Pearson correlation analysis. Moreover, stepwise regression analysis was conducted to figure out the predictive power for each impact factor, which could possibly reveal differences and similarities in students' features at both universities.

4. Research Results

4.1. The Factorial Structures of Students' Motivation and Online Self-regulation in BUPT and TUST

CFA was conducted to examine the validity and reliability of the factors in the OLLM and OSEL questionnaires. Factor loadings of all items in OLLM of BUPT students were above 0.50. Besides, the Average Variance Extracted (AVE) values of all factors ranged from 0.43 to 0.60. The Composite Reliability (CR) values of all variables ranged from 0.70 to 0.86. Moreover, the overall Cronbach's alpha value for the survey was 0.85, indicating sufficient internal consistency of the survey items. As for TUST students, all factor loadings were above 0.55. The AVE values ranged from 0.43 to 0.60 and the CR values were between 0.77 and 0.92. The overall Cronbach's alpha value was 0.92.

Meanwhile, factor loadings of BUPT students' OSEL items ranged from 0.55 to 0.80. The AVE values were beyond 0.45, and the CR values were within the range of 0.76-0.84. The overall Cronbach's alpha value was 0.95. As for TUST students, the factor loadings were between 0.56-0.88, as well as higher AVE values (above 0.46) and CR values (0.77-0.89) in terms of the BUPT statistics for each factor. Besides, the overall Cronbach's alpha value was 0.97. Therefore, the validity and reliability of the OLLM and OSEL questionnaires have been identified.

4.2. The Relationship between Students' Motivation and Online Self-regulation in BUPT and TUST

Pearson's correlation analyses showed that the OLLM factors for students were all positively related to eight OSEL factors in both BUPT and TUST. Stepwise regression analysis was conducted to explore the specific structural relationships between students' English learning motivation and online self-

regulation. Factors of the OLLM questionnaire were identified as predictors while the OSEL factors were considered outcome variables. The stepwise regression analysis results showed that for BUPT students, three factors in OLLM ("OE", "CI", and "IPO") played an important role in predicting most of the OSEL factors. "OE" appeared to be the strongest predictor, for it could positively predict seven factors in OSEL except for "ES", while "CI" had significantly positive predictive power for six factors instead of "ES" and "HS". "IPO's" prediction for five OSEL factors, "GS" ($b = 0.17$, $T = 2.72$, $p < 0.01$), "ES" ($b = 0.14$, $T = 2.20$, $p < 0.05$), "TM" ($b = 0.18$, $T = 2.96$, $p < 0.01$), "MR" ($b = 0.18$, $T = 2.88$, $p < 0.01$), and "AR" ($b = 0.15$, $T = 2.52$, $p < 0.05$) was also significantly positive. Additionally, "IPR" and "OLLE" had weaker power in predicting OSEL factors, for they showed less significance in statistics and could predict fewer OSEL factors. The results showed that BUPT students presented more obviously on Ideal L2 Self and deeply valued social appraisal.

Compared with BUPT students, all factors in OLLM were weaker in predicting the OSEL factors as most of them could significantly predict four factors of online self-regulation. Same as BUPT results, "OE" was considered one of the strongest predictors while "IPO" could positively predict four OSEL factors. However, "CI" could only positively predict "SE" ($b = 0.19$, $T = 2.71$, $p < 0.01$), while "IPR" had a significantly positive prediction for "GS" ($b = 0.26$, $T = 3.62$, $p < 0.001$), "MR" ($b = 0.32$, $T = 4.83$, $p < 0.001$), and "AR" ($b = 0.17$, $T = 2.35$, $p < 0.05$). "OLLE" could positively predict four OSEL factors, namely, "ES" ($b = 0.25$, $T = 3.33$, $p < 0.01$), "TM" ($b = 0.25$, $T = 3.65$, $p < 0.001$), "MR" ($b = 0.21$, $T = 3.08$, $p < 0.01$), and "AR" ($b = 0.28$, $T = 3.79$, $p < 0.001$). The results indicated that for TUST students, a good English learning experience could stimulate their motivation, which reflected a more noteworthy Ought-to L2 self and extrinsic motivation.

5. Discussion and Conclusion

This study explored the current characteristics and structural relationships between college students' English learning motivation and online self-regulation. On one hand, students in both universities showed significant predictions of "Others' Expectations" on their online self-regulation, reflecting the importance of social influence, which is contrary to the results of previous studies. On the other hand, BUPT students' "Instrumentality-promotion" and "Cultural Interest" in the Ideal L2 Self had significant power for positively predicting most factors of online English self-regulation, which is in concordance with the previous research (Zheng et al., 2018). The high predictive power of "CI" on "Environment Structuring" and "Help Seeking" indicated that most students in BUPT learned English for intrinsic reasons. They were inclined to set clear goals, take good advantage of time, and employ various learning strategies in online learning environments. What's more, TUST students presented higher predictive power for online self-regulation on "Instrumentality-promotion", "Instrumentality-prevention", and "Others' Expectations", indicating a strong ought-to L2 Self. The results effectively echoed previous conclusions, showing the positive impact of instrumental motivation on OSEL (Hwa, 2020; Zheng et al., 2022). This study contributed to further understanding the relationship between college students' English learning motivation and online self-regulation, and provided empirical experiences and pedagogical implications for mutual learning and educational exchanges between the two regions.

Acknowledgments

This research is funded by the National Social Science Foundation in China (Grand No. 19BYY221).

References

Barnard, L., Lan W. Y., To Y. M., Paton,V. O., & Lai. S. L. (2009). Measuring self-regulation in online and blended learning environments. *Internet & Higher Education*, *12*(1), 1-6.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in L2 acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Lopez Perles, A. (2022). Impact of study abroad on the motivation of Spanish/L2 learners: A study from the perspective of the L2 Motivational Self System. *REVISTA SIGNOS, 55*(108), 91-116.

Tseng, W. T., Dörnyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied Linguistics, 27*(1), 78-102.

Wen, X. , & Piao, M.(2020). Motivational profiles and learning experience across Chinese language proficiency levels. *System*, *90*(3), 102216.

Wong, Y. K. (2020). Effects of language proficiency on L2 motivational selves: A study of young Chinese language learners. *System*, *88*, 102181.

Yossatorn, Y., Binali, T., Weng, C., & Awuor, N. O. (2022) Relating university students' online self-regulated English learning to motivational beliefs: a structural equation modelling analysis. *Behaviour & Information Technology*. doi:10.1080/0144929x.2022.2048074.

You, C. J., & Dörnyei, Z. (2016). Language learning motivation in China: Results of a large-scale stratified survey. *Applied Linguistics, 37*(4), 495-519.

Zimmerman, B. J. (2002). "Becoming a self-regulated learner: An overview." *Theory into Practice*. *41*(2), 64-70.

Zheng, C. P., Liang, J. C., Li, M., & Tsai, C. C. (2018). The relationship between English language learners' motivation and online self-regulation: A structural equation modelling approach. *System, 76*, 144-157.

李莹莹 & 周季鸣.(2020).线上英语教学环境下学习者自我调控能力研究. *外语与外语教学*(05),45-54+149.

郑春萍,高梦雅,许玲玉 & 陈旭.(2022).在线教育环境下大学生英语学习动机与自我调控学习研究. *西安外国语大学学报*(03),67-72.

郑春萍 & 王丽丽.(2020).高校英语学习者技术接受与在线自我调控学习的结构关系研究. *外语教学*(02),64-70.

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

Bibliometric research in educational technology: a critical literature review incorporating content analysis

Cheng-Liang Wang[1,2], Yu Chen[1,2], Hao-Ming Wang[1,2*], Yu-Hui Jing[2], Ke-Ke Zhu[3]

[1] Department of Education Information Technology, Faculty of Education, East China Normal University

[2] College of Educational Science and Technology, Zhejiang University of Technology

[3]College of Foreign Languages, Zhejiang University of Technology

[*]910305038@qq.com

Abstract*: As a unique research method, bibliometrics has received much attention in educational technology. To clarify the development of this cross-boundary issue, the study conducted a retrospective analysis of educational technology journals in JCR-Q1, obtained 50 articles that analyzed educational technology using bibliometrics after retrieval, and at last coded, counted and critically analyzed them. Then, the study found that this frontier cross-boundary topic has developed rapidly in the past three years, which mainly presents as the increasing number of journals publishing this topic, articles published, nationalities of the author, and the diversification of research themes. The study also counted the tools and data used for the topic and the specific research details, analyzed and grasped the development pattern of the above elements. Besides, the study conducted a critical review of current research, made a number of suggestions on its future development in the hope of promoting the development of this cutting-edge cross-boundary topic.*

Keywords: *critical review, bibliometric, educational technology, systematic review*

1. Introduction

In 1969, the scholar Alan Pritchard (1969) coined the term "bibliometrics", and now, as a discipline and a unique research method, bibliometrics has a history of more than 50 years. However, it isn't until recently that the application in educational technology shows evidence of flourishing. Due to the bias of subjective understanding and other factors, the application of new technologies in new subject areas often encounters problems. Via a critical review, the weaknesses, contradictions, controversies and inconsistencies of the existing research could be revealed (Grant & Booth, 2009), and promising directions and constructive, insightful suggestions can also be pointed out.

2. Methodology and Materials

This study is intended to be conducted in the form of Critical review. Critical review provide an opportunity to take stock and assess the value of prior work, a way through which they can offer constructive input to other scholars, improve the existing body of knowledge and offer further suggestions for research priorities (Palvia et al., 2004).

To obtain the analytical data, academic journals rated as JCR Q1 and closely relevant to educational technology (16 journals) were reviewed for this study. We used the search formula TS = ("bibliometric*" OR "scientometric*" OR "visual analysis" OR "knowledge map*" OR "knowledge graph*") to target literature containing relevant keywords. After a three-person back-to-back manual screening, a total of 50 articles were retained after the screening.

After completing the manual screening process, two authors read the 50 articles and collected detailed information in four categories: article metadata, methodology, performance analysis, and additional research details for the articles.

3. Result and Discussion

A critical review first requires a reasonably descriptive statistics of the existing studies. Most of the studies, accounting for 72% of the total, were published between January 2021 and October 2022. This could be largely credited to the fact that the research method of bibliometrics is increasingly accepted and recognized by scholars in the field of educational technology research. In terms of the number of a country's publication, 50 articles come from 17 countries, among which 17 articles come from China, 9 articles from Spain, and 7 articles from Turkey. These three countries, contributed 33 articles. In addition to the above information, this study also calculated the number of authors of the literature. Statistics found that the average number of authors of the literature included in the analysis was 3.26, and the literature with the most authors had six authors, and the least was only one author; the number of open access literature was 17, accounting for 34% of all the literature included in the analysis; In terms of sample size included in the analysis, the study that analyzed the most literature was the study by Valtonen et al. (2022), which analyzed a total of 30,632 papers, and the least was a study on an analysis of Horizon Reports, with a sample size of seven (Dubé & Wen, 2022).

Content analysis examines the following three parts: core research focus, research tools and research databases.

The core topics can be broadly divided into three categories: the first category is the econometric analysis of specific journals over a long period of time, which respectively analyzed CHB, BJET, AJET, JCAL, EAIT. The second category is the analysis of technology applications, including the developmental analysis of the application of gamification-related technologies in education etc. The third category is the review analysis of technological scenarios and emerging concepts in educational technology. Among them the studies concerning technological scenarios are about virtual labs, museum learning, etc., and the ones concerning emerging concepts are about virtual patient, multimodal learning analytics. The development of bibliometrics is closely linked to the evolution of quantitative analysis techniques, visualization techniques and database technologies, so this places certain demands on scholars in terms of grasping the technological frontiers. To understand the logic of how bibliometric research develops at the technological level in the field of educational technology, this study counted the research tools and databases, whose usage are presented in Figure 1 respectively.

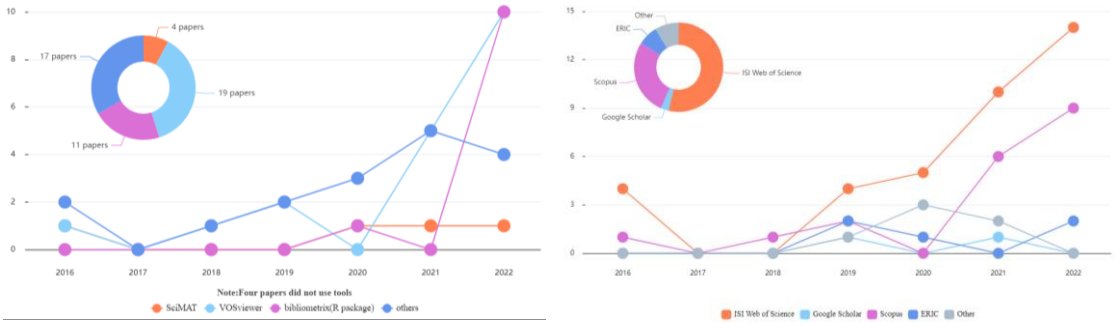


Figure 1. Evolution of bibliometrics research tools and database

4. Critical Analysis and Conclusion

Through a comprehensive review and critical examination of bibliometric research in educational technology, this study charts the current research status of the topic, while also allowing some cutting-edge questions to be raised. First, bibliometric research often requires a grand perception of the discipline. Although it is a quantitative-centered research model, it should still reflect the language system of the discipline on the basis of quantity, examine the development issues of the discipline, and propose insights for future development. Today, however, some studies lack themes and images, are mainly dependent on the piling up of quantitative information and using software, which is the wrong way to use bibliometric research in educational technology. Second, deficiencies exist in the frontier of bibliometric in educational technology. For example, in the analysis of disciplinary structure, few scholars have given attention to the use of terminology analysis rather than keyword analysis with the aim of presenting a more comprehensive picture of the discipline. Furthermore, a content analysis of the included literature

for analysis revealed that few scholars have noticed the value of the quantitative laws of bibliometrics. Their lack of a systematic understanding of bibliometrics as a discipline has led to some misunderstandings in specific analyses.

Data Availability Statements

The datasets generated during and/or analysed during the current study are available from the first author (email: wcledutech@163.com) or corresponding author (910305038@qq.com) on reasonable request.

References

Pritchard, A. (1969). Statistical bibliography or bibliometrics? *Journal of Documentation, 25*(4), 348-349.

Grant, M. J., & Booth, A. (2009). A typology of reviews: an analysis of 14 review types and associated methodologies. *Health information & libraries journal*, 26(2), 91-108.

Palvia, D. Leary, E. Mao, V. Midha, P. Pinjani, A. Salam.(2004)Research methodologies in MIS: an update, *Commun. Assoc. Inf. Syst*., 14, 526–542.

Valtonen, T., López-Pernas, S., Saqr, M., Vartiainen, H., Sointu, E. T., & Tedre, M. (2022). The nature and building blocks of educational technology research. *Computers in Human Behavior*, 128, 107123.

Dubé, A. K., & Wen, R. (2022). Identification and evaluation of technology trends in K-12 education from 2011 to 2021. *Education and Information Technologies*, *27*(2), 1929-1958.

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

Research on Key Technologies Framework of Privacy Computing for Education Big Data

Yuping Zhang [1], Tao Wang[2*],Xiaoyu Chen[3]， Ningyuan Jiang[4],Liling Lin[5]

[1, 2, 3,4&5] Central China Normal University

3435236468@qq.com

Abstract: *Educational big data is a data set generated in the whole process of educational activities and collected according to educational needs for educational development and can create great potential value.How to effectively protect the sensitive data,how to balance the development of educational big data and privacy security has become a key issue in the extensive implementation of joint analysis of educational data and the integration of multiple education.This paper constructs a key technical framework for educational big data privacy computing to solve the problems of data security and privacy protection in the field of education.*

Keywords: *educational big data; privacy computing; data life cycle; open and sharing*

1. Introduction

With the application of emerging technologies such as artificial intelligence, big data analysis and blockchain in the field of education, it has provided a strong impetus for the in-depth development of educational data. However, while obtaining valuable information from educational data, educational data also faces privacy risks and ethical issues. There are certain data security risks in all aspects of educational data collection, storage, transmission, and sharing. It is easy to be attacked by viruses and hackers, resulting in personal privacy of educational subjects or data leakage or tampering that is not suitable for public disclosure. Events such as information leakage and enrollment fraud have occurred frequently, which has caused great social hidden dangers.

Above all, the paper focuses on the risk of privacy attack in each stage of the whole life cycle of educational data, and constructs the key technical framework of privacy computing of educational big data. While making full use of the value of data, it achieves the purpose of reducing the risk of privacy leakage of educational data, and pursues the balance between data availability and privacy.

2. Build a data sharing system based on key techniques of privacy computing

Based on distributed federated learning,this paper proposes an educational data sharing framework based on Secure Multi-Party Computation (Jiang Yongbo , Zhou Yuan & Feng Tao,2022) and blockchain authorization.The system consists of five parts:task publisher,participant,trusted organization(TA)and aggregation server(AS)and supervisor,and the data sharing process can be roughly divided into four stages:

(1)Build a decentralized blockchain. The TA trusted authority selects the security parameters and the digital signature scheme before initialization,and executes the initialization algorithm to generate the public parameters and the master key.Users register in TA.Each user or institution uploads its own real identity information to register in TA.After verifying the results,TA will issue a unique identity,generate a pair of public and private keys and generate the corresponding address through Elliptic Curve Cryptography(ECC).Then,the verification information is released by calling the smart contract,and can be added to the blockchain after the verification is passed.

(2)The framework combines blockchain with InterPlanetary File System(IPFS).After storing the trained local model into IFPS,the participants will return a unique Hash.The participants access IPFS through the model Hash to obtain the corresponding model.A block includes training application records,local model Hash and global model of each aggregation operation.The initiator will publish the federated learning request through the consensus mechanism,and then upload the initial model to IPFS.Perform a multi-party data retrieval mechanism,locate the set of participants related to this federal learning request,and forward the request to the relevant participants through task allocation.The initiator writes the model Hash,key and federated learning algorithm into the smart contract,and the participants obtain the key and model Hash by calling the smart contract.

(3) Participants download the initial model from IPFS according to Hash,and use different optimization algorithms to train the model locally.In each update process,the participant uploads the local model to IPFS to get a Hash,and then use the Full Homomorphic Encrypts(Peraković Dragan et al,2022) and packages the Hash with the public key into a transaction.After the smart contract test,it is uploaded to the consortium blockchain.Each participant node in the blockchain exchanges and verifies all the model Hash.AS uses the private key to decrypt the model Hash in each iteration,locates it in IPFS according to the Hash,downloads all the actual models,and then receives the participant's encrypted local model through SMPC and performs confidential calculation of the data and continuously trains a new encrypted global model.Participants download the latest global model and decrypt it according to SMPC,and retrain the model locally.Until the global model is fitted,AS will upload the trained global model to the blockchain.

(4) Participants can download the trained global model from the alliance chain,and use the trained model to mine potential research value without obtaining the educational source data of each participant.Due to the decentralization and auditability of blockchain,various well-trained educational data sharing models can be safely accessed and analyzed.

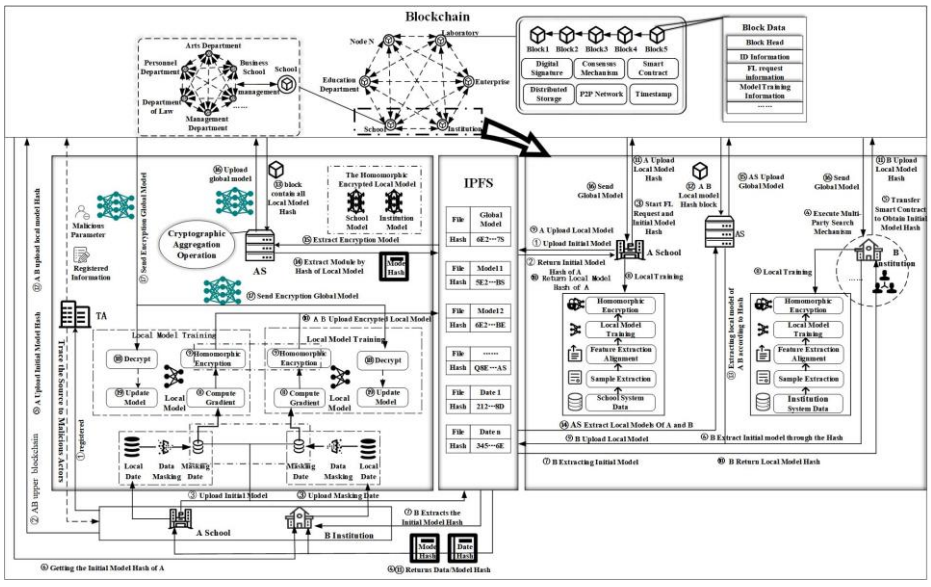


Figure 1. Build a data sharing system based on key techniques of privacy computing.

4. Conclusion

Privacy computing provides a solution for the secure sharing and circulation of educational big data,and realizes the possibility of invisible data and data-invariant models.Because the privacy risks of different links in the life cycle of data are different,the technical solutions of privacy protection are also

different.Therefore,this paper coordinates the privacy protection technologies of different links in the whole life cycle of educational big data in a large framework,and comprehensively uses different privacy protection methods to pursue the real application of technology to the educational scene and fully solve the privacy problem of educational big data.

References

Jiang Yongbo , Zhou Yuan & Feng Tao. (2022). A Blockchain-Based Secure Multi-Party Computation Scheme with Multi-Key Fully Homomorphic Proxy Re-Encryption. Information, 13(10), pp. 481-481.

Peraković Dragan et al. (2022). A Comparative Study of Privacy-Preserving Homomorphic Encryption Techniques in Cloud Computing. International Journal of Cloud Applications and Computing (IJCAC), 12(1), pp. 1-11.

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

The Influence of Flipped Classroom Teaching Mode on the Emotional Relationship between College Teachers and Students—Self-Efficacy of Learning Ability as a Mediator

Chenglong Wang [1*], Xiaoqing Gu [1]
[1] Department of Education Information Technology, East China Normal University
[*] 51214108048@stu.ecnu.edu.cn

Abstract: In order to study the influence of flipped classroom teaching mode (FCTM) on the emotional relationship between college teachers and students (TSER), and to explore the internal mechanism of the role of self-efficacy of learning ability (LASE) as a mediator, 283 college students from mainland China were investigated with the scale of LASE and TSER. The results show that: (1) there are significant differences in LASE and TSER between FCTM and the traditional classroom teaching mode, and the relevant situation is better under FCTM; (2) there is a significant correlation between teaching mode and TSER, however, after excluding the influence of LASE, there is no significant correlation between teaching mode and TSER; (3) LASE has a significant and complete mediating effect on the influence of FCTM on TSER. *This study will be beneficial* to *the research on the application effect of flipped classroom and the construction of harmonious teacher-student emotional relationship.*
Keywords: flipped classroom teaching mode, college teacher-student emotional relationship, self-efficacy of learning ability, mediating effect

1. Introduction

With the deep integration of information technology and education, flipped classroom has become a hot spot in the global education community in recent year. As a teaching mode that converts classroom lectures and individual instruction into self-study before class and in-depth communication in class, the flipped classroom reconstructs the roles of teachers and students, which makes the construction of new teacher-student relationship in the information environment face new opportunities and challenges. In fact, the assumptions and orientations of teacher-student emotional relationships are implicit in many scholars' discussions of flipped classroom. At the same time, the subjective judgments of learners' self-confidence in their own learning ability, i.e., self-efficacy of learning ability *(LASE)* occupies an important position in both the application practice of flipped classroom teaching mode *(FCTM)* and the research on teacher-student emotional relationship *(TSER) (*Li, 2021*)*. Therefore, this study aims to explore the influence of *FCTM* on *TSER*, and to investigate the mechanism of the effect using *LASE as a mediator*.

2. Methodology

Based on relevant theoretical basis and measurement tools, entries were censored and expertly validated resulting in the study's LASE scale (7 questions, belonging to the generality sub-dimension of the academic self-efficacy scale) and TSER scale (12 questions, divided into 3 dimensions of satisfaction,

trust and commitment), which both use a five-point Likert scale. Taking college students from mainland China as the research objects, a total of 283 questionnaires were distributed online and 268 valid questionnaires were returned, with an effective rate of 94.70%. Among them, the sample size of the experimental group (the selected course adopts FCTM) is 101 (37.69%), and that of the control group (the selected course adopts the traditional classroom teaching mode) is 167 (62.31%). After data collection, t-tests, nonparametric correlations, partial correlations, and regressions were conducted using SPSS 24 statistical software. The internal consistency coefficients of the above two scales were 0.935 and 0.934, respectively.

3. Results

The differences in LASE and TSER under teaching mode were analyzed using independent samples t-test. The results showed that there was a significant difference in LASE between FCTM and the traditional classroom teaching mode ($p<0.05$), and LASE in FCTM (Mean =3.801, SD=0.824) was higher than that in the traditional classroom teaching mode (Mean=3.451, SD=0.804); there was a significant difference in TSER between FCTM and the traditional classroom teaching mode ($p<0.05$), and TSER in FCTM (Mean=3.881, SD=0.733) was better than that in the traditional classroom mode (Mean=3.547, SD=0.750).

Using bivariate correlation, the correlation between teaching mode and TSER was analyzed. The results showed that there was a significant correlation between the two ($r=0.199$, $p<0.01$). The partial correlation between teaching mode and TSER was analyzed, controlling for LASE. The results showed that there was no significant correlation between teaching mode and TSER after controlling for LASE, i.e., excluding the effect of LASE ($p>0.05$).

According to the test procedure of mediating effect (Bollen, 1989), using regression analysis, the total effect c, indirect effect ab and direct effect c′ were obtained in three steps, and the mediating effect test was carried out. The results showed that teaching mode significantly and positively predicted TSER ($\beta=0.334$, $p<0.05$) as well as LASE ($\beta=0.350$, $p<0.05$); teaching mode and LASE significantly and positively predicted TSER ($\beta=0.093$, 0.688, $p<0.05$) with an explanation rate of 58.3%. When the mediator LASE was included by the regression model, the path coefficient (direct effect c′) between teaching mode and TSER was not significant, and it can be concluded that LASE has a significant full mediating effect on the influence of FCTM on TSER.

4. Discussion and Conclusion

This study confirmed the positive effect of FCTM on TSER, and LASE has a significant mediating effect. The most important feature of the flipped classroom is that it enhances communication and cooperation between students and teachers during class time, tightening the emotional bonds shared by teachers and students through intense interaction and collaborative forms of learning. In fact, students in the flipped classroom truly lead the learning process through discussion, debate, and group collaboration, enhancing their idea sharing and participation, helping to improve students' problem-solving skills, making them active learners and participants, and in turn improving LASE. When students have better LASE, they tend to actively communicate and interact with teachers, while teachers will understand students' learning better, provide personalized guidance, and promote the development of TSER (Yildiz Durak, 2018). This has provided insights into the application of flipped classroom and other intelligent teaching modes to build a harmonious emotional relationship between teachers and students in the era of "Internet+": (1) identify students' characteristics and teaching needs, and coordinate the provision of

smart equipment and teachers; (2) build a new field of flipped classroom and implement smart instruction scientifically; (3) use intelligent technology to strengthen the supply of diversified teaching contents; (4) provide formative feedback to internalize and transmit knowledge through teacher-student emotional interaction, and stimulate the development of students' autonomy, empathy and ability.

References

Bollen, K. A. (1989). *Structural equations with latent variables*. John Wiley & Sons.

Li, R. (2021). The role of teacher-student interpersonal relations in flipped learning on student engagement. *Frontiers in Psychology*, *12*, 741810.

Yildiz Durak, H. (2018). Flipped learning readiness in teaching programming in middle schools: Modelling its relation to various variables. *Journal of Computer Assisted Learning*, *34*(6), 939-959.

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

How Are Digital Games Used to Promote Collaborative Language Learning? A Review of Empirical Studies

Jing Ren [1], Xiaohui Song [1], You Su [1*]

[1]Beijing University of Posts and Telecommunications

[*] suyou@bupt.edu.cn

Abstract: *The current review analyzes recently empirical studies that implement digital games for fostering collaboration among language learners from 2018 to 2022. The results revealed that English was the most commonly examined target language. Few studies have examined digital game-based collaborative reading. In addition, three research themes have been identified, namely, language acquisition, acquisition and psychological status, and knowledge acquisition.*

Keywords: digital game-based language learning, collaborative learning, language acquisition, interaction competence

1. Introduction

Digital game-based collaborative language learning (DGCLL) refers to a multi-player game in which language learners could play together to develop language proficiency. In recent studies, few studies have reviewed the effectiveness of digital games in collaborative language learning contexts. Therefore, this review study aims to answer the following research question: what research contexts and participants, target languages, target language skills and research themes have been investigated in the selected articles?

2. Methodology

The literature search of this study focuses on four top-tier journals that specialize in technology-enhanced language learning, namely, *Computer Assisted Language Learning, ReCALL, System, and Language Learning and Technology*. The author searched the websites of these four journals with the keywords “digital game”, “language learning” and “collaborative learning” with the publication dates set from 2018 to 2022. Finally, 16 studies were retained for the systematic review and analysis in this study.

3. Results

As shown in Table 1, the results revealed that: 1) the majority of past research investigated digital game-based collaborative language learning among university students; 2) English was the most commonly examined target language; 3) As for target language skills, comprehensive language ability has gained the most attention in recent years. However, there is a dearth of reported empirical studies on reading skills.

Table 1. The overview of the selected studies

Author and year	Contexts and participants	Target languages	Target language skill	Main findings
Grimshaw & Cardoso (2018)	20 university students	English	Speaking	Showing DGCLL's positive influence on anxiety and willingness to communication.
Pitura & Pacut (2018)	25 secondary students	English	Comprehensive learning	Revealing DGCLL's positive effect on digital literacy, teamwork and language learning.
Yamazaki (2018)	11 university students	Japan	Speaking	Improving participants' communicative competencies through DGCLL.
Chen et al. (2019)	20 university students	English	Vocabulary	Indicating the positive influence of DGCLL on learners' perceptions and learning performance.
Loewen et al. (2019)	9 university students	Turkish	Comprehensive learning	Enhancing L2 measures in DGCLL contexts.
Vazquez-Calvo (2020)	5 gamers	Catalan	Translation	Revealing the importance of metalinguistic discussion for language learning in DGCLL environments.
Cheng et al. (2020)	11 university students	English	Comprehensive learning	Supplying more authentic and cross-cultural communication via digital games.
Hong et al. (2020)	96 secondary students	English	Grammar	Indicating the positive effect of DGCLL on language learning.
Lee & Park (2020)	40 college students	English	Comprehensive learning	Promoting meaningful language learning by collaborative digital game.
Chang et al. (2021)	113 primary students	Chinese	Writing	Promoting students' textual cohesion and writing attitude in DGCLL contexts.
Eryigit et al. (2021)	95 university students	Turkish	Vocabulary	Showing positive perceptions towards their learning process in DGCLL contexts.
Soyoof et al. (2021)	160 freshmen	English	Vocabulary	Improving vocabulary acquisition in DGCLL contexts.
Roohani & Vincheh (2021)	150 learners	English	Vocabulary	Integrating gaming applications into the English teaching curriculum to enhance foreign language lexical learning.
Wu (2021)	61 students	English	Comprehensive learning	Promoting students' learning attitudes, learning satisfactions and learning achievements in DGCLL contexts.
Yang et al. (2021)	96 fifth graders	English	Comprehensive learning	Enhancing students' learning with FLA in collaborative DGBL.
Liu (2022)	120 university students	English	Comprehensive learning	Suppling continuous communication between team members via collaborative digital game.

Figure 1 indicates the research themes of the selected studies. The first research theme is promoting language acquisition with DGCLL tasks. For example, Grimshaw and Cardoso (2018) indicated that players who used *Spaceteam* ESL outperformed those without digital games. The second main research theme is affective and motivation in DGCLL. For instance, Wu (2021)'s study revealed that learning motivation improved when the learners played *Pokemon Go* to learn languages together. Thirdly, researchers have also examined knowledge acquisition in DGCLL.

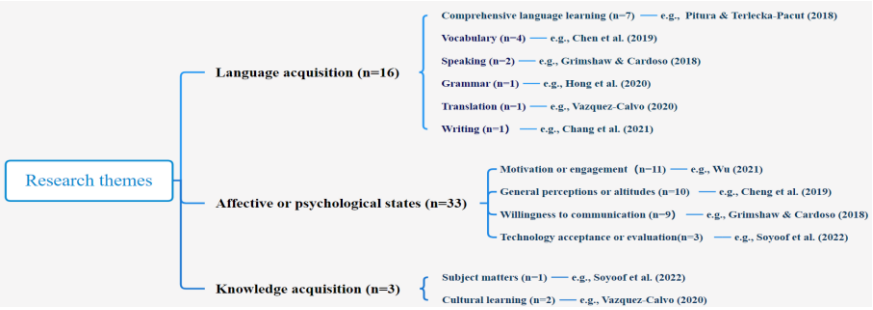


Figure 1. Research themes of the selected studies

References

Chang, W., Liao, C., & Chan, T. (2021). Improving children's textual cohesion and writing attitude in a game-based writing environment. *Computer Assisted Language Learning, 34*(1-2), 133-158.

Chen, C., Liu, H., & Huang, H. (2019). Effects of a mobile game-based English vocabulary learning app on learners' perceptions and learning performance: A case study of Taiwanese EFL learners. *ReCALL, 31*(2), 170-188.

Cheng, J., Chen, C., & Kent, S. (2020). Task engagement, learner motivation and avatar identities of struggling English language learners in the 3D virtual world. *System, 88*, 102168.

Eryigit, G., Bektas, F., Ali, U., & Dereli, B. (2021). Gamification of complex morphology learning: the case of Turkish. *Computer Assisted Language Learning*, 1-29.

Grimshaw, J., & Cardoso, W. (2018). Activate space rats! Fluency development in a mobile game-assisted environment. *Language Learning & Technology, 22*(3), 159-175.

Hong, J., Hwang, M., Liu, Y & Tai, K. (2020): Effects of gamifying questions on English grammar learning mediated by epistemic curiosity and language anxiety, *Computer Assisted Language Learning*, 1-25.

Lee, S & Park, M. (2020). Reconceptualization of the context in language learning with a location-based AR app. *Computer Assisted Language Learning, 33*(8), 936-959.

Liu, I. (2022). Gamified mobile learning: effects on English learning in technical college students. *Computer Assisted Language Learning*, 1-24.

Loewen, S., Crowther, D., Isbell, D., Kim, K., Maloney, J., Miller, Z., & Rawal, H. (2019). Mobile-assisted language learning: A Duolingo case study. *Recall, 31*(3), 293-311.

Pitura, J., & Terlecka-Pacut, E. (2018). Action research on the application of technology assisted urban gaming in language education in a Polish upper-secondary school. *Computer Assisted Language Learning, 31*(7), 734-763.

Roohani, A & Vincheh, M (2021). Effect of game-based, social media, and classroom-based instruction on the learning of phrasal verbs. *Computer Assisted Language Learning*, 1-25.

Soyoof, A., Reynolds, B., Shadiev, R., & Vazquez-Calvo, B. (2022). A mixed-methods study of the incidental acquisition of foreign language vocabulary and healthcare knowledge through serious game play. *Computer Assisted Language Learning*, 1-34.

Vazquez-Calvo, B. (2020). Guerrilla fan translation, language learning, and metalinguistic discussion in a Catalan-speaking community of gamers. *ReCALL*, 1-18.

Wu, M. (2021). The applications and effects of learning English through augmented reality: a case study of Pokémon Go. *Computer Assisted Language Learning, 34*(5-6), 778-812.

Yang, Y., Goh, A., Hong, Y., & Chen, N. (2021). Primary school students' foreign language anxiety in collaborative and individual digital game-based learning. *Computer Assisted Language Learning*, 1-21.

Yamazaki, K. (2018). Computer-assisted learning of communication (CALC): A case study of Japanese learning in a 3D virtual world. *Recall, 30*(2), 214-231.

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

Study on the improvement of children's English learning ability supported by artificial intelligence

Liqiao Nong[1], Geping Liu[2]*

[1]School of Humanities Education, Guangxi Polytechnic of Construction

[2]Faculty of Education,Southwest University

*979620226@qq.com

Abstract:This paper conducts an empirical study on the human-computer co-development teaching model based on the comprehensive perspective of education, ecology and linguistics to improve English learning ability of primary school students. Quasi-experimental research was conducted on the experimental group and the control group. The experimental results show that the English achievement and English learning ability of primary school students are significantly improved.

Keywords: Artificial intelligence; English learning ability; Empirical Research

1. Introduction

For today's education system, cultivating students' higher-order thinking and learning ability is not only the key for students to cope with future challenges, but also one of the goals of education and teaching reform. In the era of artificial intelligence, how to improve students' learning ability is an important issue that researchers need to pay close attention to.

2.Teaching Model of Human-Computer Symbiosis

The deep integration of intelligent technology and classroom teaching is to provide "digital native" learners with an intelligent learning environment and personalized learning style conducive to the development of their English learning ability, see figure 1. Therefore, this study builds a human-computer symbiosis teaching model of AI-assisted English teaching based on the human-computer symbiosis mechanism.

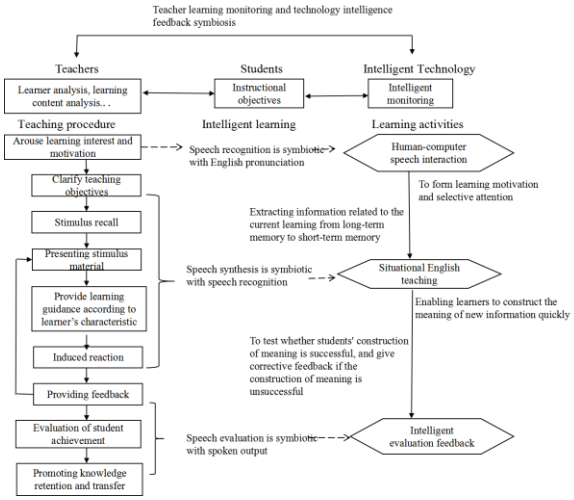


Figure 1. Human-computer symbiotic teaching model.

3. Experimental Design

The implementation of this teaching experiment aims to verify the effect of human-computer symbiosis teaching model. The implementation of the experiment is divided into three different stages, which can be seen in figure 2.

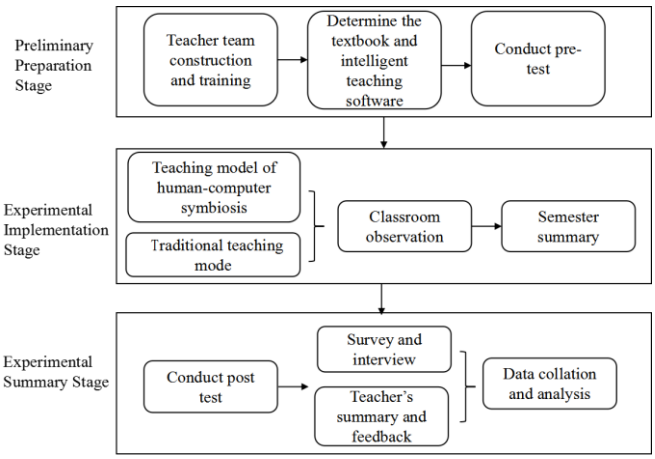


Figure 2. Experimental design of human-computer symbiosis teaching experiment.

6. Experimental Results and Analysis

After the teaching experiment of human-computer symbiosis teaching model are applied, students in experimental class's learning ability is significantly higher than the control class in four dimensions of language skills(W2), comprehensive language application ability(W3), learning strategy(W4) and Higher order cognitive ability(W5), see table 1.

Table 1. Comparison of five dimensions of English learning ability.

Classes		W1	W2	W3	W4	W5
Control class	Mean	28.2791	26.5349	26.5349	27.0930	26.8372
	Standard mean error	0.69592	0.80237	0.64900	0.76526	0.69161
	Standard deviation	4.56344	5.26149	4.25580	5.01813	4.53521
	Minimum value	15.00	15.00	17.00	16.00	16.00
	Maximum value	35.00	35.00	35.00	35.00	35.00
	Total	1216	1141	1141	1165	1154
Experimental class	Mean	28.7674	29.0930	28.9070	29.7209	29.7907
	Standard mean error	0.62598	0.56260	0.62330	0.58148	0.65089
	Standard deviation	4.10480	3.68920	4.08723	3.81305	4.26814
	Minimum value	20.00	17.00	21.00	18.00	17.00
	Maximum value	35.00	35.00	35.00	35.00	35.00
	Total	1237	1251	1243	1278	1281

7. Conclusion

The human-computer symbiotic teaching model provides students with personalized learning methods and diversified learning resources, and helps students build up their interest and motivation in independent learning. Students' learning ability were improved after experiment.

References

Bruce Joyce, et al. (2014). Teaching Model (8th Ed.) [M]. LAN Ying, et al., trans. Beijing: : China Renmin University Press, 5-9.

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

The Impact of Students' Self-regulated Learning on Their Satisfaction with Online General Elective Courses

Huan Huang[1*], Shuai Yuan[2], Ruoxing Xiong[3]
[1] School of Education, South-Central Minzu University
[2] College of Computer, Hubei University of Education
[3] School of Education, Guizhou Normal University
[*] huanghuan@mail.scuec.edu.cn

Abstract: *Online general elective courses are being adopted by more and more university in China, because they can largely increase the amount and types of general elective courses but at a very low cost. Although online courses have existed few years, the online general elective courses for the full-time university students in China only emerged in recent years. As an emerging thing, do students adapt to it and be satisfied with it, and what factors affect their satisfaction? In order to explore these problems, this study investigated Chinese university students' satisfaction with the online general elective courses, and deeply analyzed the impact of students' self-regulated learning on their learning satisfaction. The results suggest that 1) students' satisfaction with online general elective courses is at a medium level, which needs to be further improved; 2) students' self-regulated learning in online general elective courses is also not high, which has a significant positive impact on their learning satisfaction. Especially the persistence dimension and help-seeking dimension are significant predictors of learning satisfaction. According to these conclusions, we discuss the implications for administrators and teachers to develop online general elective courses.*

Keywords: Learning Satisfaction, Online General Elective Courses, Self-regulated Learning, Regression Analysis

1. Introduction

General elective course is an indispensable part of undergraduate curriculum system in the China. It is a main way to implement general education, and its importance is self-evident. In recent years, with the rapid development of online education, many colleges and universities in the China are incorporating some online courses into their general elective curriculum system, which are usually called online general elective courses. They are different from the massive open online courses (MOOCs), and are also different from the small private online courses (SPOCs). Different from MOOCs, the university students will get some credits after they finish the required tasks and pass the exams. Furthermore, the exams are usually offline and organized by their own universities. Different from SPOCs, the online general elective courses are designed and developed by the professors of other universities, and the students generally can't interact with the professors. They can only interact with the assistants in their own universities.

The emergence of online general elective courses not only largely enriches the amount and types of general elective courses, but also makes the learning more flexible. Students can not only study the excellent courses provided by professors from many famous universities all over the world, such as the Harvard, MIT, Tsinghua, Peking, but they can also study these courses at anytime and anywhere. Although online general elective courses have such advantages, it is a new emerging thing for the

university students in the China, and university students may not be satisfied with them. Learner's satisfaction is the reflection of student's perception of learning experience (Kuo et al., 2014). It is not only an essential outcome for learners, but also a very important variable for teachers, because it is a crucial indicator for teachers to improve their instruction (Hew et al., 2020). Furthermore, as the general elective courses, learner's satisfaction is even more important than the academic achievement because their contents are usually not difficult, and their goal is to stimulate students' interest and expand their knowledge scope. Therefore, it is very valuable to investigate the university students' satisfaction with the online general elective courses, and deeply explore the important factors that influence this satisfaction.

Self-regulated Learning (SRL) was proposed by the American educational psychologist Zimmerman (Zimmerman, 1989). It is an active and constructive learning process, in which learners are actively set goals, constantly adjust those goals, and regulate their cognition, motivation and behaviors to archive their goals in constrained learning environments (Pintrich, 2000). Zimmerman described SRL as a cyclical process that includes three phrases: forethought, performance, and self-reflection (Zimmerman, 2000). In the forethought phase, learners are involved in task analysis and self-motivation beliefs. In the performance phrase, learners are involved in self-control and self-observation. In the self-reflection phrase, learners are involved in self-judgment and self-reaction. These phases repeat in a cyclical manner throughout the learning process. Recently many studies have shown that the use of self-regulated learning strategies is significantly related to online academic success (Broadbent, Poon, 2015; Li et al., 2020). Inspired by this conclusion, we suppose that students' self-regulated learning in online general elective courses may have an important impact on their learning satisfaction. Therefore, this study aims to investigate students' satisfaction with the learning of online general elective courses, and explore whether students' self-regulated learning has significant impact on their learning satisfaction. If so, how does the students' self-regulated learning affect their learning satisfaction?

2. Literature Review

In the last decades, many studies have explored the factors that influence students' satisfaction with online learning in different context. The main factors that were considered are student's demographic features, online learning self-efficacy, self-regulated learning, and three types of interaction (learner-content interaction, learner-instructor interaction and learner-leaner interaction), environmental variables.

Many studies explored the impact of student's gender, age, grade, initial computer skills, learning style on their satisfaction with the online courses. For example, Hong (2002) explored the impact of students' gender, age, scholastic aptitude, learning style and initial computer skills on their satisfaction with a Web-based course. Cakir (2014) also explored the impact of students' gender, age, computer literacy level, Internet accessibility, and distance learning experience on their satisfaction with the online courses for the Nursing Complementary Undergraduate Program. Cole et al. (2014) conducted a three-year study of graduate and undergraduate students' level of satisfaction with online instruction at one university. In this study, they also analyzed the impact of students' gender, age, level of study on their satisfaction with the fully online courses.

In addition to students' demographic features, student's online learning self-efficacy and self-regulated learning were explored by many studies. For example, Puzziferro (2008) explored the impact of students' online technologies self-efficacy and self-regulated learning on their satisfaction with the online courses. Kuo et al. (2014) explored the predict power of students' Internet self-efficacy, self-regulated learning and three types of interaction on their satisfaction with the online courses. Shen et al.

(2013) explored the dimensions of online learning self-efficacy and their influence on students' satisfaction with the online courses. Alqurashi (2019) recently also explored the impact of students' online learning self-efficacy and three types of interactions on their satisfaction with the online courses.

Interaction is important in all forms of education, regardless of whether technology is involved (2014). Based on the transactional distance theory (1989), many studies explored the impact of learner-content interaction, learner-instructor interaction, and learner-learner interaction on student's satisfaction on the online courses. For example, Hong (2002) explored the impact of student-student interaction and student-instructor interaction on student's satisfaction with a Web-based course. Kuo et al. (2014) tested the predicting power of the three types of interaction on student's satisfaction with the online courses. Alqurashi (2019) also explored the impact of the three types of interaction on student's satisfaction with the online courses and got the similar results.

As the implementation of online course largely depends on the Internet and online learning platform, therefore some studies have also explored the influence of some environmental variables on learners' satisfaction with the learning of online courses. For example, Sun et al. (2008) analyzed the critical factors that affect students' online learning satisfaction from six dimensions, which are learners, teachers, courses, technology, design and environment. Cole et al. (2014) summarized the factors contributed to learner's satisfaction and dissatisfaction with the fully/partially online course from learners' responses to the open-ended question: what made your experience with the online course/s satisfactory or unsatisfactory?

From the above studies, we can see that only few studies have explored the influence of students' self-regulated learning on their satisfaction with the online courses. Furthermore, the results of these studies are inconsistent, and they do not deeply explore the impact of various dimensions of self-regulated learning. Based on these two research gaps, this study investigates students' satisfaction with the learning of the online general elective courses in university, and explores the impact of various dimensions of students' self-regulated learning on their satisfaction.

3. Methods

3.1. Participants

The participants of this study were 131 undergraduates from the South-Central Minzu University. In recent years, this university has purchased a batch of high-quality online general elective courses that deployed on the *Erya* platform (http://erya.mooc.chaoxing.com/) and the *Treenity* platform (https://www.zhihuishu.com/) for undergraduates. These online courses further enrich the number of general elective courses and improve the quality of general education. All undergraduate students are required to learn two of these online elective courses before graduation. Usually, students will complete this task during their first year or second year. In this study, 150 undergraduate students were selected from the undergraduate students in grade one or grade two by using the snowballing non-random sampling method, and their self-regulated learning and satisfaction with the online general elective courses were investigated. However, 19 of the questionnaires were invalid, so the final participants were the remaining 131 students.

3.2. Instruments

This study involves two variables: participants' self-regulated learning engagement in the online general elective courses and their satisfaction with the learning of these online general elective courses.

According to the connotation of these two variables, the Self-regulated Online Learning Questionnaire (SOL-Q) designed by Jansen et al. (2017) was adopted to measure the participants' self-regulated learning engagement in the online general elective courses, and the Online Learning Satisfaction Questionnaire for College Students (OLS-Q) designed by Li et al. (2020) was adopted to measure students' satisfaction with the learning of these online general elective courses.

The SOL-Q is a questionnaire that was designed to measure learners' self-regulated learning engagement in a fully online learning context like MOOCs. It was designed based on the MSQL (Pintrich et al., 1991), MAI(Schraw, Dennison, 1994), LS (Warr, Downing, 2000) and OSLQ (Barnard et al., 2009). The questionnaire consists of 36 items which involve five dimensions: metacognitive skills, time management, environment construction, persistence and academic help. The overall reliability coefficient of SOL-Q is 0.968, and the reliability coefficient of each dimension is also greater than 0.800, indicating the reliability of SOL-Q is good and the measurement results are reliable.

The OLS-Q was developed to investigate college students' satisfaction with the online learning experience during the COVID-19 pandemic period (Li et al., 2020). Based on the three types of learning behaviors in the online learning environment, Li et al. designed the questionnaire from three dimensions: academic performance, interpersonal gain and intention feedback. The questionnaire contains 10 items in total. The overall reliability coefficient of the questionnaire is .93, and the reliability coefficients of the three dimensions are all greater than .80, which indicates that the questionnaire has good reliability and the measurement results are reliable.

3.3. Data analysis

After obtaining the participants' self-regulated learning engagement in the online general elective courses and their satisfaction with the learning of these courses, we input the data into the SPSS 24.0, and conducted the descriptive statistics, correlation analysis, regression analysis and variance analysis on the data in turn. In order to understand university students' satisfaction with the learning of these online general elective courses and their self-regulated learning engagement in these courses, we first conducted the descriptive statistical analysis of these two variables. In order to test the influence of students' self-regulated learning engagement on their learning satisfaction, Pearson correlation coefficient was used to analyze the correlation between students' self-regulated learning engagement and their learning satisfaction, and step-step regression was used to analyze the effect of students' self-regulated learning engagement on their learning satisfaction.

4. Results

4.1. Descriptive statistics

Table 1 summarizes the minimums, maximums, means, and standard deviations of variables measuring students' self-regulated learning engagement in the online general elective courses and their satisfaction with the learning of these courses.

Table 1. Descriptive results of students' self-regulated learning engagement and their satisfaction with the learning

	Min	**Max**	**M**	**SD**
Metacognitive skill	1.33	7.00	4.50	1.14
Time management	1.00	7.00	4.80	1.18
Environment construction	1.40	7.00	4.66	1.11
Persistence	1.00	7.00	4.51	1.23
Academic help-seeking	1.00	7.00	4.46	1.20

Evaluation of academic performance	1.00	5.00	3.60	.75
Evaluation of interpersonal gain	1.00	5.00	3.52	.81
Intention feedback	1.50	5.00	3.63	.73

Note: Min = Minimum; Max = Maximum; M = Means; SD = Standard Deviation.

As shown in Table 1, all the means of subscales of students' self-regulated learning are in between 4 to 5, and all the means of subscales of students' satisfaction are in between 3 to 4, which indicates that students' self-regulated learning engagement in the online general elective courses and their satisfaction with the learning are only at the medium level. Furthermore, the means of time management and environment construction are slightly higher than the means of metacognitive skills, persistence, and academic help seeking, which suggests that students used more time management strategies and environment construction strategies when learning the online general elective courses.

4.2. Correlation analysis

In order to analyzing the influence of students' self-regulated learning engagement on their satisfaction with the learning of online general elective courses, we first used the Pearson correlation coefficient to examine the correlation relationship between students' self-regulated learning engagement and their satisfaction. Before calculating the Pearson correlation coefficients, we respectively aggregated all the subscales of self-regulated learning engagement and all the subscales of learning satisfaction, and further used normal Q-Q diagram to test whether the two aggregated variables conform to the normal distribution. The normal Q-Q diagrams indicate that students' self-regulated learning engagement and learning satisfaction conform to the normal distribution, which means Pearson correlation coefficient can be used to analyze the correlation relationship between these two aggregated variables. Through calculating the Pearson correlation coefficient of self-regulated learning engagement and learning satisfaction, it is found that there is a moderate and significant positive correlation between them ($r = 0.59$, $p < .01$).

In order to further explore the correlation relationship between each subscale of the self-regulated learning and learning satisfaction, we also used the normal Q-Q diagram to test the normal distribution characteristic of each subscale of self-regulated learning. The results suggested that they all conform to the normal distribution. After that, we further calculated the Pearson correlation coefficients between each subscale of self-regulated learning and the learning satisfaction. The results are shown in Table 2.

Table 2. Correlation coefficient between each subscale of self-regulated learning and the aggregated learning satisfaction

	MS	**TM**	**EC**	**P**	**AHS**	**S**
Metacognitive skill (MS)	1					
Time management (TM)	.40**	1				
Environment construction (EC)	.71**	.51**	1			
Persistence (P)	.59**	.37**	.58**	1		
Academic help-seeking (AHS)	.72**	.42**	.72**	.61**	1	
Satisfaction (S)	.51**	.25**	.47**	.58**	.53**	1

Note: ** $p < .01$.

As can be seen from Table 2, five components of the self-regulated learning are positively correlated with the learning satisfaction. More specially, four components are moderately correlated with the learning satisfaction. Besides, we can also find that the components of self-regulated learning are positively co-correlated with each other, and some Pearson correlation coefficients are up to the high level. For example, the Pearson correlation coefficient between the metacognitive skill and the academic help-seeking, between the academic help-seeking and the environment construction are up to .72 ($p < .01$);

the Pearson correlation coefficient between the metacognitive skill and the environment construction is up to .71 ($p < .01$).

4.3. The impact of self-regulated learning engagement on students' satisfaction

After analyzed the correlation relationship among the five components of the self-regulated learning engagement and student's satisfaction, we further used stepwise multiple regression to explore to what extent the five components of the self-regulated learning engagement affect student's satisfaction with the learning of online general elective courses. The results are shown in Table 3 and Table 4.

Table 3. Summary of the regression models

Models	R	R^2	Adjusted R^2	Std. Error of the Estimate	F	Sig.
1	.58[a]	.34	.33	.55	65.72	.000
2	.62[b]	.39	.38	.53	40.02	.000

a. predictors: (constant), persistence

b. predictors: (constant), persistence, academic help-seeking

Table 4. Coefficients of regression models

Models		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	2.16	.18		1.81	.000		
	Persistence	.32	.04	.58	8.11	.000	1.00	1.00
2	(Constant)	1.89	.20		9.50	.000		
	Persistence	.23	.05	.41	4.74	.000	0.63	1.59
	Academic help-seeking	.16	.05	.27	3.13	.002	0.63	1.59

According to the Table 3, we can see that the final regression model (model 2) includes two predictors, which are persistence and academic help-seeking; the F value of the model is 40.02 and its significance level is $p<.01$. That means learning persistence and academic help-seeking are two important predictors of student's satisfaction with the learning of online general elective courses. More specially, the persistence explains about 33% of the variance of student's satisfaction (adjusted R2 = 0.33), and when it was combined with the academic help-seeking, the explained variance of student's satisfaction increased to 38% (adjusted R2 = 0.38). It can be further seen from the Table 4 that the predictive power of learning persistence is greater (beta = 0.41, $p < 0.005$) than the academic help seeking (beta = 0.27, $p < 0.005$), and there is no multicollinearity between them (Tolerances > .2 and VIFs < 10). Based on the above analysis, the regression equation is: Student's satisfaction =1.89+ 0.23 × Persistence + 0.16 × Academic help-seeking.

5. Discussion

Online general elective course is a new type of course in many universities in China. It is seen as an ideal and low-cost way to enrich the quantity and improve the quality of general courses in universities. However, as a new type of curriculum, whether the students are satisfied with it and what factors affect student's satisfaction, these questions are of great significance for improving the design of online general elective courses. Student's self-regulated learning is considered to be a critical factor that influences their performance. This study explores the following two questions.

First, we investigated students' satisfaction with the online general elective courses and their self-regulated learning level in such courses. According to the survey results of 131 university students, we found that students' satisfaction with the online general elective courses is at a medium level as a whole, and their self-regulated learning level in such courses is also at a medium level as a whole, and both of which need to be further improved. This result is similar to the results of (Kuo et al., 2014). Kuo et al. (2014) investigated the satisfaction and self-regulated learning level of 222 students who participated in a fully online credit course, and found that their self-regulated learning level and learning satisfaction were at the medium level. Specifically, they used 5 items to assess learners' satisfaction with the course, and used the metacognitive self-regulation subscale come from the MSQL to assess the extent to which the planning, monitoring, and regulating strategies learners utilized during learning. They found that both the average score of students' satisfaction and self-regulated learning were slightly higher than the midpoint. Based on our results and the results of (Kuo et al., 2014), it can be concluded that although online general elective courses have many advantages, students' satisfaction and their self-regulated learning level are not very high, and both of them need to be improved.

Second, we explored the relationship between students' satisfaction and their self-regulated learning level during learning the online general elective courses. According to the results of correlation analysis and regression analysis, we found that students' self-regulated learning level during learning the online general elective courses has a significant impact on their learning satisfaction. Especially, the two dimensions of persistence and academic help-seeking have a significant positive predictive effect on students' learning satisfaction. This result is not consistent with the results of (Kuo et al., 2014). Kuo et al. used hierarchical linear modeling (HLM) to explore the influence of three types of interaction, Internet self-efficacy, and self-regulated learning on students' learning satisfaction. They found that students' self-regulated learning level was not a significant predictor of their learning satisfaction, only the interaction between students and content and the interaction between students and teachers had significant predictive effects on their learning satisfaction. The reason for this inconsistency may be that the measurements of self-regulated learning are different. Self-regulated learning is a concept with multiple dimensions. Our study includes five dimensions, which are metacognitive skills, time management, academic help, persistence and environment construction. However, Kuo et al. only selected the metacognitive self-regulation dimension when measure students' self-regulated learning. In this study, metacognitive self-regulation is also not the significant predictor of students' satisfaction. Based on above analysis, it can be seen that when exploring the relationship between self-regulated learning and learning satisfaction, we should go deep into all dimensions of self-regulated learning.

6. Conclusions

At present, online general elective courses have been widely used in colleges and universities in China, and have become one of the trends of general education reform. As a new type of course, online general elective courses have many advantages, but there are also some problems in the implementation process, which drive us to pay attention to the satisfaction of students and its influencing factors. This study takes students in a university as samples to deeply investigate students' satisfaction with online general elective courses, and takes students' self-regulated learning level as a breakthrough point to deeply explore its influence on students' learning satisfaction. The results show that students' satisfaction with online general elective courses is at a medium level. Students' self-regulated learning level during learning the online general elective courses is also not high, which has a significant impact on students' learning satisfaction. More specifically, among the five dimensions of self-regulated learning, persistence

and academic help-seeking have significant predictive effects on students' learning satisfaction. These conclusions have two implications for the design of online general elective courses: first, online general elective courses have good adaptability to all kinds of students, and colleges and universities should continue to promote its construction; Second, appropriate strategies and methods (Wong et al., 2019) should be adopted to effectively intervene students' self-regulated learning level and promote students to use more self-regulated strategies, so as to improve students' satisfaction with the learning experience of online general elective courses.

Although this study has drawn some useful conclusions, it also has some limitations. In terms of sample selection, the samples are from only one university, and the sample size is not very large, which will have a certain impact on the reliability of the research conclusion. In terms of data collection, self-report questionnaire was adopted in this study, and there may be some deviation in the self-regulated learning level collected by this method (Winne, 2010). In terms of analysis methods, this study only uses correlation analysis and multiple linear regression to analyze the impact of each dimension of self-regulated learning on learning satisfaction, which can not reveal the mechanism of each dimension of self-regulated learning on learning satisfaction. In view of the above problems, future research can further expand the range and increase the number of samples, use learning analysis technology to measure students' self-regulated learning level from the log data of online learning platform (Li et al., 2020), and use structural equation model, multi-layer linear regression and other methods to deeply explore the impact of students' self-regulated learning on their learning satisfaction with online general elective courses.

Acknowledgements

This study was supported by the Fundamental Research Funds for the Central Universities, South-Central Minzu University (No. CSQ22008).

References

Alqurashi E. (2019). Predicting student satisfaction and perceived learning within online learning environments. *Distance Education, 40*,133-148.

Barnard, L., Lan, W. Y., To, Y. M., Paton, V. O., & Lai, S.-L. (2009). Measuring self-regulation in online and blended learning environments. *The Internet and Higher Education, 12*, 1-6.

Broadbent, J., Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education, 27*, 1-13.

Cakir, O. (2014). The factors that affect online learners' satisfaction. *The Anthropologist, 17(3)*, 895-902.

Cole, M., Shelley, D., Swartz, L. (2014). Online instruction, e-learning, and student satisfaction: A three year study. *The International Review of Research in Open and Distributed Learning, 15(6)*, 110-119.

Hew, K. F., Hu X., Qiao C., & Tang Y. (2020). What predicts student satisfaction with MOOCs: A gradient boosting trees supervised machine learning and sentiment analysis approach. *Computer & Education, 145*, 103724.

Hong, K.-S. (2002). Relationships between students' and instructional variables with satisfaction and learning from a Web-based course. *The Internet and Higher Education, 5(3)*, 267-281.

Jansen, R. S., Leeuwen1, A., Jansen, J., Kester, L., & Kalz, M. (2017). Validation of the self-regulated online learning questionnaire. *Journal of Computing in Higher Education, 29*, 6-27.

Kuo, Y.-C., Walker, A. E., Schroder, K. E. E., Belland, B. R. (2014). Interaction, Internet self-efficacy, and self-regulated learning as predictors of student satisfaction in online education courses. *The Internet and Higher Education, 20*, 35-50.

Li, Y.-Y., Zhang, H.-M., Zhang, H.-Z. (2020). Model construction and empirical test of college students' satisfaction with online learning during epidemic prevention and control period: Based on the survey of 15 universities in Shanghai. *Open Education Research, 26(4)*, 102-111.

Li, Q., Baker, R., & Warschauer, M. (2020). Using clickstream data to measure, understand, and support self-regulated learning in online courses. *The Internet and Higher Education, 45*:100727.

Moore, M. G. (1989). Three types of interactions. *The American Journal of Distance Education, 3(2)*, 1-6.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In Boekaerts, M., Pintrich, P. R., Zeidner, M. (Eds.), *Handbook of self-regulation* (1st ed., pp. 452-520), San Diego: Academic Press.

Pintrich, P. R., Smith, D. A., Garcia, T., & McKeachie, W. J. (1991). A manual for the use of the Motivated Strategies for Learning Questionnaire (MSQL). Michigan, National Centre for Research to Improve Postsecondary Teaching and Learning, Ann Arbor.

Puzziferro,M. (2008). Online technologies self-efficacy and self-regulated learning as predictors of final grade and satisfaction in college-level online courses. *The American Journal of Distance Education, 22*, 72-89.

Schraw, G.; Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19(4)*:460-475.

Shen, D., Cho, M.-H., Tsai C.-L., Marra, R. (2013). Unpacking online learning experiences: Online learning self-efficacy and learning satisfaction. *The Internet and Higher Education, 19*, 10-17.

Sun, P. C., Tsai,R. J., Finger, G., Chen, Y. Y., & Yeh, D. (2008). What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction. *Computers & Education, 50(4)*, 1183-1202.

Warr, P., & Downing, J. (2000). Learning strategies, learning anxiety and knowledge acquisition. *British Journal of Psychology, 91(3)*, 311-333.

Winne,P. H. (2010). Improving measurements of self-regulated learning. *Educational Psychologist, 45(4)*: 267-276.

Wong, J., Baars, M., Davis, D., Zee, T. V. D., Geert-JanHouben & 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.

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology, 81(3)*, 329-339.

Zimmerman, B. J. (2000). Attaining self-regulation: a social cognitive perspective. In Boekaerts, M., Pintrich, P. R., Zeidner, M. (Eds.). *Handbook of Self-Regulation* (1st ed., pp. 13-40), San Diego: Academic Press.

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

Effects of a Motion Capture and Interactive Learning System-based Classroom on Students' learning Performance, Motivation and Self-efficacy in Dance Lessons

Liujie Xu[1], Jing Wu[1], Jingdong Zhu[1]
[1]School of Educational Technology and Science, Zhejiang University of Technology, Hangzhou, China
xulj2004@zjut.edu.cn

Abstract: It is important to provide one-to-one tutoring and timely feedback for students learning dance. Artificial intelligence-based motion capture technology is considered to improve the efficiency of dance training. However, there is a lack of empirical research on the application of motion capture technology in dance education in real educational scenarios. In this study, a motion capture and interactive learning system (MCILS) was applied in a dance classroom, and a MCILS-based learning approach was developed. A total of 40 participants from a vocational school in Hangzhou studying dance were randomly divided into two groups: MCILS-based learning (experimental group) and traditional teaching (control group). The results indicate that the MCILS-based learning approach can significantly improve students' dance skills and self-efficacy. However, there is no significant effect on students' learning motivation. In addition, students with higher learning motivation and self-efficacy benefit more from MCILS-based learning as compared to those with lower learning motivation and self-efficacy. Finally, this study demonstrates that it is possible to create a smart classroom by applying MCILS to the teaching activities of dance education, physical education, and other disciplines.

Keywords: *Interactive learning environment, technology assisted dance education, feedback visualization, learning performance, motion capture system*

1. Introduction

Motion capture can be defined as a technical process of recording three-dimensional (3D) positioning and direction information of moving objects, which are effective methods to digitize action behavior. With the continuous improvement of high-speed camera performance and machine learning algorithms, motion capture technology (MCT) is gradually becoming mature and able to recognize coherent actions. MCT can be used for motion comparison and evaluation and identifying emotions through people's actions. Many researchers combine human bone monitoring technology, human posture recognition method, and tracking technology to design efficient motion capture algorithms combined with virtual reality technology, and constantly improve motion capture accuracy. In dance training, MCT can capture the actions of students and compare them with the standard dance template actions to provide intuitive feedback for students and intuitive error correction tips for coaches. Some research has been done to improve the performance of the motion capture system in terms of technology and algorithm and testing the accuracy of an algorithm based on dance training. However, few studies show whether the

application of MCT in real classrooms affects students' motion skills and learning performance. Further, there is a lack of empirical research on vocational school students. Therefore, in this study, we develop a motion capture and interactive learning system (MCILS) and implement for dance teaching and learning. We hope this study can broaden the empirical evidence of MCT in the classroom and provide the basis for future research.

2. Literature review

Dance requires students to master active literacy, conceptualization, self-reflection, and creativity (El Raheb et al., 2019). The objectives of dance teaching are to develop students' dance skills, dance innovation ability, cognitive skills and decision-making ability, and aesthetic awareness. Most dance teachers use the imitation method. The teachers' tutoring, feedback, and students' reflection are essential in this teaching model. However, in a one-to-many dance classroom, a single teacher can provide only minimal tutoring and feedback to all students. Therefore, it is necessary and relevant to apply MCT and interaction visualization technology. Interaction visualization technology supports students to learn dance by observing and imitating intelligent agents (Gao & Xu, 2021a). MCT identifies and evaluates the students' dance movements, and helps learners criticize, analyze, and reflect on their movements (Chan et al., 2011).

Historically, teachers have noticed the importance of visual interaction technology for dance learning. For example, traditionally, mirrors are usually used as visual tools for teachers to demonstrate and provide feedback, and for students to practice and reflect in dance lessons (Andersson, 2014). Students who use the mirror to learn dance can find their own mistakes and adjust them in time. However, when students do complex actions such as jumping and rotating, they cannot observe their actions through the mirror to receive the feedback information, so it is difficult for students to spot and reflect on their mistakes (Ehrenberg, 2010). Therefore, video-based visualization technology has become an essential tool for students to learn dance. Videos can provide more information to compensate for mirrors' shortcomings in information recording. However, Students must remember their errors and correct them in the following practice. This is known as lag feedback, which is characterized by lag and delay. Thus, it is not conducive for students to correct their actions by video feedback. Moreover, mirrors and videos require students to spot their own mistakes. However, due to the limitation of students' professional skills, students cannot spot every incorrect action and comprehensively reflect on their actions.

MCT can make up for the shortcomings of mirror and video teaching, and can be used for art performance with real-time visual motion (Andreadis et al., 2010). It has been applied to dance training practice to evaluate students' dance movements and provide feedback for students (Ofori et al., 2022). Combining MCT with virtual reality technology can support students to detect the difference between target movements and practice movements in real time, and give timely feedback to students. At the same time, it can measure the learning results in real time, which can help students improve their dancing skills and interest in learning (Kyan et al., 2015). The combination of MCT and automated evaluation technology can support students in learning dance, which is recognized and affirmed by dance teachers and experts.

MCP has proven beneficial to dance learning and can improve dance skills. Guo et al. (2022) designed a comprehensive visualization system, DanceVis, which helps in evaluating students' dance movements and the interaction between students and the system. It has the characteristics of objective evaluation, fine-grained analysis, high efficiency, accuracy, and a clear training path. It can track the dynamic changes of groups and individuals from coarse to acceptable, global to local, and various time

dimensions as well. El Raheb et al. (2019) posited that a digital interactive dance system should include functions such as providing students with visual feedback, evaluating motion, and supporting technical methods of dance learning.

Due to its high research value, motion capture has become a general research field in recent years. However, research on the combination of MCT and dance movements is still in its infancy (Ni & Yao, 2021). Although the above literature has reached a consensus on the importance of MCT in dance education, there are no verified empirical studies on the educational effect of using MCT for dance learning in a classroom environment. Therefore, this study aims to explore the effect of an MCILS-based learning approach on vocational school students' dance performance. The research questions of this study are as follows:(1) Can an MCILS-based learning approach improve students' dance skills to a greater extent as compared to traditional dance teaching approaches?(2) Are there differences in dance skills among students with different levels of self-efficacy and learning motivation within an MCILS-based learning environment?(3) Can an MCILS-based learning approach improve students' learning motivation to a greater extent as compared to traditional dance teaching approaches?(4) Can an MCILS-based learning approach improve students' self-efficacy to a greater extent as compared to traditional dance teaching approaches?

3. MCILS-based dance learning environment

Motion capture can be used in the teaching of motion training. Based on the motion capture and interactive learning system, we constructed a smart dance learning classroom for dance teaching. MCILS can support students' autonomous learning and support teachers to carry out blended learning. Its function mainly comprises the following aspects.

(1) Learning resources module: As Figure 1 shows, MCILS provides rich dance learning resources for students, including micro-course resources, dance movements education resources, dance action education resources, and artistic temperament training resources. Students can use computer equipment and smart devices to watch the learning resources.

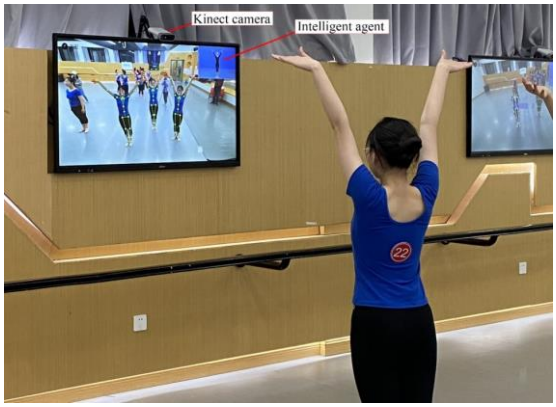


Figure 1. Student practices dancing with MCILS.

(2) Students' learning space module: Students can record their dance videos and upload the videos to their own learning space on the MCILS. Peers can watch the dance videos uploaded by each student.

(3) Assisting students' practice dance module: Students can watch dance learning videos in an MCILS-based classroom and imitate dance videos for practice.

(4) Visualizing the evaluation results: The MCILS can evaluate every dance exercise, show students' scores in each step of motion training, and provide students with detailed dance movement evaluation results.

4. Methods

4.1. Second Level Headings

The study sample comprised 40 students from a vocational school who were randomly assigned to either an experiment group (EG) or a control group (CG). There were 19 students in the EG, including 2 boys and 17 girls. The CG had 21 students, including 2 boys and 19 girls. The EG adopted the MCILS-based learning approach, while the CG adopted the traditional dance teaching approach.

4.2. Experimental procedure

The two groups had two consecutive dance classes each week. Each lesson lasted 45 minutes, for a total of 90 minutes. Figure 2 shows the experimental procedure. In the first seven weeks, both groups of students used the traditional teaching approach to learn basic dance theory and dance movement knowledge. In the eighth week, both groups of students conducted a week-long questionnaire survey on dance skills, learning motivation, and self-efficacy. Then, the eight-week teaching experimental stage commenced.

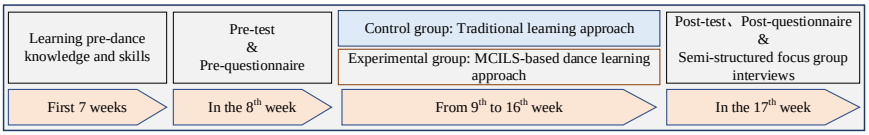


Figure 2. The experiment process.

Before the experiment, the EG students were trained for 45 minutes. The EG was required to be familiar with the functions and specific operations of MCILS, which helped to reduce the novel effects that might affect students' learning when using new technologies. As there was not enough teaching equipment based on MCILS for each student, they were randomly assigned into five groups of three to four students. From the 9th week to the 16th week, two groups of students had two dance lessons every week, separately. Table 1 shows an example of a 90-minute lesson of dance learning activities in an MCILS-based learning environment. The CG was randomly divided into seven groups of three students each. The CG students learnt the same materials through the traditional teaching approach in ordinary classes. Every week, the teacher used Microsoft PowerPoint to present and introduce teaching objectives, syllabus, and new knowledge, and organize the students to use computers and projectors to watch dance leaning videos. Then, each group followed the dance video to practice and observed their dance movements in the mirror. After the dance exercise, the group members discussed their dance movements. Next, the teacher organized all the students to practice and perform a complete dance. The teacher evaluated and tutored the performance of students. Following this, the teacher provided the students with a reflective outline. Then, the students communicated in groups and reflected on their learning. Finally, the teacher summarized the learning content and practice results and gave homework.

Table 1. An example of a 90-minute teaching activity in MCILS-based environment.

Step	Activity	Time
1	The students danced a dance that they learned in the last class, and the teacher evaluated and guided the students.	10 min
2	The teacher introduces the teaching objectives, syllabus, and learning contents of this lesson to the students.	10 min
3	Students watch dance teaching videos and learn new dance knowledge and	10 min

	movements.	
4	Students practice dance moves, check feedback information, and discuss with group members.	15min
5	Teachers check students' dance practice results, explain important knowledge points and dance demonstration.	10 min
6	Students practice new dance moves, check visual evaluation results of MCILS, reflect, and communicate in groups.	15 min
7	Students collectively dance the newly learned dance moves, and teachers evaluate and give feedback.	15 min
8	Teachers summarize and assign homework.	5 min

The control group was randomly divided into seven groups of three students to learn the same material through traditional teaching methods. Every week, the teacher used Microsoft PowerPoint to present and introduce teaching objectives, syllabus, and new knowledge, and organize the students to use computers and projectors to watch dance leaning videos. After the dance practice, the group members exchanged and discussed with each other. After that, the teacher organizes all the students to practice dancing and complete a complete dance. The teacher evaluates and guides the overall performance of the students. And then, the teacher provides the students with a reflection outline, so that the students can communicate in groups and complete the reflection outline. Finally, the teacher summarizes the learning content and practice results, and assigns homework. Except for the MCILS environment, the learning process (including dance learning activities and learning time) was the same in both groups.

After the teaching experiment, all participants were required to conduct a post test of dance skills and complete the self-efficacy and learning motivation questionnaires. Lastly, a semi-structured focus group interview was conducted in the following week.

4.3. Instruments

The instruments used in this study include the dance skills assessment scale, learning motivation scale, self-efficacy questionnaire, and interview outline. The dance skill test was adapted from the dance performance evaluation criteria proposed by Campus (2019). The self-efficacy questionnaire was adapted from Pintrich et al. (1991), including eight items with a 5-point Likert scale. The motivation questionnaire was modified by Wang and Chen (2010) according to the measure proposed by Pintrich et al. (1991). The outline comprised seven issues. All the students were divided into five groups, and each group participated in a 10-minute semi-structured focus group interview.

4.4. Data analysis method

The t-test method was adopted to analyze the differences between the EG and CG in the individual dimensions of the pre-test (see Table 2). No significant difference was found between the EG and CG for the "Dance skills", "learning motivation", or "self-efficacy" dimensions. This indicates that the students were at the same level in dance skills, learning motivation, and self-efficacy. Then the variance is used to analyze the dance skills, learning motivation, and self-efficacy of the students in the EG and CG. The effective quantity $\eta 2$ was calculated to evaluate the significant differences in dance skills, learning motivation, and self-efficacy between the EG and CG. In addition, a two-factor ANOVA was applied to evaluate the impact of the interaction between high and low levels of learning motivation and learning styles in different environments on dance skills, and the impact of the interaction between high and low levels of self-efficacy and learning styles in different environments on dance skills.

Table 2. Independent samples *t* tests on group differences in the pre-test.

	Group	*N*	Mean	*SD*	*t*	*p*
Dance skills	EG	19	8.895	3.4624	1.394	0.177
	CG	21	7.714	1.3470		
Learning motivation	EG	19	4.2632	.52209	1.559	0.129
	CG	21	4.0397	.36096		
Self-efficacy	EG	19	3.9608	.87296	0.877	0.387
	CG	21	3.7440	.66436		

5. Results

5.1. Analysis of Students' Dance Skills

We carried out covariance analysis (ANCOVA) on the score of dance skills in the EG and CG to test the differences between the students' dance skills in different learning environments. We considered the students' pretest dance skills as covariance and post-test dance skills as dependent variables. In the homogeneity of regression test, the F value was not significant (F = 0.774, p = 0.384 > 0.05). The homogeneity assumption of two sets of variances was satisfied. Table 3 shows the relevant results. After excluding the influence of the pretest, there was a significant difference between the EG and CG (F = 7.368, p = 0.010 < 0.05). Further analysis showed that the dance skills of the EG (M = 14.316) were significantly higher than those of the CG (M = 11.333). Thus, it can be seen that the dance teaching in a classroom based on MCILS is better than the traditional classroom teaching method for improving students' dance skills.

Table 3. The ANCOVA result of the dance skills scores.

Groups	*N*	*Mean*	*SD*	*Adjusted M*	*F*	*p*
CG	21	11.333	2.556	11.693	7.368	0.010
EG	19	14.316	3.400	13.918		

To further analyze the influence of learning motivation and self-efficacy on dance skills, participants who scored less than 50% in the pretest of learning motivation and self-efficacy were classified into the high-motivation group and high-self-efficacy group, respectively. In contrast, the remaining participants were put into the low-motivation group and low-self-efficacy group, respectively. Then, with learning motivation or self-efficacy and learning approach together as a fixed factor. We performed a two-way ANOVA to analyze the interaction of MCILS-based learning approach and learning motivation (Table 4), and learning self-efficacy (Table 5). There was no impact of the interaction between motivation and learning approach on the students' dance skills. However, the influence of learning approach and motivation on dance skills was significant. The effect size of the learning approach and motivation was large.

Table 4. The two-way ANOVA result of the dance skills with factors of learning approach and motivation.

Source	*SS*	*df*	*MS*	*F*	*p*	η^2
Learning approach	66.234	1	66.234	8.693	.006	.195
Motivation	58.698	1	58.698	7.704	.009	.176
Learning approach * Motivation	7.867	1	7.867	1.033	.316	.028
Error	274.279	36	7.619			

Table 5. The two-way ANOVA result of the dance skills with factors of learning approach and self-

efficacy.

Source	*SS*	*df*	*MS*	*F*	*p*	η^2
Learning approach	69.744	1	69.744	8.436	.006	.190
Self-efficacy	39.857	1	39.857	4.821	.035	.118
Learning approach * Self-efficacy	2.122	1	2.122	.257	.616	.007
Error	297.643	36	8.268			

According to Figure 3, the dance skills of the students with high learning motivation are higher than that of those with low learning motivation. With an increase in learning motivation, the increase in dance skill of the EG students is significantly greater than that of the CG students. Therefore, in the MCILS-based learning approach, motivation plays a role in influencing the learning effect.

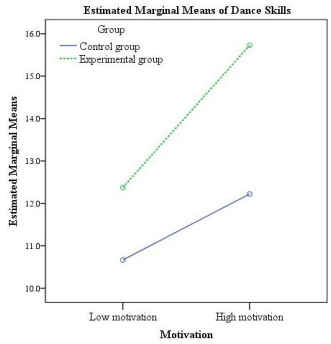


Figure3. Interaction between different motivation and dance skills.

According to Figure 4, the dance skills of students with high self-efficacy were higher than that of those with low self-efficacy. With an increase in self-efficacy, the increase in the dance skill score of the EG students is significantly more significant than that of the CG students. Students with high self-efficacy are more likely to benefit from the MCILS-based learning approach.

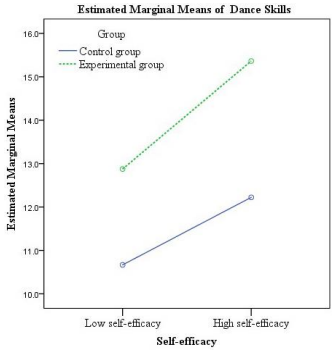


Figure 4. Interaction between different self-efficacy and dance skills.

5.2. Analysis of students' learning motivation

To exclude the influence of the pre-test scores, the learning motivation scores were tested by the homogeneity of the regression coefficients; the F value was not significant ($F = 0.660$, $p = 0.422 > 0.05$), indicating that the hypothesis of homogeneity of the regression coefficients was met and ANCOVA could be performed. Table 6 shows the relevant results. After excluding the influence of pre-test learning motivation, there was no significant difference in learning motivation between the EG and CG, indicating that the traditional and MCILS-based learning approaches did not affect the students' learning motivation. In addition, on the 5-point Likert scale, the average scores of the two groups were all higher than 4, showing that all the students had positive attitudes towards dance learning.

Table 6. The ANCOVA result of the learning motivation.

Groups	*N*	*Mean*	*SD*	*Adjusted M*	*F*	*p*
CG	21	4.214	0.703	4.260	2.454	0.126
EG	19	4.597	0.413	4.546		

5.3 Analysis of Students' self-efficacy

We conducted a different test on the post-test score of the students' self-efficacy to determine the effect of the MCILS-based learning approach. According to the homogeneity test results of the regression test, the F value was not significant (F = 0.858, p =.0.360 >.05). The homogeneity assumption of two sets of variances was satisfied. Therefore, ANCOVA could be performed. Table 7 shows the relevant results. After excluding the effect of the pretest, there were significant differences in self-efficacy between the EG and CG (F = 4.672, p = 0.037 < 0.05). The results showed that the EG (M = 4.257) significantly outperformed the CG (M = 3.774). The findings indicated that the MCILS-based learning approach could have a significant and better effect on improving students' self-efficacy as compared to the traditional learning approach.

Table 7. The ANCOVA result of the self-efficacy.

Groups	N	Mean	SD	Adjusted M	F	p
CG	21	3.774	0.794	3.804	4.672	0.037
EG	19	4.257	0.405	4.223		

6. Discussion and Conclusion

In this study, we proposed an MCILS-based learning approach, which provides students with video learning, activity capture and correction, visual evaluation, and timely feedback, and allows students to reflect on dance learning and practice activities. To explore the effectiveness of this learning approach, we conducted experiments in a secondary vocational school. This study demonstrates that the application of MCILS in teaching can improve students' dance skills. The benefits of MCILS may be due to the following reasons. First, the EG's higher dance skills may result from motion capture and real-time visual feedback. Second, MCILS provides students with dance teaching videos, allowing them to watch and practice repeatedly. Lastly, MCILS supports learning interaction, and the intelligent agent in the MCILS system can act as a learning tutor for students.

Regarding learning motivation, there is no significant difference between the EG and CG. This may be because students desire good academic performance and praise from teachers, parents, and other personnel regardless of an MCILS-based class or a traditional teaching environment. In addition, the study examined the main effects of learning motivation on dance performance. The results showed that students with high learning motivation exhibited better dance performance as compared to those with low learning motivation. In terms of self-efficacy, there was a significant difference between the EG and CG. Specifically, the proposed learning approach could improve students' self-efficacy and increase their confidence in dance learning. In this study, through repeated practice, the students' dance moves gradually became proficient and reached a higher level. The score assigned to the students by the MCILS will gradually increase and they will see their scores continue to increase; thus, their self-confidence will continuously improve as well. Teachers mainly evaluated the CG. Although students can improve the proficiency of dance movements through repeated practice, their self-efficacy will decrease when they get a negative evaluation due to the lagging evaluation of teachers and teachers' summary evaluation.

The experimental results showed that this innovative way of learning improves students' dance skills and enhances their self-efficacy. This study provides evidence for teachers and educators to make teaching decisions in dance teaching or other disciplines, and can support teachers and educators to carry out teaching activities based on artificial intelligence. In addition, in the use of technology to carry out teaching, improving students ' motivation and self-efficacy can enhance the role of technology in improving students ' academic performance. However, there are some limitations in this study. First, the sample size was small. Therefore, caution is needed when applying and generalizing the findings. Future studies will expand the sample size to further validate the proposed learning approach. Second, the experimental setup only in dance learning domain. It is worth implementing this innovative learning approach in other learning activities. Finally, this study explored the effects of MCILS-based learning approach in terms of dance skills, motivation and self-efficacy. It is not clear how students learn dance. In the future, we aim to explore what kind of behavioral patterns and rules they exhibit when learning activities using an intelligent dance learning system and explore how students learn.

References

Andersson, N. (2014). Assessing dance: a phenomonological study of formative assessment in dance education. *Nordic Journal of Art and Research,* 3(1).

Andreadis, A., Hemery, A., Antonakakis, A., Gourdoglou, G., Mauridis, P., Christopoulos, D., & Karigiannis, J. N., (2010). Real-time motion capture technology on a live theatrical performance with computer generated scenery, *2010 14th Panhellenic Conference on Informatics*, 2010, pp. 148–152.

Chan, J. C. P., Leung, H., Tang, J. K. T., & Komura, T. (2011). A Virtual Reality Dance Training System Using Motion Capture Technology. *IEEE Transactions on Learning Technologies*, 4(2), 187-195.

Ehrenberg, S. (2010). Reflections on reflections: Mirror use in a university dance training environment. *Theatre, Dance and Performance Training*, 1(2), 172–184.

El Raheb, K., Stergiou, M., Katifori, A., & Ioannidis, Y. (2019). Dance interactive learning systems: A study on interaction workflow and teaching approaches. *ACM Computing Surveys*, 52(3), Article 50.

Gao, Y., & Xu, D. (2021a). Application of posture recognition service system based on information fusion smart sensor in dance training. *Journal of Sensors,* Article 4284249.

Guo, H., Zou, S. C., Xu, Y. L., Yang, H., Wang, J., Zhang, H. X., & Chen, W. (2022). DanceVis: toward better understanding of online cheer and dance training. *Journal of Visualization*, 25:159-174.

Kyan, M., Sun, G. Y., Li, H. Y., Zhong, L., Muneesawang, P., Dong, N., Elder, B., & Guan, L. (2015). An Approach to Ballet Dance Training through MS Kinect and Visualization in a CAVE Virtual Reality Environment. *ACM Transactions on Intelligent Systems And Technology*, 6(2).

Ni, S. S., & Yao, D. W. (2021). Sports Dance Action Recognition System Oriented to Human Motion Monitoring and Sensing. *Wireless Communications & Mobile Computing*.

Ofori, E. K., Wang, S. J., & Bhatt, T. (2021). Validity of Inertial Sensors for Assessing Balance Kinematics and Mobility during Treadmill-Based Perturbation and Dance Training. *Sensors*, 21(9).

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

Does the position and mode of the embedded questions in an instructional video affect learning outcome? A moderation effect of cognitive load level.

Yun pan [1], Yi Zhang[1], Yuewen Shi[1], Jiumin Yang[1*]
[1]Central China Normal University
[*] yjm@ccnu.edu.cn

Abstract: *The study investigates how the position and mode of the embedded questions in an instructional video impact learning and then tests mental effort as a moderator of the relationship between the position of embedded questions and learning from video lectures. Students (N = 100 undergraduate and graduate students) were assigned to watch one of four videos in a 2(mode: segmented, concentrated) ×2 (position: pre-embedded, post-embedded) between-subject design. The ANOVA results of ANOVAs showed that students in the post-embedded group reported higher satisfaction and mental load than those in the pre-embedded group. Also, the students in the segmented group achieved better than those in the concentrated group. The findings have practical implications: embedded post-questions in video learning can increase learning satisfaction and cognitive load than video-embedded pre-questions, and segmented embedded questions help students achieve better scores compared to concentrated embedded questions. For the students who reported less mental effort, post-embedded questions performed better intention scores than pre-embedded ones.*

Keywords: embedded questions, learning performance, learning satisfaction, instructional video, cognitive load

1. Introduction

Videos are widely used in education but sometimes contribute little to learning (MacHardy & Pardos, 2015). Many factors may contribute to the ineffectiveness of video learning, for example, the long length and poor interaction of videos (Mayer et al., 2018). To solve these problems, some studies came up with the ideas that segmented and embedded questions in the video (Koumi, 2015; Lim & Wilson, 2018).

Embedded questions are an important form of interaction in instructional videos (Tweissi, 2016). Studies have shown that embedded questions in the video can improve learning performance and effectiveness of video-recorded lectures (van, Mei & Böckmann, 2021, Callender & McDaniel,2007; Lawson et al., 2006). According to the position, the embedded questions can be divided into pre- and post-questions (Yaohui X et al., 2021). Although studies compared the pre-questions with the post-questions in a classroom setting, whether the results will be similar still need to be explored in video learning.

Despite the position, there are two common forms of embedded questions according to the mode: concentrated and segmented (van, Mei & Böckmann, 2021). Several studies have found that embedded segmented and concentrated questions can improve students' learning performance and be an effective

study method (Callender & McDaniel,2007; Lawson et al., 2006). However, whether there are some differences between the concentrated embedded questions and segmented embedded questions in the videos is not clear.

No study has explored how the position and mode of embedding questions in instructional videos impact learning performance. In the present study, we investigated the impacts of embedded questions' position and mode on learning performance in an instructional video.

1.1. The position of embedding questions

Some studies compared the pre-questions with the post-questions in different situations and gained different results (Jafarigohar et al., 2015; Frase, Patrick, & Schumer, 1970). One study found no significant difference in word recognition and comprehension for elementary school students in the post-questions or pre-questions condition (Frase, Patrick, & Schumer, 1970). However, another study found that the pre-and post-question groups differed significantly in text learning; in other words, students in the pre-question group outperformed the post-question (Jafarigohar et al., 2015). The different conditions might explain the possible inconsistent results, and the position of embedded questions may cause the different learning performance, which was still unclear.

1.2. The mode of embedded questions

Among the existing studies, the mode of embedded questions was generally the embedding segmented questions and concentrated questions, which may lead to video segmentation. Compared to studying the material again, students who learned the video with embedded segmented questions performed significantly better in the learning performance (Jafarigohar et al., 2015). Students in segmented question conditions positively affected learning, but no significance was found in learning satisfaction (Schmitz, 2020). Concentrating questions can also promote students' engagement and learning performance (Carpenter, Rahman & Perkins, 2018). However, it is still unclear whether segmented questions can achieve better learning performance than concentrated ones, and it still needs further exploration.

The position and the mode of embedded questions can influence the cognitive load (Tweissi, 2016; Carpenter, Rahman & Perkins, 2018). Thus, the effects of position and mode on cognitive load and learning performance should be further investigated when they are utilized together.

1.3. Mental effort as moderator

Although the various advantages of embedded questions, some empirical evidence does not find the benefits of embedded questions for students' learning (Meutstege, 2019; Hirsch et al., 2020). The discussed conflicting results suggest that different mechanisms and boundary conditions determine the effects of the embedded questions on learning. For example, contrary to the benefits of the embedded questions, the study did not find any significant difference in knowledge gain between the students who studied the embedded question video and traditional video (Meutstege, 2019). A possible reason may be the low cognitive load level in both embedded and without questions.

Cognitive load was defined as a multidimensional construct representing the load that performing a particular task imposes on the learner's cognitive system (Paas & van Merriënboer, 1994). According to Paas and van Merriënboer, mental load and mental effort are the aspects of conceptualized cognitive load (1994). Mental load related to task or environmental demands. Mental effort refers to cognitive load when learners work on a task (Paas et al., 2003).

Since investing enough cognitive resources into learning tasks is the basic condition to ensure successful information processing, if students invest few mental efforts, they would not benefit from any design. The mental effort will be investigated as a potential moderator for the current investigation.

1.4. Research questions and hypotheses

The experiment tested the effects of the embedded questions' position and embedded mode on students' learning using neural measures of cognitive load. Based on segmenting theory and previous studies, our questions are as follows:

Q1: Do the position, and embedded mode affect students' learning satisfaction, learning performance, and cognitive load?

Q2: Does mental effort moderate the effects of the embedded position on students' learning performance?

2. Method

2.1. Participants and experimental design

Participants were 100 undergraduate and master's degree students. A 2 (Pre-embedded questions vs. Post-embedded questions) × 2 (concentrated appearance vs. segmented appearance) between-subjects experimental design was used to control differences in the learning materials across the experimental conditions.

2.2. Materials

The main knowledge point of this study is the concept and application of genetic engineering and its related steps. The teaching content covers two topics: the concept and application of genetic engineering and the four steps of genetic engineering, and lasts 6 min 33s. The content of the video is preceded and followed by relevant questions. The only difference between pre-and post-embedded groups was inserting questions before or after the seven knowledge points. The only difference between the concentrated and segmented groups was presenting the seven questions simultaneously or after each part of the lecture, including five multiple-choice items and two fill-in-the-blank.

2.3. Measures

2.3.1. Demographic questionnaire

Participants were asked to report gender, grades, major, and age.

2.3.2. Prior knowledge test

The prior knowledge test was developed by the instructors according to the textbook and was designed to examine the students' mastery of basic knowledge related to genetic engineering. The eight items included five multiple-choice items with five options (20 points) and three fill-in-the-blanks. (*Cronbach'a* α = 0.73). There was no difference between the four groups on this prior knowledge test (F (3, 96) = 1.05, p = .373, η_p^2 = 0.032).

2.3.3. Learning satisfaction questionnaire

The learning satisfaction questionnaire was adapted from the questionnaire developed by Jiumin Yang (2014). The questionnaire was a 3-item measure designed to assess students' satisfaction. Participants answered the items on a five-point Likert scale ranging from 1 (extremely low) to 9 (extremely high). The final score was calculated in summary (*Cronbach'a* α = 0.87).

2.3.4. Learning performance tests

The instructor created this test, and all items were derived from the learning material presented in the video. The retention test included four multiple-choice items (8 points) and six fill-in-the-blank items (9 blanks, 18 points). The total possible score was 26 points. The transfer test included a multiple-choice item (2 points) and seven fill-in-the-blank items (11 blanks, 22 points). The final score was the summary of each item. The total possible score was 24 points (*Cronbach'a* α = 0.85).

2.3.5. Cognitive load questionnaire

The cognitive load questionnaire was developed by Paas and Van Merrienboer (1994). The questionnaire was a 2-item measure designed to assess students' cognitive load mental load, and mental effort. Each dimension includes one item. Participants answered the items on a nine-point Likert scale ranging from 1 (extremely low) to 9 (extremely high). The final score was calculated in summary (*Cronbach'a* α = 0.74).

2.4. Procedure

The study was conducted in a laboratory and took approximately 30min. The procedure is shown in *figure 1*.

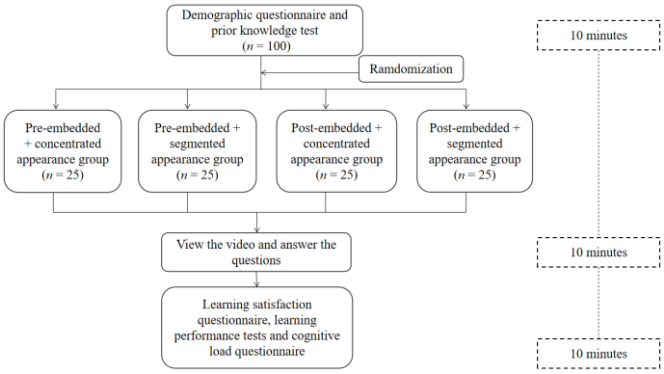


Figure 1: The procedure of the experiment

2.5. Data analysis

Descriptive statistics for all variables are shown in Table 1.

Table 1: Means and standard deviations of all variables

Dependent variable	Pre-embedded				Post-embedded			
	concentrated appearance		segmented appearance		concentrated appearance		segmented appearance	
	M	SD	M	SD	M	SD	M	SD
Prior knowledge test	5.76	3.28	4.56	3.77	4.72	3.60	4.16	2.51
Learning satisfaction	10.76	3.44	10.88	2.56	12.32	2.27	12.04	2.25
Retention	16.32	5.94	18.96	4.20	18.16	5.56	19.68	3.86
Transfer	13.76	6.04	15.84	4.43	16.72	6.40	16.08	5.49
Mental load	5.28	1.84	5.96	1.84	5.40	1.68	6.12	1.88
Mental effort	6.32	1.91	6.80	1.56	7.20	1.23	6.92	1.98

3. Results

3.1. Learning satisfaction questionnaire

To test Q1, the two-way ANOVA was conducted with the learning satisfaction score as the dependent variable. The main effect was found for a position, F (1,98) = 6.48, p = .013, η_p^2 = 0.06. However, no main effect was found for embedded mode, F (1,98) = 0.02, p = .881, η_p^2 < .001. No interaction was observed, F (1,98) = 0.14, p = 0.71, η_p^2 = 0.001. The results suggested that compared to the pre-embedded groups, students in the post-embedded groups reported higher learning satisfaction, MD = 1.36, p = .013.

3.2. Learning performance tests

3.2.1. Retention scores

To test Q1, the two-way ANOVA was conducted with the retention score as the dependent variable. No main effect was found for a position, F (1,98) = 1.66, p = .201, η_p^2 = 0.02). However, embedded mode had a main effect, F (1,98) = 4.38, p = 0.039, η_p^2 = 0.04. No interaction was observed, F (1,98) = 0.32, p = .574, η_p^2 = 0.003. The results suggest that compared to the concentrated appearance groups, students in the segmented appearance group performed better, MD = 2.08, p = .039.

3.2.2. Transfer scores

To test Q1, the two-way ANOVA was conducted with the transfer score as the dependent variable. No main effect was found for a position, F (1,98) = 2.01, p = .159, η_p^2 = 0.02), or for embedded mode, F (1,98) = 0.41, p = .525, η_p^2 = 0.004. No interaction was observed, F (1,98) = 1.45, p = .231, η_p^2 = 0.02. That is, the position and embedded mode did not significantly impact students' transfer performance.

3.3. Cognitive load questionnaire

3.3.1. Mental load

To test Q1, the two-way ANOVA was conducted with the mental load score as the dependent variable. No main effect was found for a position, F (1,98) = 0.16, p = .687, η_p^2 = 0.002. However, embedded mode had a main effect, F (1,98) = 4.07, p = .046, η_p^2 = 0.04. No interaction was observed (F (1,98) = 0.003, p = .954, η_p^2 < 0.001. The results suggest that the students in the post-embedded group significantly increased learners' mental load, MD = 0.70, p = .046.

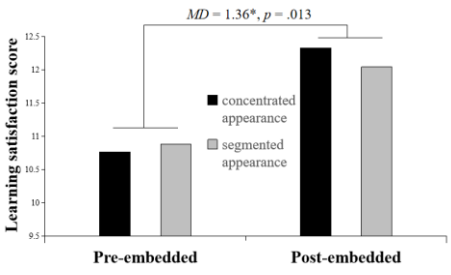


Figure 2: Main effect on learning satisfaction score (*p < .05)

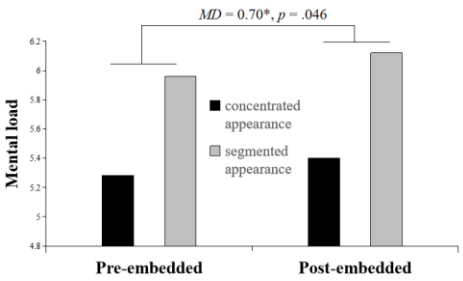


Figure 3: Main effect on mental load (*p < .05)

3.3.2 Mental effort

To test Q1, the two-way ANOVA was conducted with the transfer score as the dependent variable. No main effect was found for a position, F (1,98) = 2.18, p = .143, η_p^2 = 0.02, or for embedded mode, F (1,98) = 0.09, p = .768, η_p^2 = 0.001. No interaction was observed, F (1,98) = 1.26, p = .265, η_p^2 = 0.01. That is, the position and embedded mode had no significant effect on students' mental effort.

3.4. Mental effort as moderator

Before the moderation analysis, Correlation analyses of mental load, mental effort learning satisfaction, and learning performance were conducted. As shown in Table 2, the Pearson correlation analysis of the dependent variables showed that mental effort was significantly correlated with retention score (r = 0.33, p = 0.001 < 0.05) and transfer scores (r = 0.30, p = 0.003 < 0.05); mental effort and learning satisfaction were significantly correlated (r = 0.25, p = 0.012 < 0.05).

Table 2: Correlation coefficient of all variables

	1	2	3	4	5
1. Learning Satisfaction	1.00				
2. Retention scores	0.13	1.00			
3. Transfer score	0.03	.74**	1.00		
4. Mental load	0.09	0.07	0.02	1.00	
5. Mental effort	0.25	.33**	.30**	0.01	1.00

* p < 0.05, ** p < 0.01

To test Q2, one separate hierarchical multiple regression model was tested to evaluate the moderating effect of mental effort on the relationship between position (pre-embedded question vs. post-embedded question) and retention scores. Concerning mental effort, the position × retention scores interaction explained significant additional variance in mental effort, providing evidence of moderation. Simple slope tests were conducted for each outcome variable to interpret these interactions. As shown in Figure 4, when they reported lower mental effort levels, the students in the post-embedded group achieved better than those in the pre-embedded group (t = 3.63, p = .001). Still, there was no significance when they reported high mental effort levels.

The result suggested that when the students invested lower mental effort into the task, the post-embedded questions could help them achieve better retention scores than the pre-embedded questions, but when they invested higher mental effort into the task, there were no differences.

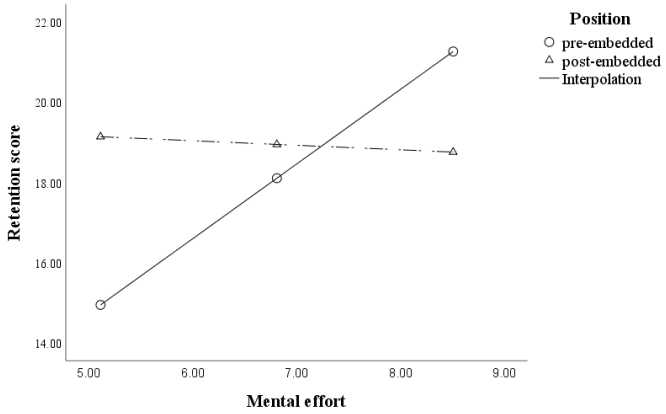


Figure 4: Interaction effect of position × mental effort on retention score

4. Discussion

This study examined the effects of position and embedded mode in questions on video learning. The results revealed that post-embedded questions in instructional videos significantly increased learner satisfaction and cognitive load. Also, the embedded segmented questions significantly increased

retention scores compared to the embedded, concentrated questions. Furthermore, we found that students' mental effort moderated the position of embedded questions.

Compared to pre-embedded questions, post-embedded questions significantly increased learner satisfaction and mental effort, consistent with previous research (Altinpulluk, 2020). The videos with post-embedded questions are more helpful for students to deepen their memory and summarize their knowledge. In contrast, pre-embedded questions are prone to cognitive conflict due to the students' own a priori knowledge level, which leads to lower learner satisfaction.

Compared to concentrated questions, segmented embedded questions can significantly improve students' learning performance, which is consistent with previous research findings (Schacter & Szpunar, 2015). Also, the results are consistent with cognitive load theory. Embedding questions segmented are helpful for students to decrease their cognitive load and perform better than the concentrated questions.

However, this study did not find an interaction effect between the position and embedded mode. The main impact of embedded mode was also nonsignificant. A possible explanation is that the embedded questions type is not appropriate. A previous study found the main effect of embedded mode with embedding the "WH" question (Jafarigohar et al., 2015). It is reasonable to infer that the result might be due to the embedded question type. Future work can test the embedded different question types to examine our findings.

Further, we conducted a moderation analysis of the mental effort, and the results may explain the inconsistent results (Meutstege, 2019; Hirsch et al., 2020). We found that the mental effort moderated the position of embedded questions' effect on students' retention scores. Specifically, when students engage less (in less mental effort), the post-embedded question will benefit them more than the pre-embedded question. Still, when they engaged higher (in higher cognitive load level), there was no significance between the two groups. The possible reason might be that the pre-embedded questions tend to contradict students' prior knowledge and fail to achieve better retention scores when the mental effort is low.

4.1. Limitation, implication, and future directions

This study has investigated the embedded questions in video learning and found that the position and mode of embedded questions can promote students' learning, as shown by the fact that segmented embedded questions encourage students' learning performance, increasing cognitive load, and post-embedded questions are beneficial to enhance students' learning satisfaction compared to the concentrated embedded. Based on this, the study suggests that in the video learning process, teachers can enhance students' learning satisfaction and performance by using post-embedded segmented questions.

Although the present research offers valuable insights into QEVs, there are two limitations. First, we have focused only on procedural knowledge but not declarative knowledge in the study. Previous studies have shown that embedded questions in declarative knowledge can improve students' learning performance (Schacter & Szpunar, 2015). Thus, future research could consider the impact of embedding videos with different knowledge types in segmented videos. Second, this study was conducted in a laboratory setting, and whether it can be replicated in a real-world situation remains to be tested.

References

Altinpulluk, H., Kilinc, H., Firat, M., & Yumurtaci, O. (2020). The influence of segmented and complete educational videos on the cognitive load, satisfaction, engagement, and academic achievement levels of learners. Journal of Computers in Education, 7(2), 155-182.

Callender, A. A., & McDaniel, M. A. (2007). The benefits of embedded question adjuncts for low and high structure builders. Journal of Educational Psychology, 99(2), 339.

Carpenter, S. K., Rahman, S., & Perkins, K. (2018). The effects of prequestions on classroom learning. Journal of Experimental Psychology: Applied, 24(1), 34.

Frase, L. T., Patrick, E., & Schumer, H. (1970). Effect of question position and frequency upon learning from text under different levels of incentive. Journal of Educational Psychology, 61(1), 52.

Hirsch, S. E., Chow, J. C., Randall, K. N., Nemer, S. L., & McKown, G. (2020). Evaluating the Effect of Embedded Responses in Multimedia-Based Instruction With Preservice Teachers. Behavioral Disorders, 46(1), 18-28.

Jafarigohar, M., Hemmati, F., Soleimani, H., & Jalali, M. (2015). An investigation of the role of explicit and implicit instruction in second language acquisition: A case of English embedded question. International Journal of Applied Linguistics and English Literature, 4(3), 98-108.

Koumi, J. (2015). Learning outcomes afforded by self-assessed, segmented video–print combinations. Cogent Education, 2(1), 1045218.

Lawson, T. J., Bodle, J. H., Houlette, M. A., & Haubner, R. R. (2006). Guiding questions enhance student learning from educational videos. Teaching of Psychology, 33(1), 31-33.

Lim, K. H., & Wilson, A. D. (2018). Flipped learning: Embedding questions in videos. Mathematics Teaching in the Middle School, 23(7), 378-385.

MacHardy, Z., & Pardos, Z. A. (2015). Evaluating the Relevance of Educational Videos Using BKT and Big Data. International Educational Data Mining Society.

Mayer, R. E., Howarth, J. T., Kaplan, M., & Hanna, S. (2018). Applying the segmenting principle to online geography slideshow lessons. Educational Technology Research and Development, 66, 563-577.

Meutstege, K. (2019). Keep One's Attention at Recorded Lectures: What effect do embedded questions in recorded lectures have on mind wandering and knowledge gain? (Master's thesis, University of Twente).

Paas, F. G., & Van Merriënboer, J. J. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. Educational psychology review, 6, 351-371.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. (2016). Cognitive load measurement as a means to advance cognitive load theory. In Educational psychologist (pp. 63-71). Routledge.

Schacter, D. L., & Szpunar, K. K. (2015). Enhancing attention and memory during video-recorded lectures. Scholarship of Teaching and Learning in Psychology, 1(1), 60.

Schmitz, W. H. G. (2020). Embedded questions in text and video-based lectures (Master's thesis, University of Twente).

Tweissi, A. (2016). The effects of embedded questions strategy in video among graduate students at a middle eastern university (Doctoral dissertation, Ohio University).

van der Meij, H., & Böckmann, L. (2021). Effects of embedded questions in recorded lectures. Journal of computing in higher education, 33(1), 235-254.

Yang Jiumin. (2014). The influence of teachers on the learning process and effectiveness in online video courses (Doctoral dissertation, Wuhan: Central China Normal University).

Yaohui, X. I. E., Jiumin, Y. A. N. G., Zhongling, P. I., Chenyan, D. A. I., & Caixia, L. I. U. (2022). Effects of In-video Questions and Feedback on Learning Performance. Frontiers of Education in China, 17(1).

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

Detecting Online Learners' Moods before Learning and Exploring Adaptive Strategies

Shuai Yuan[1,2], Huan Huang[3*],
He Yang[1,2], Rui Hou[4]
[1] College of Computer, Hubei University of Education, Wuhan, China
[2] Hubei Cooperative Innovation Center for Basic Education Information Technology Service, Wuhan, China
[3] School of Education, South-Central University for Nationalities, Wuhan, China
[4] College of Computer Science, South-Central University for Nationalities, Wuhan, China
[*] huanghuan@mail.scuec.edu.cn

Abstract: *Many studies in psychology have suggested that learners' emotions have an important effect on their learning performance. In the past decade, many researchers have conducted in-depth studies on learners' emotion modeling, especially proposing many techniques to detect learners' emotions. However, most of these methods focus on learners' emotions during learning, and few focus on learners' moods before learning. In this paper, we propose a modeling method and an automatic detection approach for learners' moods before learning. The experimental results show that the proposed detection approach can detect some moods to some extent, especially happy moods. Based on the proposed modeling method and detection approach, this study also explores the application scenarios of learners' moods before learning and proposes an adaptive navigation mechanism and adaptive resource recommendation mechanism that integrate learners' moods before learning. The proposed modeling method, detection approach and adaptive strategies have important implications for the development of personalized learning systems.*

Keywords: personalized learning, mood detection, adaptive strategies, sentiment analysis

1. Introduction

Personalized learning has always been one of the important goals pursued by educational researchers. To realize this goal, a large number of educational researchers and practitioners have devoted themselves to the research of adaptive learning systems. An adaptive learning system is a hypermedia learning environment that can provide personalized learning support services for learners according to their individual characteristics (Karampiperis & Sampson, 2005). From the perspective of system architecture, an adaptive learning system consists of three key components: a domain knowledge model, learner model and adaptive model (Karampiperis & Sampson, 2005). The learner model stores various characteristics of learners, such as sex, age, highest degree, learning style, and cognitive capability, which are the bases for adaptive learning systems to provide adaptive learning support services. Learners' emotions are also a very important characteristic and should be modeled, which has received extensive attention from many researchers in recent years.

In the past decade, with the development of affect computing (AC) technology, many researchers have conducted in-depth research on learners' emotion modeling, especially proposing many techniques

to detect learners' emotions. According to the data sources on which they depend, they can be categorized as physiological information-based methods, facial expression-based methods, speech intonation-based methods, body posture-based methods, interactive behavior-based methods and multimodal data fusion-based methods (Calvo & D' Mello, 2010). However, most of these methods focus on learners' emotions during learning, and few studies focus on learners' moods before learning. Many studies have shown that learners' moods before learning have an important influence on the subsequent learning process and learning results (Isen et al., 1987; Brand et al., 2007; Cui et al., 2021). Therefore, how to detect learners' moods before learning and how to provide personalized support services to them according to their moods when they enter into the digital learning system are two important problems that need to be deeply considered in the research of modeling learning emotion.

Compared with emotions during learning, modeling moods before learning faces a major challenge, as it is difficult to find a suitable data source. Because the mood before learning is often not caused by learning but by other events occurring before learning, it is difficult to find clues that can be used to infer learners' moods before learning. In view of this problem, this study proposes an approach to model learners' moods before learning based on online social media and discusses adaptive strategies to improve the adaptability of digital learning platforms. The reason why online social media is used to model learners' moods before learning is that existing reports have shown that learners often express their views, attitudes and feelings through online social media, especially among college students (Ortigosa et al., 2014). In addition, with the development of sentiment analysis technology in recent years, many effective sentiment analysis methods have been proposed, which can accurately identify the emotions hidden in texts.

2. Formalization of Learners' Moods before Learning

To enable the digital learning system to detect learners' moods when entering, the first problem that needs to be solved is the formalization of learners' moods before learning. In the field of psychology, researchers have already carried out in-depth research on emotion description and proposed two types of emotion description methods. The first is usually called the classification method, and the second is usually called the multidimensional method (Feidakis et al., 2014). The classification method uses discrete categories to describe people's emotions at a certain time, such as happiness, sadness, and anger. This method assumes that people's emotions are composed of several relatively independent basic emotions and that a variety of compound emotions are formed on this basis. Using discrete emotion categories is a quick, intuitive, and effective way to formalize people's emotions at a certain time. However, how many different emotions there are has not been resolved until now. The multidimensional method uses a continuous set of values to describe people's emotions at a certain time, which each represent a data point in the multidimensional space. Each dimension of the multidimensional space represents an inherent characteristic of emotion, such as polarity and intensity. Obviously, this multidimensional method assumes that emotion is continuous rather than discrete. Based on the above analysis, combined with the actual needs of modeling learners' moods before learning, we have proposed an improved method that combines the classification method and multidimensional method to formalize learners' moods before learning.

Formalization Definition Assume that the learner's mood before learning is denoted by E, then $E = (H, S, A, F)$, in which H represents the intensity of happiness, S represents the intensity of sadness, A represents the intensity of anger, and F represents the intensity of fear. H, S, A, $F \in [0,1]$.

The reason why we propose this method is mainly based on the following considerations. First, the goal of modeling learners' moods before learning is to provide personalized support services for learners, so the classification method is helpful for the design of personalized support strategies. Second, a learner's mood before learning is usually not triggered by learning. Therefore, it is more appropriate to adopt the four common basic emotions proposed by Ortony et al. to describe learners' moods before learning rather than cognitive emotions (Ortony et al., 1988). Third, the previous classification method cannot reflect the intensity of emotion, which is undoubtedly very important when using learners' emotions to provide personalized support services. Therefore, we further adopt four continuous values to represent the intensity of four basic emotions. Fourth, the previous classification method cannot represent the co-occurrence of multiple basic emotions, but this phenomenon objectively exists. Therefore, we use a four-dimensional vector instead of one numerical value to represent a learner's mood before learning.

3. Detecting Learners' Moods before Learning

After designing an appropriate data structure to formalize learners' moods before learning, the second problem that needs to be solved is acquisition of the value of each dimension in a timely manner. Because learners usually express their moods in a timely manner on online social media, such as Twitter and Facebook, we propose analyzing learners' messages on online social media to implement modeling of moods before learning. To mine learners' moods from the messages, we need to use sentiment analysis technology. Sentiment analysis is a text mining technology that uses computers to automatically detect the emotions or opinions contained in text (Paltoglou & Thelwall, 2012).

There are two types of approaches to implement sentiment analysis: machine learning-based approaches and linguistic cue-based approaches (Paltoglou & Thelwall, 2012). Machine learning-based approaches treat sentiment analysis as a classification problem, using classification algorithms to determine the opinion or emotion category contained in texts. Linguistic cue-based approaches mainly use linguistic cues, such as emoticons and emotional words, and linguistic rules to infer the opinions or emotions contained in texts. Because linguistic cue-based approaches do not require large-scale labeled data, only some low-cost language resources, and can simulate the intensity of sentiment, we use this approach to detect the emotions expressed by learners in messages of online social media and then infer learners' recent moods. In the following, we will take *Weibo (https://weibo.com/*) as an example to detail the three important phases of learners' mood detection.

3.1. Identify the messages containing learners' emotions

The messages posted by learners on *Weibo* often involve a wide range of topics. In addition to expressing their feelings, learners will also share their life and work trivia and express their opinions on hot events. Therefore, we first need to identify the messages containing emotional expressions. Based on the observation of a large-scale data set developed by Zhao et al. (Zhao et al., 2012), we summarized some high-confidence discriminant rules to identify emotional messages. These discriminant rules include two categories: filtering rules and determining rules.

The filtering rules can filter out many nonemotional messages. Through observing the nonemotional messages, we summarized four filtering rules. 1) Reposted messages without any comments should be filtered out because learners did not provide any linguistic cues. 2) Messages with fewer than six Chinese words should be filtered out. Although these messages may contain emotional expressions, analysis of this type of message is difficult. 3) Messages containing *URLs* should be filtered out because these

messages are probably summaries of longer objective information on a web page directed by the *URL*. 4) Messages containing full-width square brackets should be filtered out.

In addition to using the above rules to filter out nonemotional messages, we further summarized two determining rules to identify the emotional messages. The first determining rule is that messages containing emoticons or emotional slangs are identified as emotional messages. To identify these messages, we developed a dictionary that contains 13 frequently used emoticons and 32 emotional slang terms from the data set developed by Zhao et al. (Zhao et al., 2012). The second determining rule is that messages containing emotional words or phrases are identified as emotional messages. Emotional words and phrases are important cues to identify emotional messages. To judge whether a message contains emotional words or phrases, we developed a dictionary that includes 596 emotional words and phrases. These emotional words and phrases mainly come from the lexicon developed by the Institute of Computational Linguistics of Peking University and the lexical set for sentiment analysis released by HowNet.

3.2. Analyze the category and strength of emotion in a message

After identifying the emotional messages, it is necessary to further analyze the category and strength of emotion contained in each message. To achieve these goals, four dictionaries have been developed, and a series of rules have been defined based on them. The first dictionary contains frequently used emoticons and emotional slang, which have been introduced before. The second dictionary contains frequently used emotional words, such as happy, angry, and sad. In these two dictionaries, every item is annotated with a letter and an integer. The letter represents the emotional category, and the integer represents the strength of emotion. When annotating the strength of emoticons, emotional words and emotional slang, we adopted the solution in (Paltoglou & Thelwall, 2012). In addition to the two emotional dictionaries, we have also developed two auxiliary dictionaries. The first contains frequently used negative words and phrases, and the second contains frequently used degree adverbs and phrases. These words and phrases are typically used to limit or modify emotional words and can enhance or weaken their strength. Similarly, we also annotated the influence of these degree adverbs and phrases on the strength of emotional words and phrases.

Based on the above dictionaries, the process of calculating the emotional strength of a message is shown in Figure 1. It includes five steps:

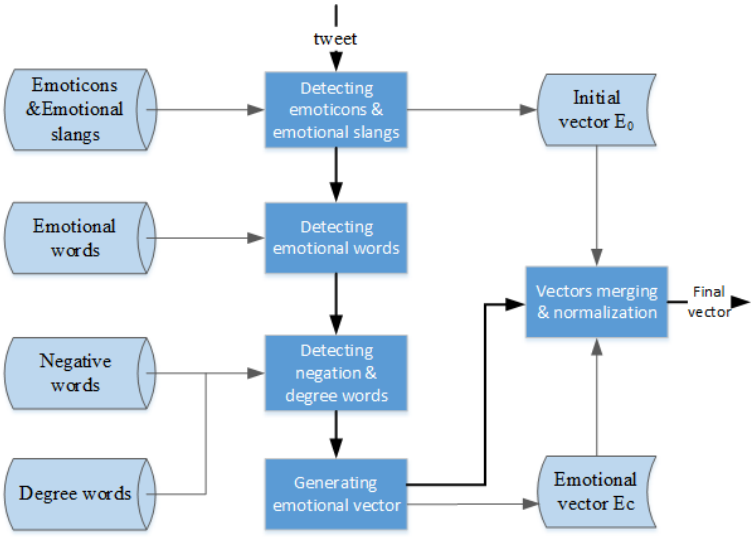


Figure 1. Process of calculating the emotional strength of a message.

(1) Detecting the emoticons and emotional slang in the message and generating an initial vector E_0 based on the annotated emotional strength in the dictionary.

(2) Carrying out word segmentation and detecting the emotional words in the message. Based on the annotated emotional strength in the dictionary, each emotional word will have a corresponding vector E_i ($i \in \{1,2,3\ldots,N\}$).

(3) Detecting the negative and degree words in the neighborhood of each emotional word and modifying the corresponding vector. In this study, the scope of the neighborhood is defined as the three words before the emotional word according to the expression habits of Chinese.

(4) Categorizing the emotional words according to their emotion and generating the aggregated vector E_c. The vector E_c is the sum of vectors that have been regulated.

(5) E_0 and E_c are added and normalized to obtain the final vector representing the emotional strength of the message.

3.3. Infer learners' recent moods

Because it is incredibly unlikely for learners to post a tweet just before learning and express their mood in that tweet, it is necessary to further comprehensively analyze the emotions contained in the multiple tweets released by learners during a period before learning. Generally, the greater the emotional intensity induced by an event, the greater its influence on the current state of mind of learners, and vice versa; the closer an event happens to the present time, the greater its influence on the learner's current state of mind, and vice versa. Based on these facts, this study uses the following two aspects to infer learners' moods before learning: 1) the time when learners post their tweets and 2) the emotional intensity expressed in the tweets.

We assume that a learner's current mood can be inferred from the tweets he or she has posted in the last T hours, and when the T hours are divided into N periods, the tweets published in one period are considered to have the same effect on the learner's current mood. Based on these assumptions, if a learner post M_i tweets in the i-th period and the strength of the j-th tweet is E_{ij}, then the current mood of the learner can be calculated using formula 1.

$$E = \frac{1}{1\times M_1}\sum_{j=1}^{M_1} E_{1j} + \frac{1}{2\times M_2}\sum_{j=1}^{M_2} E_{2j} + \cdots + \frac{1}{N\times M_N}\sum_{j=1}^{M_N} E_{Nj} \quad (1)$$

4. Evaluating the Mood Detection Approach

To validate the accuracy of the proposed mood detection approach, we conducted two experiments on the data set developed by Zhao et al. (Zhao et al., 2012). The aim of the first experiment is to validate the accuracy of identifying emotional tweets, and the aim of the second experiment is to validate the accuracy of calculating the strength of emotional tweets.

4.1. The accuracy of identifying the emotional tweets

The first experiment includes three steps. First, we randomly selected 200 samples from the data set using the first two filtering rules in section 3.1 and manually marked whether these samples were emotional tweets. Second, we used the latter two filtering rules and the two determining rules in section 3.1 to predict whether each sample is an emotional tweet. Finally, we compared the predicted results and the manually marked results.

The manually marked results are shown in Table 1. The results suggest that more than 65 percent of tweets are nonemotional, and most tweets that contain URLs or Chinese square brackets are

nonemotional. This means that it is necessary to identify the emotional tweets when using tweets to detect learners' moods before learning, and the filtering rules we proposed can filter many nonemotional tweets.

Table 1. The manually marked results.

	Number of tweets containing URLs	Number of tweets containing Chinese square brackets	Number of other tweets
Emotional	1	1	68
Nonemotional	12	4	114

To validate the effect of the determining rules, we further compared the predicted results and the manually marked results on the tweets that do not contain URLs or Chinese square brackets. The results are shown in Table 2.

Table 2. The confusion matrix of the tweets that do not satisfy all filtering rules.

Manually marked results	Predicted results		**Recall**
	Emotional	Nonemotional	
Emotional	43	25	**63.24%**
Nonemotional	10	104	**91.23%**
Precision	**81.13%**	**80.61%**	

According to the data in Table 2, we can see that the precision of the predicted emotional tweets is 81.13%, but the recall of the predicted emotional tweets is only 63.24%; the precision of the predicted nonemotional tweets is 80.61%, and the recall of the predicted nonemotional tweets is up to 91.23%. According to these results, we can conclude that the proposed two determining rules can further identify emotional tweets, but the emotional dictionaries should be extended to recall more emotional tweets.

4.2. The accuracy of calculating the emotional strength of a tweet

The second experiment also includes three steps. First, we used the proposed rules to randomly select 200 emotional tweets from the data set and manually marked their emotional strength. Every tweet is marked with a four-dimensional vector, and the value of each dimension is an integer between 0 and 5, which represents the strength of an emotion. Second, every manually marked tweet was processed by a Python program that implements the proposed approach in section 3.2, and the program outputs the predicted vector. Finally, we compared the predicted results to the manually marked results and analyzed their consistency.

When analyzing the consistency between the predicted results and the manually marked results, we first analyzed the consistency of emotion type and then analyzed the consistency of strength. When analyzing the consistency of emotion type, it is considered to be consistent when the value of the specified dimension in the predicted result and the value of the corresponding dimension in the marked result are greater than zero. When analyzing the consistency of strength, we used the correlation coefficient of each dimension. The results are shown in Table 3.

Table 3. The consistency between manually marked results and predicted results

	Happiness	Sadness	Anger	Fear
Number of manually marked tweets	106	39	26	9
Number of predicted tweets	116	30	43	9
Number of consistent tweets	95	21	19	3

Precision	**82%**	**70%**	**44%**	**33%**
Recall	**90%**	**54%**	**73%**	**33%**
Correlation coefficient	**0.35**	**0.17**	**0.21**	**-0.31**

As seen from the data in Table 3, the proposed mood detection approach can recall more than 90% of tweets that express happiness, and more than 80% of the tweets are consistent with the manually marked results. This means that the proposed method can accurately detect the tweets that express happiness. In addition, although the recall of tweets that express sadness is only 54%, the precision is up to 70%. This suggests that the predicted tweets that express sadness are also credible. However, the proposed approach does not accurately detect tweets that express anger and fear emotions. For the correlation coefficient, only the value of happiness is greater than 0.3, and the value of fear is even less than 0. This suggests that the calculation of emotional intensity is very complex, and the predicted result is not credible.

5. Application Scenarios of Detecting Mood before Learning

The main purpose of detecting learners' mood before learning is to enable the learning environment to provide students with more appropriate support services. Therefore, in addition to exploring the detection techniques, this study further combined the actual needs of adaptive learning systems and proposed two application scenarios and corresponding strategies to integrate the mood detection approach into adaptive learning systems.

5.1. Using the mood before learning to improve adaptive navigation

Adaptive navigation is a core function of adaptive learning systems, which can provide adaptive learning topics for learners according to their characteristics. In most existing adaptive learning systems, the system mainly recommends appropriate learning topics for learners based on their knowledge mastery level, and the recommendation is mainly based on Vygotsky's zone of proximal development (ZPD). Specifically, based on the knowledge mastery level recorded in the learner model and the knowledge logic relationship in the domain knowledge model, the topics suitable for the current knowledge state of learners are determined and recommended to learners. When the learner's mood before learning can be detected, is it necessary to adjust the navigation strategy of the adaptive learning system? Research in psychology on emotion and cognitive flexibility has suggested that the answer is yes.

In the field of psychology, many researchers have long conducted in-depth studies on the relationship between emotion and cognitive flexibility. In an earlier study, Isen and colleagues examined the effect of positive emotions on creative problem solving (Isen et al., 1987). The results show that positive emotions help people think out of the box and improve the flexibility of cognitive processing, which in turn helps people solve creative problems better. In addition, Brand et al. studied the effect of negative emotions on cognitive flexibility (Brand et al., 2007). The results show that negative emotions have a certain inhibitory effect on creative problem solving. From the results of the above two studies, it can be seen that learners' mood before learning has an important influence on their cognitive processing.

5.2. Using the mood before learning to improve the resource recommendation

The adaptability of an adaptive learning system is reflected not only in providing adaptive learning navigation for learners but also in providing adaptive learning resources for learners. When different

learners learn the same knowledge, the adaptive learning system will recommend different types of learning resources for learners according to their characteristics. At present, most existing adaptive learning systems mainly recommend learning resources of different media types for learners according to their learning styles. For example, some researchers developed an adaptive web-based learning system that can unobtrusively identify student cognitive styles through his or her browsing behaviors and adaptively recommend learning content to the student based on his or her cognitive style (Lo et al., 2012; Bentaib et al., 2021). When the learners' mood before learning can be detected, does the strategy of learning resource recommendation need to make some adjustments? According to research on multimedia learning, the answer is yes because different types of learning resources have different effects on the emotions of learners with different styles.

Chen and Sun studied the effects of different multimedia learning materials on the emotions and performance of students with verbal learning styles and visual learning styles (Chen & Sun, 2012). The results show that students with verbal learning styles and visual learning styles show significant differences in both positive and negative emotions when using different multimedia materials. Inspired by the above conclusions, we propose that the adaptive learning system should consider not only learners' learning style but also their current emotional state when recommending learning resources to learners. The recommendation principle is that video-based learning resources should be recommended to learners regardless of their cognitive style when the learners' mood is negative. When the learners' mood is positive, recommendations will be made according to their cognitive style.

6. Conclusions and Future Work

Interest in detecting and applying learners' emotional states in online learning has increased in the past decade. However, most studies focus on detecting and applying learners' emotional states in the learning process but do not pay enough attention to learners' mood before learning. Many studies have suggested that mood before learning also has an important effect on the subsequent learning process. Therefore, this study analyses the characteristics of mood before learning and proposes a modeling method and an automatic detection approach. The experimental results show that the proposed detection approach can detect some moods to some extent, especially happy moods. Based on the proposed modeling method and detection approach, this study also explores the application scenarios of learners' moods before learning and proposes an adaptive navigation mechanism and adaptive resource recommendation mechanism that integrate learners' moods before learning.

Although this study proposes a feasible approach to modeling and detecting learners' moods before learning, there are still some issues that need to be explored further. In addition, there are still some learners who rarely use online social media, so there is a certain blind spot in using the tweets released by learners on online social media to detect their moods before learning. In addition, this method also involves learners' personal privacy. The proposed adaptive mechanisms have not been implemented in adaptive learning systems, and their effects on learning performance need to be validated in the future.

Acknowledgements

This work is supported by the Ministry of Education of Humanities and Social Science project (No. 22YJCZH224), Hubei Cooperative Innovation Center for Basic Education Information Technology Service (No.ESRC20220045), and Hubei Provincial Teaching Research Project (No. 2020674)

References

Bentaib, M., Aitdaoud, M., Namir, A., & Talbi, M. (2021). Adaptive help system based on learners "digital traces" and learning styles. *International Journal of Emerging Technologies in Learning (iJET): 16(10)*, 288-294.

Brand, S., Reimer, T., & Opwis, K. (2007). How do we learn in a negative mood? Effects of a negative mood on transfer and learning. *Learning and Instruction: 17(1)*, 1-16.

Calvo, R. A., & D' Mello, S. (2010). Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications. *IEEE Transaction on Affective Computing: 1(1)*, 18-37.

Chen, C. M., & Sun, Y. C. (2012). Assessing the effects of different multimedia materials on emotions and learning performance for visual and verbal style learners. *Computers & Education: 59(4)*, 1273-1285.

Cui, Y., Wang, S., & Zhao, R. (2021). Machine Learning-Based Student Emotion Recognition for Business English Class. *International Journal of Emerging Technologies in Learning (iJET): 16(12)*, 94-107.

Feidakis, M., Caballé, S., Daradoumis, T., Jiménez, D. G., & Conesa, J. (2014). Providing emotion awareness and affective feedback to virtualised collaborative learning scenarios. *International Journal Continuing Engineering Education and Life-Long Learning: 24(2)*, 141-167.

Isen, A. M., Daubman, K. A., & Nowicki, G. P. (1987). Positive affect facilitates creative problem solving. *Journal of Personality and Social Psychology: 52(6)*, 1122-1131.

Karampiperis, P., & Sampson, D. (2005). Adaptive Learning Resources Sequencing in Educational Hypermedia Systems. *Educational Technology & Society: 8(4)*, 128-147.

Lo, J. J., Chan, Y. C., & Yeh, S. W. (2012). Designing an adaptive web-based learning system based on students' cognitive styles identified online. *Computers & Education: 58(1)*, 209-222.

Ortigosa, A., Martín, J. M., & Carro, R. M. (2014). Sentiment analysis in Facebook and its application to e-learning. *Computers in Human Behavior: 31*, 527-541.

Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions*. UK, Cambridge: Cambridge University Press, 1988.

Paltoglou, G., & Thelwall, M. (2012). Twitter, MySpace, Digg: Unsupervised Sentiment Analysis in Social Media. *ACM Transactions on Intelligent Systems and Technology: 3(4)*, 66-84.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology: 39(6)*, 1161-1178.

Zhao, J., Dong, L., Wu, J., & Xu, K.(2012). MoodLens: An emoticon-based sentiment analysis system for Chinese tweets in Weibo. *In Proc. 12th ACM SIGKDD Conference on Knowledge Discovery and Data Mining: 2012*, 1528-1531.

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

Oral or written? The modality of self-explaining influences learning from videos

Caixia Liu [1], Ronghua Yang [1], Yi Zhang [1], Jiumin Yang [1*]
[1] Faculty in Artificial Intelligence in Education, Central China Normal University
[*] yjm@mail.ccnu.edu.cn

Abstract: *This study combined qualitative and quantitative methods to explore the effect of the modality of self-explaining (oral vs. written) on the learning process and outcomes (indicated by attention allocation, immediate and delayed learning performance, cognitive load, metacognition, the quality of explanations, self-perception, and behavioral patterns) when learning from self-paced videos. The results showed that learning by self-explaining in the written form increased students' attention to the text, facilitated their immediate learning performance, metacognition, and content organization, and reduced the cognitive load. Furthermore, the informal results found that the written group reported more positive self-perception than the oral one. Moreover, the analysis of the behavioral patterns indicated that the written group paused the video more often to think while the oral group tended to pause the video briefly and then play it again. The results suggested the positive effect of written explanation on the learning process and learning outcomes when learning self-paced video using a self-explaining strategy.*
Keywords: instructional videos, self-explaining, written explanation, attention allocation, behavioral patterns

1. Introduction

Instructional videos are widely used in formal and informal learning. Moreover, learning strategy is an important factor influencing the effects of learning from videos, and learners are encouraged to actively engage in the learning. Learning by self-explaining is one of the eight generative learning strategies, which referred to students generating statements that clarify the meaning of the learning material by relating it to their existing knowledge (Fiorella & Mayer, 2016; Chi, Bassok, Lewis, Reimann, & Glaser, 1989).

Recent studies have shown that learning by self-explaining is an effective learning strategy, which can improve learning performance and learning experience (Fiorella, Stull, Kuhlmann, & Mayer, 2019; Fiorella & Mayer, 2016; Pi, Liu, Meng, & Yang, 2022). For example, in a study conducted by Fiorella et al. (2019), college students watched an instructional video on how the human kidney works. The video was divided into five segments, and after each segment, participants explained to themselves or re-watched the video. The results showed that the self-explaining group performed better on the academic achievement tests than the control group.

Oral and written explanations are two common modalities of learning by explaining, and previous studies explored the different effects of oral and written explanations on learning (Lachner, Ly, & Nuckles, 2018; Jacob, Lachner, & Scheiter, 2021). Jacob et al. (2021) found that oral explanation is effective when explaining to others because it triggers a unique generative process during interpretation, possibly due to increased social presence. Lachner et al. (2018) claimed that written explanation was more effective than an oral explanation to support the content of student organizations' explanations. In contrast, the oral

explanation can trigger students' elaboration process more obviously than the written explanation. However, Lachner et al. (2021) found that when learning by writing explanations, explaining to a fictitious student was more effective than self-explaining.

Taken together, the modality may influence the effects of learning from videos. However, the previous mainly focused on immediate learning performance and did not examine the effects of the explanation modality on students' long-term memory. Moreover, the influences of explanation modality on attention allocation and learning behavior patterns are also unknown. So we posed the research question: Does the modality of self-explaining (oral vs. written) influence learning process and outcomes (indicated by attention allocation, immediate and delayed learning performance, cognitive load, metacognition, the quality of explanations, self-perception, and behavioral patterns) when learning from self-paced videos?

2. Method

2.1. Participants and design

Fifty-six Chinese college students participated in the experiment, ranging from 17 to 26 years old (M_{age} = 20.73, SD_{age} =2.21; 52 females). They specialized in a wide range of variety, including psychology, educational technology, linguistics, politics, and geography. At the start of the experiment, participants were randomly assigned to two conditions (n = 28 in each group): the oral-explanation group, which was required to explain the content orally, and the written-explanation group, which was required to explain the content by type. All participants were voluntary and received 20 RMB after completing the experiment. This study received approval from the local ethics committee and informed consent signed by all participants.

2.2. Instructional videos

The topic of the instructional video was "Human respiratory system", which consisted of an introductory paragraph and four subtopics: "The composition of the respiratory system", "How the nervous system controls the respiratory system", "The structure of the chest cavity", and "The three stages of the breathing process". The video was a recorded PPT with a female instructor's voice, with text on the left and pictures on the right. No instructor's image was presented during the whole process (Figure 1). The video lasted four minutes. Participants were allowed to control the video, but the maximum learning time was eight minutes.

Figure 1. Screenshot and two areas of the instructional video.

2.3. Measurements

2.3.1. Prior knowledge test

All the questions were derived from the instructional video but did not duplicate those in the learning performance test. It consisted of three fill-in-the-blank questions (1 point per blank, 11 points in total) and three multiple-choice questions (2 points per question, 6 points in total). The total score was 17 points (M = 7.73, SD = 3.39). The test had a moderate internal consistency (Cronbach's $\alpha = 0.66$).

2.3.2. Learning performance test

The learning performance test included an immediate test and a delayed test. The former consisted of two multiple-choice questions (2 points per question, 4 points in total) and four fill-in-the-blank questions (1 point per blank, 15 points in total). The total score was 19 points, with a moderate internal consistency (Cronbach's $\alpha = 0.63$).

The delayed test was the same as the immediate one. Participants were asked to complete it one week later. The total score was also 19 points, with a moderate internal consistency (Cronbach's $\alpha = 0.59$).

2.3.3. Cognitive load scale

The cognitive load scale was based on the measures proposed by Paas et al. (1994). It investigated participants' mental effort and perceived difficulty across two items. All items were rated on a 9-point Likert scale, where 1 represented an extremely small amount, and 9 represented an extremely large amount. The scale had a moderate internal consistency (Cronbach's $\alpha = 0.68$).

2.3.4. Judgment of learning (JOL) scale

In this study, the Judgment of Learning Scale in the metacognition domain was used to test the participants' judgment of learning (Lindner, Eitel, Barenthien, & Köller, 2021). The test consisted of three questions (Cronbach's $\alpha = 0.95$). Participants were asked to rate their confidence in correctly answering questions and explaining the video content from 0 to 100 and their accuracy in the subsequent learning performance test. The average score of the three items was the final score of the JOL scale.

2.3.5. Explanations coding scheme

All oral or written explanations were recorded verbatim by the researchers. Based on the widely used coding schemes in previous studies, this study adapted an explanation quality coding scheme, which included four dimensions: concept, paraphrase, content organization, and elaboration (Lachner et al., 2020; Jacob et al., 2021). A professional researcher further refined each dimension's definition and verified the coding scheme's feasibility (see Table 1). Two independent coders counted the first 20% of all explanations, and the inter-coder reliability was good (*Kappa*s > 0.75). So, one coder coded all remaining explanations.

Table 1. Explanations coding scheme.

Code	Description
Concept	The number of concepts mentioned by the participants in the explanation process.
Paraphrase	The number of text segments correctly explained through retelling or paraphrasing text from video.
Elaboration	The number of knowledge points correctly and in detail explained with prior knowledge.
Organization	The explanation is logical and coherent.

2.3.6. Informal interview

This study used an informal interview to investigate participants' self-perception of oral or

written explanations. The question was: "Do oral or written explanations influence learning differently? If so, what's the difference?"

2.3.7. Log data coding system

Based on the participants' log data, a coding system was developed to encode participants' interactive behaviors while watching the instructional video (see Table 2; Li, 2019; Pi et al., 2022). A professional researcher further verified the feasibility of the coding scheme. Furthermore, two independent coders counted the first 20% of video learning behavior, and the inter-coder reliability was high (*Kappa* > 0.99). So, one coder coded the rest of the learning behavior.

Table 2. The log data coding system.

Code	Behavior	Description
Pl	Play	Clicked the video section to play the video.
Pa	Pause	Clicked the video section to pause the video.
GF	Go forward	Clicked the progress bar to jump to a later time point.
GB	Go backward	Clicked the progress bar to jump to a previous time point.
Re	Replay	Replayed the video from the beginning
TS	Think in silence	Hold still for more than 5 seconds while paused.

2.4. Apparatus and eye movement data analysis

An Eyelink1000 eye tracker was used to record participants' eye movements with a sampling rate of 1000 Hz. The instructional video was displayed on 1024 × 768 pixel screen. The participants sat 70 centimeters away from the monitor.

To measure the attention allocation, we created two areas of interest (AOIs): the text area and the picture area (see Figure 1). The eye movement indexes selected in this study included the percentage of fixation duration on the text area, the percentage of fixation duration on the picture area, and the transitions between text and picture.

2.5. Procedure

The experiment was conducted in an eye movement lab and lasted about 40 minutes. Firstly, the researchers introduced the basic information and procedure of the experiment to the participants and led them to sign the informed consent (5 min). Afterward, the participants filled out the demographic questionnaire (e.g., gender, age, and major) and the prior knowledge test (5 minutes). Then, all participants were randomly assigned to one of two conditions. In both conditions, they first watched an instructional video. As they watched, their eye movements and log data were recorded (10 minutes). After watching the video, they were asked to explain what they had learned by speaking or typing (5 minutes). Then they completed the JOL test, the immediate learning performance test, the cognitive load test, and the informal interview (10 minutes). A week later, they completed the delayed learning performance test online.

3. Results

Before examining the differences between the two groups, the study first tested whether there were differences in performance on the prior knowledge test. The results showed no significant difference in prior knowledge between the two groups [$t(54)=1.39$, $p = .169$, Cohen's $d = 0.37$].

Since the data were normally distributed, a series of independent samples *t*-tests were conducted with the explanation modality as the variable. The dependent variables were attention allocation, learning

performance, cognitive load, judgment of learning, and explanation quality. Descriptive statistics and the results of *t*-tests are shown in Table 3.

Table 3. Declarative statistical results for all dependent.

Dependent variables	Oral (*n* = 28)		Written (*n* = 28)	
	M	***SD***	***M***	***SD***
Prior Knowledge (max. = 17)	7.11	3.25	8.36	3.47
Attention Allocation				
Percentage of fixation duration on text	57.26	11.11	62.46	8.85
Percentage of fixation duration on picture	31.58	12.55	28.92	9.37
Transitions between text and picture	67.64	20.75	64.04	16.50
Learning Performance				
Immediate learning performance (max. = 19)	14.82	2.91	16.29	2.79
Delayed learning performance (max. = 19)	14.93	2.91	16.00	2.72
Cognitive Load (max. = 9)	6.63	1.42	5.54	1.30
Perceived difficulty	5.82	1.74	4.18	1.79
Mental effort	7.43	1.37	6.89	1.32
JOL (max. = 100)	63.46	3.65	76.10	2.17
The Quality of Explanations				
Concept (max. = 20)	13.29	3.05	13.32	3.49
Paraphrase (max. = 18)	6.14	2.72	5.75	3.19
Elaboration (max. = 10)	2.57	1.69	2.36	1.19
Content organization (max. = 10)	2.89	1.69	3.89	1.87

3.1. Attention allocation

In order to determine whether the explanation modality affected the participants' attention allocation, we conducted independent sample *t*-tests on the percentage of fixation duration on text and picture area, as well as the transition between text and picture. It was found that the written-explanation group was associated with slightly significantly more focus on the text area [$t(54)=1.94$, $p = .058$, Cohen's $d = 0.52$]. However, the percentage of fixation duration on the picture area [$t(54)=0.90$, $p = .373$, Cohen's $d = 0.24$] and the transitions between text and picture [$t(54)=0.72$, $p = .475$, Cohen's $d = 0.19$] did not reach a significant level.

3.2. Learning performance

There existed a marginally significant difference in the immediate performance [$t(54)=1.92$, $p = .060$, Cohen's $d = 0.52$]. Both groups scored highly on the immediate test, but the written explanation led to better learning. However, the benefits of this explanation modality did not reflect in the delayed test [$t(54)=1.42$, $p = .160$, Cohen's $d = 0.38$].

3.3. Cognitive load

Similar analyses were repeated with cognitive load and its two dimensions. The analysis indicated that the cognitive load in the oral explanation group was significantly higher than the written-explanation group [$t(54)=3.00$, $p = .004$, Cohen's $d = 0.37$]. Specifically, the oral explanation group perceived more difficulty than the written explanation group [$t(54)=3.48$, $p = .001$,

Cohen's d = 0.93], but they put in similar mental effort [t(54)=1.49, p = .142, Cohen's d = 0.40].

3.4. Metacognition

The judgment of learning is an important form of metacognitive judgment. So, we used the score of the JOL test to measure participants' metacognition. The results showed that there were significant differences in metacognition between the two groups [t(43.96)=2.97, p = .005, Cohen's d = 0.89]. The written-explanation group scored significantly higher than the oral-explanation group.

3.5. The quality of explanations

To test the difference in the quality of oral explanation and written explanation, we conducted independent sample t-tests on four dimensions, including concept, paraphrase, elaboration, and content organization. However, the results showed no significant differences in the three dimensions except content organization (ps > .05). In the dimension of content organization, the score of the written explanation group was significantly higher than that of the oral explanation group [t(54)=2.10, p = .040, Cohen's d = 0.56].

3.6. Self-perception

To gain further insight into participants' views on oral and written explanations, we analyzed the informal interview data of 56 learners. Most learners believed that written explanation and oral explanation had different effects on their learning (76.80%).

As shown in Figure 2, further analysis showed that oral explanation was characterized by its convenience and comprehensiveness. Learners believed that the process of oral explanation could help them understand and grasp knowledge, deepen their memory, and create a sense of face-to-face communication. On the other hand, the written explanation was more of a deliberate product. It required thinking through the content and organizing it into coherent language, then writing them down. Unlike oral explanation, the written explanation could be revised at any time, was more logical, and could deepen the learner's memory. It was also reflected in their higher content organization and learning performance scores.

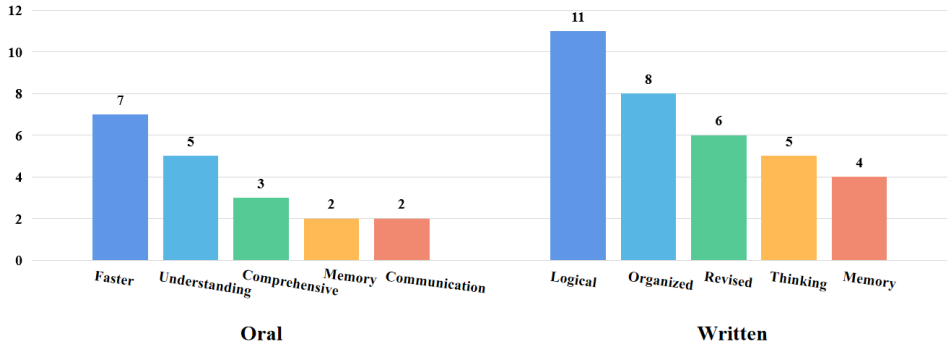


Figure 2. Participants' self-perception of oral and written explanations.

3.7. Behavioral patterns

The radar chart in Figure 3 showed the code frequencies between the two groups, which indicated that the frequency of code was similar between the two groups. The results showed that the frequency of code was similar between the two groups. They had the highest frequencies of GF (Go forward), indicating that learners often fast-forward videos. Second, learners also frequently used Pl (Play), Pa (Pause), and GB (Go backward). So learners may constantly monitor their learning status and pause to review what they have learned. However, the use of Re (Replay) was the lowest. It may be because learners have checked the previous content by pausing and retreating frequently. So they no longer have to replay the whole video. In addition, the frequency of TS (Think in silence) was slightly different

between the two groups, with the written-explanation group pausing the video more often to think. The possible reason is that they need to organize and integrate their learning content timely. This process takes up too many cognitive resources, which is also reflected in their high cognitive load.

Furthermore, we conducted a sequence analysis using GSEQ 5.1 to test whether these sequential relationships reached significant levels. All sequences that achieved statistical significance (*Z*-score > 1.96) were plotted as behavioral transition diagrams. The value represents the *Z*-score of the sequence and the arrow points in the direction of the transition. As shown in Figure 3, the behavior patterns of the two groups in the learning process were different. It mainly reflected in that the oral-explanation group had two more significant behavior sequences than the other one: Re→Pl (Replay→Play) and Pa→Pl (Pause→Play). These behavior paths mean that learners in the oral-explanation group tended to watch the instructional video again, and were more inclined to pause the video briefly and then play it again

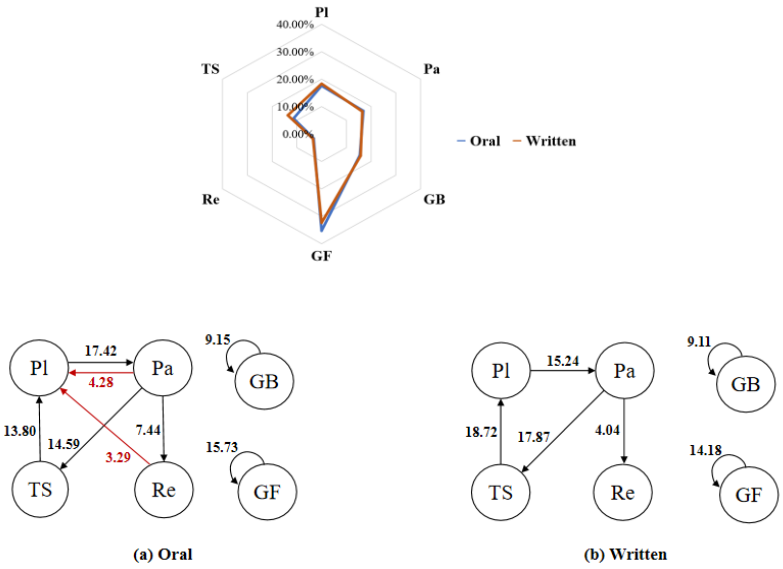


Figure 3. A radar chart of codes distribution and two behavioral transition diagrams (the numbers represent the Z-scores, and the arrows point in the direction of the transfer) for the (a) oral-explanation, and (b) written-explanation groups. Pl = Play; Pa = Pause; GF = Go forward; GB = Go backward; Re = Replay; TS = Think in silence.

3.8. Exploratory analysis

Next, we used moderating effect analyses to explore whether prior knowledge would be a moderator factor affecting the relationship between explanation modality and all variables. Only the moderated effect of prior knowledge between modality and attention allocation was found. Specifically, the interaction of explanation modality with prior knowledge could significantly positively predict the percentage of fixation duration on text area [$F(1, 52) = 5.02$, $p = .029$]. And it could also significantly negatively predict the transitions between text and picture [$F(1, 52) = 5.12$, $p = .028$]. However, for the percentage of fixation duration on picture, the negative predictive effect of the interaction between explanation modality and prior knowledge only reached a marginal significance level [$F(1, 52) = 3.75$, $p = .058$]. Further simple slope analysis results were shown in Figure 4. For high-prior-knowledge learners, oral explanation prompted them to pay more attention to the picture ($p = .046$) and less to the text ($p = .005$), and switch between the text and text more often ($p = .047$). But for learners with low prior knowledge, explanation modality had no significant effect on their attention allocation ($ps > .005$).

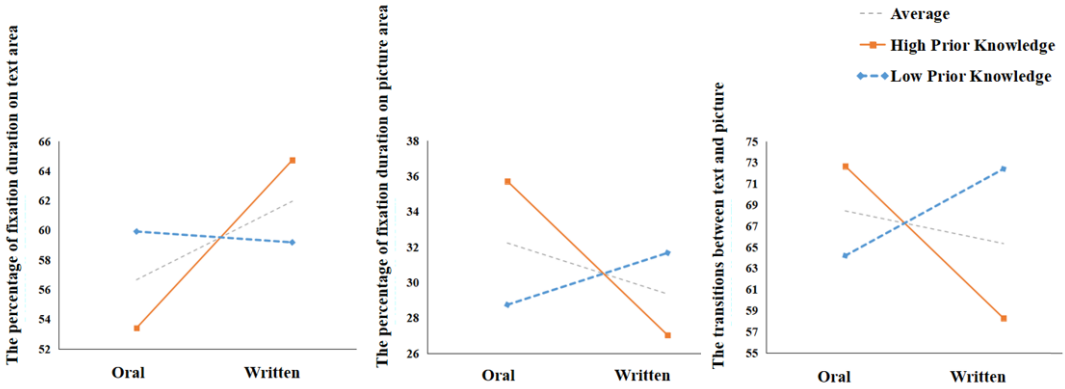


Figure 4. Moderating effect of prior knowledge between explanation modality and attention allocation.

4. Discussion

This study explored the effects of the modality of self-explaining on the learning process and outcomes (indicated by attention allocation, immediate and delayed learning performance, cognitive load, metacognition, the quality of explanations, self-perception, and behavioral patterns) when learning from self-paced videos. Specifically, written explanations increased students' attention to the text, facilitated their immediate learning performance, metacognition, and content organization, reduced the cognitive load, and triggered more positive self-perception. Moreover, the written group paused the video more often to think deeply.

Regarding attention allocation, we found that the written-explanation group allocated significantly more attention to the text area. The possible interpretation is that written explanation is presented in the form of text, so learners need to combine the text content with prior knowledge. However, the oral explanation group may be more dependent on the oral explanation in the video, which was also supported by the fact that the oral explanation group exhibited more Replay→Play and Pause→Play sequences. Regrettably, we failed to find the difference in transitions. The number of transitions is regarded as associated with the integration of the text and pictures, so the results indicated that the modality did not influence students' integration of knowledge (Coskun & Cagiltay, 2021).

Furthermore, we explored the moderated effect of prior knowledge between modality and attention allocation. We found that for high-prior-knowledge learners, oral explanation prompted them to pay more attention to the picture and less to the text, and switch between the text and text more often; while for low-prior-knowledge learners, the modality had no significant effect on their attention allocation. One possible explanation was that learners who have more prior knowledge devoted cognition to the organization and integration of knowledge rather than the selection of knowledge.

In addition, we found that the written explanation can prime students' metacognition during learning and thus improve learning performance, and help students organize the information in the explanation better than in the oral explanation (McNamara & Magliano, 2009). Higher-level organization in turn helps students acquire deeper conceptual knowledge, which is also confirmed by students' reported self-perception (Lachner et al., 2018). However, written explanations can increase students' cognitive load, especially in the dimension of perceived difficulty. Regrettably, we did not find a difference in the modality on delayed learning performance, which may be because the delayed test was so easy that the students got higher scores.

In conclusion, the results found that written explanation has a significant advantage in learning by self-explaining. Our findings have implications for the implementation of self-paced instructional videos: students are encouraged to use a self-explaining strategy in the written format to improve the learning process and outcomes.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant [62177027], and the Fundamental Research Funding for the Central Universities under Grant [2022CXZZ031].

References

Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145-182.

Coskun, A., & Cagiltay, K. (2021). A systematic review of eye-tracking-based research on animated multimedia learning. *Journal of Computer Assisted Learning, 38*(2), 581-598.

Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review, 28*, 717-741.

Fiorella, L., Stull, A. T., Kuhlmann, S., & Mayer, R. E. (2020). Fostering generative learning from video lessons: Benefits of instructor-generated drawings and learner-generated explanations. *Journal of Educational Psychology, 112*(5), 895-906.

Jacob, L., Lachner, A., & Scheiter, K. (2021). Does increasing social presence enhance the effectiveness of writing explanations?. *Plos one, 16*(4), e0250406.

Lachner, A., Backfisch, I., Hoogerheide, V., van Gog, T., & Renkl, A. (2020). Timing matters! Explaining between study phases enhances students' learning. *Journal of Educational Psychology, 112*(4), 841-853. c

Lachner, A., Ly, K., T., & Nuckles, M. (2018). Providing written or oral explanations? Differential effects of the modality of explaining on students' conceptual learning and transfer. *The Journal of Experimental Education, 86*(3), 344-361.

Lachner, A., Jacob, L., & Hoogerheide, V. (2021). Learning by writing explanations: Is explaining to a fictitious student more effective than self-explaining? *Center for Open Science*. http://dx.doi.org/10.31234/osf.io/8hjcx

Li, L. Y. (2019). Effect of prior knowledge on attitudes, behavior, and learning performance in video lecture viewing. *International Journal of Human Computer Interaction, 35*(4–5), 415-426.

Lindner, M. A., Eitel, A., Barenthien, J., & Köller, O. (2021). An integrative study on learning and testing with multimedia: Effects on students' performance and metacognition. *Learning and Instruction, 71*, 101100.

McNamara, D. S., & Magliano, J. P. (2009). Self-explanation and metacognition: The dynamics of reading. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Ed.). *Handbook of Metacognition in Education.* New York and London: Routledge.

Paas, F. G., Van Merriënboer, J. J., & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills, 79*(1), 419-430.

Pi, Z., Liu, C., Meng, Q., & Yang, J. (2022). Co-learner presence and praise alters the effects of learner-generated explanation on learning from video lectures. *International Journal of Educational Technology in Higher Education, 19*(1).

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

A Systematic Review of Stimulated Recall (SR) in Education from 2012 to 2021

Xiaoyan Chu [1*], Xuesong Zhai[1]
[1] Zhejiang University
[*] xiaoyan_chu@zju.edu.cn

Abstract: *Stimulated Recall (SR) has been widely employed in educational studies for a long time as an efficient method of introspection. However, with the advent of online and blended learning, as well as the application of ever more cutting-edge technologies to educational settings over the last decade, it is unclear how SR has progressed and how researchers have implemented it effectively. The content analysis revealed: (1) most articles used SR to explore participants' thought patterns, investigate the effect of educational strategies, and facilitate metacognition; (2) SR video stimuli get further advancement, and the stimulus sources became more diverse, with physiological data incorporated; (3) In SR interviews, more strategies are applied, including flexible intervals and questioning styles. This article provides in-depth insights on SR research in the education domain, presenting a prospective avenue for educational researchers to use SR to investigate educational issues in the post-pandemic era.*
Keywords: stimulated recall, metacognition, cognitive process, eye tracking

1. Introduction

Stimulated Recall (SR), based on vivid prompts and the concept of reflective thinking proposed by Dewey (1933), can revive participants' memories by presenting clues or stimuli occurred in original scenario (Bloom, 1953). Extensive empirical studies have elucidated that stimulated recall is a practical approach to exploring teachers' and students' implicit viewpoints and promoting their retrospection and metacognition competence (Zhai et al., 2022). In addition to probing and boosting participants' thinking, stimulated recall also assists researchers in validating and exploring the effects of some teaching and learning strategies such as technology-mediated learning, various feedback strategies etc.(Van der Kleij et al., 2017; Cao et al., 2019). Besides diverse educational functions and research purposes, employing SR in education and educational research has a long history dating back to last century. Bloom (1953) initially identified the concept and definition of Stimulated Recall, and the method was utilized in educational research in 1970 by Stanford University researchers at the outset. Since then, it has been a vital instrument in pedagogical research and has been used to investigate a wide range of issues in educational research. The last decade witnessed the advancement of ICT-supported teaching and learning, as well as the alteration of stimulated recall. However, little research has summarized studies involving the stimulated recall method..

The stimulated recall technique is an introspective method for eliciting cognitive processes and other thinking strategies when participants conducting an activity or task (Lyle 2003). Behind exterior behaviors, participants' internal thought patterns are generally tacit and difficult to be directly observed or measured by researchers from an external perspective. But stimulated recall can help us to acquire information inside the participants' minds. With external behaviors (e.g., teaching instructions, verbal

communication, and motor movements) recorded as effective stimuli for retrospection, participants are inclined to verbalize their inner thoughts that drove these behaviors (Gass & Mackey, 2000; Mackey and Gass, 2015). Usually, SR consists of two steps: providing stimuli and posing recall questions (Fig. 1). To boost recall, researchers will first select specific artifacts which display participants' behavior or cognitive tasks to assume the role of stimuli, such as notes, audio recordings, video recordings, etc. Then researchers conduct strategic interviews to encourage participants to recount and verbalize their inner thoughts, mental activities, or personal sentiments at moment when stimuli are created (Calderhead, 1981; Lyle, 2003). With these two critical procedures, SR enjoys an edge in helping teachers and learners recall more details without interfering with the normal teaching and learning process (Omodei & McLennan, 1994; Gaydos and Devane, 2019).

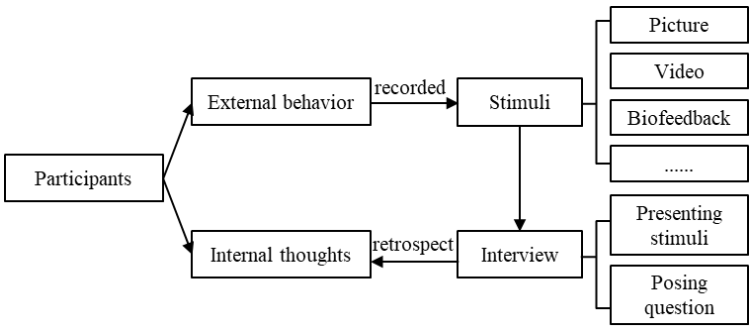


Figure 1. The flow diagram of the SR procedure

On top of the expeditious diffusion of Internet-connected technologies and the exponential growth of online learning, stimuli are constantly evolving to contain richer information about participants' activities. On the one hand, the integration of ICT exerts an influence on stimuli selection by giving rise to the transformation of the learning context. In physical settings, stimuli in SR, such as audio or video recordings, can favorably respond to interactive activities between instructors and learners. However, classes incorporating digital technology are apt to be more static, with instructional activities carried out via electronic devices and unable to be captured easily. On the other hand, technological advancements also enhance stimuli in SR by enriching its data collection channels and capacities. Versatile stimuli sources such as weblogs, computer screen recordings, bio-feedback data and so on grant researchers assistance to unearth information about the learner's inner workings.

After deciding on and obtaining elucidative stimuli, performing tactical and in-depth interviews is another crucial stage of stimulated recall. Because this segment emphasizes estimating internal thinking processes and determines the extent to which the stimulated recall method can energize teachers' and students' retrospection and probe into their inner thought. In previous studies, researchers were inclined to conduct after-class interviews as soon as possible and adopt common open-ended questions to ensure participants' autonomy in their reflection(Lyle, 2003: 863; Gass and Mackey, 2000). Nevertheless, it is noted that some studies chose the disparate way. Kuri and his colleagues (2016) did not invite teachers to participate in the interview until two weeks later. For presenting questions in the interview, some researchers also reckon that together with generic ones, follow-up specific questions concentrated closely on the research aim are equally essential (Hu and Wu, 2020).

Synthesize the above assumption, with teachers' and students' retrospection activated by stimuli and interview, SR has assumed a prominent role in investigating their cognition and behaviour patterns in the field of education. There is a lack of a holistic review of the development of stimulated recall in the last decade, impeding its maximum potential. Therefore, this research aims to investigate how SR has

developed over the decades through an extensive systematic review. The research questions that guided this review are as follows:

(1) What research aims are proposed when SR were implemented in the past decade?

(2) What have the stimuli been developing when SR were implemented in the past decade?

(3) What interview strategies have been employed according to the selected SR research.?

2. Method

2.1. Paper selection

To ensure that the papers selected are of a high quality, the literature sources for this review were selected from the Web of Science, Scopus, and IEEE Xplore. Considering the teaching environment and methods having experienced significant changes in the recent decade, we set the time span from 2012 to 2021.

The procedures implemented to identify the research papers of this study can be classified into two stages. In the first stage, the keyword selected was "stimulated recall", and the disciplinary was refined to "education and educational research". Two hundred twenty-seven articles were yielded in this process. In the second stage, researchers manually and systematically screened the abstracts and the full text of the selected articles according to the inclusion criteria.

Each full text of all the identified papers was read and screened individually by three-panel members with doctoral degrees or professorships in the field of instructional technology. Finally, 243 representative papers were identified as the research sample pool of this review. We adopted a robust and comprehensive approach to conduct the systematic review, recommended by Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA, Page et al., 2021). Following Moher et al. (2009), Fig.2 depicts the screening process flow.

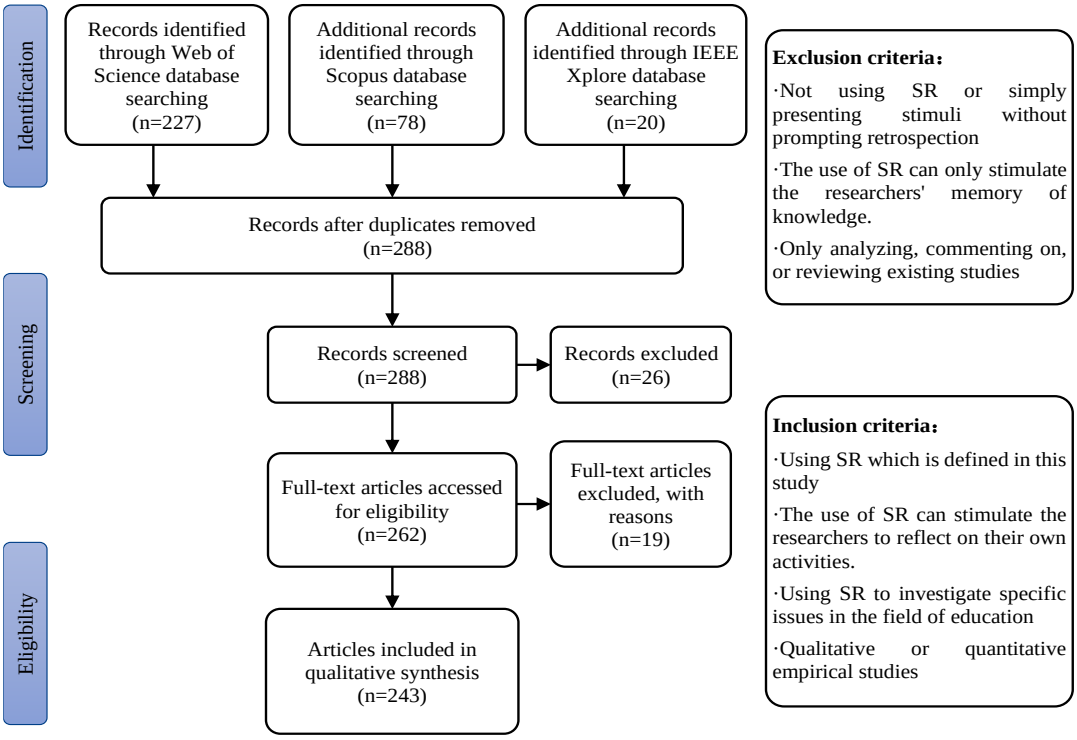


Figure 2. PRISMA flow diagram of the research

2.2. Paper selection

To enable a clear and in-depth analysis of the application of SR in education, we systematically coded these articles. Based on Gass and Mackey's (2000) definition of stimulated recall, we identified the key components that emerged in the study and formed a coding framework, including the purpose of the study, stimuli, question strategy, and question interval (the sub-items of which are listed in the next section). For the reason that stimuli recall is a commonly used teaching and research tool, coding the discipline and learning environment helps us clarify how to apply SR according to various situations. Although we reviewed studies with teachers or learners as subjects, we did not distinguish between the two categories in that SR mainly focuses on the inquiry into the consciousness and thinking behind participants' behaviour.

Table 1. Research foci

Research foci	Description
Research aim	Purposes that the researcher expected to explore with SR
Stimuli	Materials used to stimulate participants to reflect on prior activities
Question strategy	Questioning techniques used to stimulate participants' reflection during SR
Question interval	The time interval between the actual activities and the SR interview

3. Results and discussions

According to the above coding criteria and content analysis, a total of 243 papers are carefully reviewed. The following sections describe our findings and respond to the research questions.

Table 2. Characteristics of included studies (n=243).

Description	*n*	%
Research aim		
Patterns of non-cognitive processing	82	32.41
Patterns of cognitive processing	56	22.13
Effect of learning strategies	47	18.58
Effect of instructional strategies	34	13.44
Other	17	6.72
Improvement of metacognition	9	3.56
Patterns of higher-order thinking	8	3.16
Stimuli		
Video recordings	169	69.55
Text	25	10.29
Pictures	5	2.06
Not mentioned	5	2.06
Audio recordings	12	4.94
Eye movement data	11	4.53
Text and audio/video recordings	10	4.12
Text and pictures	4	1.65
Pictures and audio/video recordings	2	0.82
Questioning strategy		
Not mentioned	94	38.68
Specific on research aim	66	27.16
General	58	23.87
General first and then specific	24	9.88
Specific first and then general	1	0.41
Questioning Interval		

Instant reflection	115	47.33
Not mentioned	108	44.44
Delay reflection	20	8.23

Note: Since some articles hold more than one research aim, the sum of articles with distinct research purposes is more than the total number of review articles in the current research.

3.1. Research aim

3.1.1. Exploring thought patterns

As illustrated in Table 2, more than half of the studies included adopted stimulated recall to probe participants' implicit thought patterns, including non-cognitive process, cognitive process, and higher-order thinking. For patterns of noncognitive processing, eighty-two of reviewed studies have explored this issue, involving motivational, emotional, and cultural orientation etc. (Lichtinger and Kaplan, 2015). This approach can also explore factors that influence willingness to communicate or the ethics of teaching behaviors during language learning (Peng, 2020). For patterns of cognitive processing, fifty-six studies probed the epistemic thinking of different participants (both instructors and learners) under various subjects (Language learning, STEM, Art etc.) (Cherrington and loveridge, 2014). In addition, based on the stimulated recall, some studies have also simultaneously gained insight into cognitive and noncognitive processes by integrating multivariate data (Lambert and Zhang, 2019). For patterns of higher-order thinking, SR has also played an prominent role in investigating participants' critical and creative thinking, along with the collaborative process (Schindler and Lilienthal, 2020).

3.1.2. Investigating the effect of educational strategies

Our review suggests that exploring the impact of instructional design or learning models on participants' learning processes and learning experiences is another purpose for researchers to adopt SR methods. 47 articles used stimulus recall to explore the effects of using specific learning strategies in instructional scenarios, while 34 studies examined instructional strategies. After adopting a new teaching or learning strategy, SR can help researchers learn more about how teachers and students understand and apply these strategies. For example, SR can help to understand the pedagogical knowledge base of novice teachers, teachers' use of dashboards, and teachers' provision of feedback to students (Molenaar and Knoop-van Campen, 2019). For the effectiveness of learning strategies such as computer-supported learning for self-directed learning, SR provides more accurate and in-depth information on student perspectives. Extensive empirical studies indicates that SR-acquired data can improve the interpretability of single-outcome data such as test scores and produce more valuable information, helping teachers and students to better evaluate and improve strategies for better learning outcomes (Van der Kleij et al., 2017).

3.1.3. Improving metacognition

According to the results of the review, nine articles mentioned the purpose of adopting SR to enhance participants' metacognition. Metacognition is an awareness of one's thought processes and comprehending the patterns that underpin them (Flavell 1979). It consists of metacognitive knowledge and metacognitive activities. Educational psychologists have paid much attention to metacognition in the learning process and assume it significantly impacts students' academic performance in laboratory and classroom settings. A study on self-regulated learning noted that metacognitive activities during the self-reflection phase involve evaluating their task performance and understanding or memory. SR participants are asked to explain or comment on their previous behavior rather than merely remembering the knowledge. Therefore, their metacognitive skills are improved. it's worth noting that involving students themselves in video-stimulated recall of feedback conversations contributed to students' self-reflection

of their involvement in the feedback process, encouraging them to make their voices heard and participate in feedback as a dialogic practice (Van der Kleij et al., 2017).

3.2. Stimuli

3.2.1. Optimizing video stimuli

Video recordings are the most frequently-used stimuli selected by 70% of reviewed research (169 articles). Video-based stimuli are generally composed of real-life scenes based on classroom or laboratory learning or screen recordings of computer-supported learning scenarios. Video recordings as incentives require high quality, and researchers should control the length of the video to prevent participant fatigue (Lee, 2020). Researchers have also proposed several improvement options: The first is to select clips from the full-length video footage as the final stimuli. This approach effectively reduces the duration of SR and allows participants to focus on their behavior relevant to the research question. Secondly, researchers record stimuli videos with multiple cameras, which provides various perspectives of the classroom learning situation when played in split screens as visual stimuli. For example, Jackson and Cho (2018) recorded both teacher and student behaviors and performance in the classroom simultaneously. A split-screen video was created and played on a computer screen. In addition, some researchers also manipulate head-mounted video cameras for video recording. This video-recording perspective visually reproduces what participants see and hear in learning scenarios. The head-mounted video provides much assistance in studies concerned with interpersonal communication, such as research exploring interactions between instructors and students (Agricola et al., 2021) or teacher intervention in early childhood peer conflict (Myrtil et al., 2021).

3.2.2. Combining multiple sources

16 studies incorporating more than one material as finally stimuli. On the one hand, the inclusion of stimuli from various sources such as video, text, and pictures broaden the time-span and breadth of the triggers reflect, on top of ensuring that the video records authentic details. Videotapes only capture a fixed amount of time, but teaching preparation and learning behaviors do not merely occur in the classroom. Teachers' preparation notes, guide sheets, and students' class notes can also serve as cues along with classroom learning videos to stimulate participants' recall better. On the other hand, multiple sources of stimuli provide participants with a more comprehensive perspective that includes both subjective and objective aspects. Because video is an external viewpoint, it only records the participants' behavior at class time. Still, the teachers or the students themselves will also keep some notes or logs, compensating for the video's inability to record subjective data and bringing participants a more intimate and familiar memory cue. In a study exploring metacognitive interventions in implementing twenty-first-century writing pedagogies, stimuli consisted of a literacy autobiography, a unit plan, and teaching materials in addition to a classroom tour video to gather the participant's perceptions (Jensen, 2019).

3.2.3. Utilizing biofeedback data

As physiological feedback techniques continue to develop in educational research, physiological feedback data are gradually becoming one of the stimuli choices for SR. Unlike the traditional photo, audio, and video stimuli, physiological feedback provides more accurate and detailed data. The physiological feedback technique integrated with SR now is eye-tracking technology which is utilized in 11 research articles among reviewed studies. According to the Eye-Mind hypothesis, researchers assume that eye movements correspond to our mental operations (Obersteiner & Tumpek, 2016, p. 257). Cognitive processes can consequently be inferred from gaze patterns. Combining eye-tracking data with the SR method reduces the possible ambiguity and uncertainty of eye-tracking techniques (Schindler and

Lilienthal, 2019) and provides a more comprehensive picture of the learning process for retrospection (Michel et al., 2020). In addition to eye-movement data, other physiological signals have also functioned as informative stimuli in the SR method. Zhai (2018) and his colleagues used both eye-movement and EEG physiological signals as stimuli to significantly improve online self-directed learners' reading comprehension and cognitive abilities.

3.3. Interview strategies

3.3.1. Setting up flexible intervals

In general, the time interval between completing classroom instruction and conducting SR interviews is relatively nimble, and different researchers have different options. In majority (117 articles) of the reviewed research where the time interval was explicitly stated, instant reflection was preferred, which means the stimulated recall interviews were conducted immediately after participants complete their learning tasks (Shintani, 2016). This is because a shorter time interval ensures that participants still have a clearer memory of their cognitive processes at the time they completed the task, thus enhancing the accuracy of the interview data (Gass & Mackey, 2000). Therefore, interviews are usually conducted instantly or with a slight delay depending on the school's curriculum schedule.

On the other hand, instead of using a practice of reducing the number of interviews conducted without delay, some researchers (20 articles) extended the time interval and asked participants to recall only after a period of time, for example, 2-4 weeks afterwards. This design may have been motivated by optimizing the study itself and reducing the impact on participants. For researchers, they need more time to analyse and process previously recorded raw data and footages, such as selecting video clips that are closely related to the purpose of the study (Kurki et al., 2016). For participants, delayed stimulated recall interviews can avoid influencing their subsequent teaching and learning activities so some interviews will not start until the academic semester ends.

3.3.2. Posing appropriate questions

In SR interview, researchers stimulate participants' border reflection in an open dialogic space by means of presenting questions. The common pattern employed is a series of general, reflective and open-ended questions that require no straightforward answers, which is observed in 82 articles, including studies that ask only general questions and those that ask general and then focused questions. During the interview, researchers serve more as a listener than an inquirer, playing the role of training, facilitating, and illuminating, and avoid putting forward with leading questions (Ramnarain and Modiba, 2013; Egi, 2008; Sato, 2019). For example, the researcher would avoid the yes-no questions that would lead the participants to respond in a particular direction. Also, such questions make the purpose of stimulated recall interviews less evident to the participants to minimize the risk of presentational responses (Rassaei, 2020; Thararuedee and Wette, 2020). Although SR questioning should leave enough room for participants to reflect, it still facilitates addressing the research aims, so 24 articles claim that the initial questions can be open, but they become more specific as the interview proceeds (Lichtinger and Kaplan, 2015). Researchers utilize follow-up questions depending on participants' responses as a prompt to ensure that the entire interview does not deviate from the research topic and that it goes deeper.

In addition to the questioning scope, guidance and neutrality of the questions also count. Even though participants have a great deal of autonomy and freedom in answering questions, they should receive some training before the interview begins, including clear and standardized instructions regarding how to think back to previous cognitive processes (Rassaei, 2013; Cherrington and Loveridge, 2014). When posing questions, the interviewer should stay as neutral as possible to capture the effect of mere

confrontation with the perspective provided by the stimuli and a facilitator's support in recalling (Consuegra et al., 2016). Because if participants perceive a bias in the question or some other value judgment, they would rationalize and make up explanations under pressure instead of accurately reporting their interactive thinking. Therefore, the interviewer needs to carefully design the wording of the questions and adopt a positive and supporting attitude that merely indicates interest in participants' descriptions rather than judging them (Schindler and Lilienthal, 2019).

4. Conclusions

This study provides a systematic review of 243 empirical articles on the use of SR in education from 2012 to 2021. This paper analyzes some changes of the SR method in the current environment when the learning environment is moving to online and virtual space, and more and more technological tools are involved in the educational process. First of all, this study found that researchers often utilized SR to explore participants' inner views and thoughts, and their metacognitive abilities got improved in the process. Secondly, based on the selection and processing of stimuli, video stimuli with the highest frequency of use are constantly being upgraded. The sources of stimuli are becoming more diversified, physiological feedback data included. Finally, participants need some space to respond to questions in the interview. Future studies can build on this foundation for a more scientific and effective research design, and explore the combination of more cutting-edge technological tools such as physiological feedback and SR.

References

Agricola, B. T., Prins, F. J., van der Schaaf, M. F., & van Tartwijk, J. (2021). Supervisor and Student Perspectives on Undergraduate Thesis Supervision in Higher Education. *Scandinavian Journal of Educational Research*, 65(5), 877-897.

Bloom, B. S. (1953). Thought-processes in lectures and discussions. *The Journal of General Education*, *7*(3), 160-169.

Calderhead, J. (1981). Stimulated recall: A method for research on teaching. *British journal of educational psychology*, *51*(2), 211-217.

Cao, Z., Yu, S., & Huang, J. (2019). A qualitative inquiry into undergraduates' learning from giving and receiving peer feedback in L2 writing: Insights from a case study. *Studies in Educational Evaluation*, 63, 102-112.

Cherrington, S., & Loveridge, J. (2014). Using video to promote early childhood teachers' thinking and reflection. *Teaching and Teacher Education*, *41*, 42-51.

Consuegra, E., Engels, N., & Willegems, V. (2016). Using video-stimulated recall to investigate teacher awareness of explicit and implicit gendered thoughts on classroom interactions. *Teachers and Teaching*, *22*(6), 683-699.

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. DC Heath.

Egi, T. (2008). Investigating stimulated recall as a cognitive measure: Reactivity and verbal reports in SLA research methodology. *Language Awareness*, *17*(3), 212-228.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist*, *34*(10), 906.

Gaydos, M. J., & Devane, B. M. (2019). Designing for identity in game-based learning. *Mind, Culture, and Activity*, *26*(1), 61-74.

Hu, J., & Wu, P. (2020). Understanding English language learning in tertiary English-medium instruction contexts in China. *System*, 93, 102305.

Jackson, D. O., & Cho, M. (2018). Language teacher noticing: A socio-cognitive window on classroom realities. *Language Teaching Research*, *22*(1), 29-46.

Jensen, A. (2019). Fostering preservice teacher agency in 21st century writing instruction. *English Teaching: Practice & Critique*.

Kurki, K., Järvenoja, H., Järvelä, S., & Mykkänen, A. (2016). How teachers co-regulate children's

emotions and behaviour in socio-emotionally challenging situations in day-care settings. *International Journal of Educational Research*, 76, 76-88.

Lambert, C., & Zhang, G. (2019). Engagement in the use of English and Chinese as foreign languages: The role of learner-generated content in instructional task design. *The Modern Language Journal*, *103*(2), 391-411.

Lee, C. (2020). A study of adolescent English learners' cognitive engagement in writing while using an automated content feedback system. *Computer Assisted Language Learning*, 33(1-2), 26-57.

Lyle, J. (2003). Stimulated recall: A report on its use in naturalistic research. *British educational research journal*, *29*(6), 861-878.

Mackey, A., & Gass, S. M. (2015). *Second language research: Methodology and design*. Routledge.

Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback?. *Studies in second language acquisition*, *22*(4), 471-497.

Michel, M., Révész, A., Lu, X., Kourtali, N. E., Lee, M., & Borges, L. (2020). Investigating L2 writing processes across independent and integrated tasks: A mixed-methods study. *Second Language Research*, *36*(3), 307-334.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group*. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine*, *151*(4), 264-269.

Molenaar, I., & Knoop-van Campen, C. A. (2018). How teachers make dashboard information actionable. *IEEE Transactions on Learning Technologies*, *12*(3), 347-355.

Myrtil, M. J., Lin, T. J., Chen, J., Purtell, K. M., Justice, L., Logan, J., & Hamilton, A. (2021). Pros and (con) flict: Using head-mounted cameras to identify teachers' roles in intervening in conflict among preschool children. *Early Childhood Research Quarterly*, *55*, 230-241.

Obersteiner, A., & Tumpek, C. (2016). Measuring fraction comparison strategies with eye-tracking. *ZDM*, *48*(3), 255-266.

Omodei, M. M., & McLennan, J. (1994). Studying complex decision making in natural settings: using a head-mounted video camera to study competitive orienteering. *Perceptual and motor skills*, *79*(3_suppl), 1411-1425.

Peng, J. E. (2020). Teacher interaction strategies and situated willingness to communicate. *ELT Journal*, 74(3), 307-317.

Ramnarain, U. D., & Modiba, M. (2013). Critical friendship, collaboration and trust as a basis for self-determined professional development: A case of science teaching. *International Journal of Science Education*, *35*(1), 65-85.

Rassaei, E. (2013). Corrective feedback, learners' perceptions, and second language development. *System*, *41*(2), 472-483.

Rassaei, E. (2020). The separate and combined effects of recasts and textual enhancement as two focus on form techniques on L2 development. *System*, *89*, 102193.

Sato, R. (2019). Fluctuations in an EFL teacher's willingness to communicate in an English-medium lesson: an observational case study in Japan. *Innovation in Language Learning and Teaching*, *13*(2), 105-117.

Schindler, M., & Lilienthal, A. J. (2019). Domain-specific interpretation of eye tracking data: towards a refined use of the eye-mind hypothesis for the field of geometry. *Educational Studies in Mathematics*, *101*(1), 123-139.

Schindler, M., & Lilienthal, A. J. (2020). Students' creative process in mathematics: insights from eye-tracking-stimulated recall interview on students' work on multiple solution tasks. *International Journal of Science and Mathematics Education*, *18*(8), 1565-1586.

Shintani, N. (2016). The effects of computer-mediated synchronous and asynchronous direct corrective feedback on writing: a case study. *Computer Assisted Language Learning*, 29(3), 517-538.

Thararuedee, P., & Wette, R. (2020). Attending to learners' affective needs: Teachers' knowledge and practices. *System*, *95*, 102375.

Van der Kleij, F., Adie, L., & Cumming, J. (2017). Using video technology to enable student voice in assessment feedback. *British Journal of Educational Technology*, 48(5), 1092-1105.

Zhai, X., Fang, Q., Dong, Y., Wei, Z., Yuan, J., Cacciolatti, L., & Yang, Y. (2018). The effects of biofeedback-based stimulated recall on self-regulated online learning: A gender and cognitive taxonomy perspective. *Journal of* Computer Assisted Learning, 34(6), 775-786.

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

A Review of Academic Emotions Recognition and Analysis with Emphasis on E-Learning Environment

Jie Gao [1], Yantao Wei [2*], Yafei Shi [3], Yuqin Yang [4]
Hubei Research Center for Educational Informationization Central China Normal University
[*] yantaowei@ccnu.edu.cn

Abstract: *Academic emotions are ubiquitous in learning settings. They correlate to students' motivation, cognitive resources, learning strategies, and academic achievement. This study discusses the category of academic emotions, the academic emotions recognition methods in E-learning environment, and empirical studies on the intervention of academic emotions. This review paper will discuss the current research on academic emotions in three parts. For the first part, starting from the theory and category of academic emotions, we expound the classic Cognitive-Motivational model and Control-Value theory of academic emotions and then sort out the category and dimensions of academic emotions in detail. For the second part, we have made a detailed arrangement of the existing seven methods of identifying academic emotions. For the last part, we introduce the empirical research of current academic emotion intervention from two aspects: qualitative analysis and quantitative research. Finally, we summarize the current work on academic emotions and look forward to future research. The existing research shows that: 1) the current research on the category of academic emotions has not been unified; 2) the existing measurement methods of academic emotions are challenging to scale application; 3) the empirical study of academic emotion intervention needs further discussion.*

Keywords: academic emotion category, emotion recognition methods, academic emotion analysis

1. Introduction

Academic emotions play an important role and are ubiquitous in learning settings. However, academic emotions in early studies are often neglected by educational psychology. This research originated from Reinhard Pekrun's initial exploration of students' anxiety in test and learning motivation and Raymond Perry's exploration of academic control and motivation (Pekrun, 1991; Pekrun, 1992; Perry & Magnusson, 1989). We found that students' academic emotions in the learning process were largely neglected. Pekrun groups students' academic emotions according to their valence in positive and negative (Pekrun, Goetz, & Titz, 2002), and found that students' positive emotions occurred more frequently than negative emotions in epistemic activities (Pekrun, Vogl, Muis, & Sinatra, 2017). The researchers have shown that academic emotions significantly impact students' motivation, cognitive resources, learning strategies, self-regulation, and academic achievement (Pekrun, Goetz, & Titz, 2002). The main research content of this review and discussion is the theoretical basis of academic emotions and category structure, measurement of learners' academic emotions, and empirical study of intervention on learners' academic emotions.

The main literature sources of this review paper are scientific academic journals and various conference papers, and the retrieval time is from 2002 to 2022. The main electronic database resources retrieved by scientific academic journals come from Science Direct, most of these journals are indexed

by SCI, SSCI and EI. Conference papers are mainly retrieved from the Institute of Electrical and Electronics Engineers (IEEE). This paper consists of three parts. For the first part, we search the keywords 'emotion theory' and 'emotion model' from ScienceDirect and found a total of 90,868 results, from which we filtered the articles in the context of learning scenarios. For the second part, we searched the keyword 'academic emotion recognition' on IEEE for nearly 20 years of research and focused on the recognition of students' academic emotions in learning scenarios, showing a total of 115 related articles. The search for keywords such as 'self-report method', 'physiological signals', and so on to find articles with specific emotion recognition methods. For the third part, we mainly search the keywords 'emotion analysis', and 'academic emotion analyses' from ScienceDirect, focusing on qualitative research and quantitative analysis of students' academic emotions in learning scenarios.

2. Academic Emotions Theory and Category

Emotions play an important role in students' learning process. In this article, academic emotions summary research, academic emotions theory is the premise of dividing structure and category. To our knowledge, emotion theory mainly groups into two categories: subjective manifestation and emotion structure (Lopatovska & Arapakis, 2011). From the perspective of subjective manifestation, it is divided into subcategories of cognitive factors (i.e., thought or cognitive judgment) and somatic factors (i.e., bodily response) under the subject category and lay theories under the observer category. The emotion structure is divided into discrete structure (i.e., basic emotions) and continuous structure (i.e., dimensional approaches). This paper's theory of academic emotions mainly combines educational psychology with the learner's cognitive factors, including the cognitive-motivational model and control-value theory. Summarize the category of academic emotions by combining discrete emotions and continuous emotions.

2.1. Academic Emotions Theoretical Frameworks

2.1.1. Cognitive-Motivational Model

The cognitive-motivational model was proposed by Pekrun (Pekrun, 1992c), which assumes that learner cognition and some motivational mechanisms regulate the influence of emotions on learning and achievement. Among them, the most important are learning motivation, strategies, cognitive resources, academic achievement and learning self-regulation. For example, instructors could trigger, maintain or inhibit learning motivation and related volitional processes by guiding students' specific emotions. Positive academic emotions help motivate learners to use more flexible and creative learning strategies, while negative emotions mean emotional deactivation, which leads to students' shallow processing of learning information. Furthermore, academic emotions can also direct the subject's attention to emotional objects, which means that emotions use cognitive resources and distract the subject's attention from the task (Pekrun, 2002).

2.1.2. Control-Value Theory

The learning environment is full of achievement emotions, such as enjoyment of learning, hope, pride, anger, anxiety, shame, hopelessness, and boredom (Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011). Control-value theory is used to discuss students' achievement emotions. It believes that human emotion can be induced and regulated in various ways (Pekrun, 2000), which include genetically based emotions, neurophysiological mediated emotions, conditioned emotions, cognitively mediated emotions, and habitualized emotions.

Genetically based emotions mean humans experience emotions early in life, for example, altophobia, fear of some animals, etc. There is also a neuro-mediated emotion, generally caused by neurochemical mechanisms in the individual body, such as depression. Conditioned emotions refer to the direct connection between situational awareness and marginal emotional response, which is generated by early conditional reflex. Cognitively mediated emotions are caused by individuals' cognitive appraisals of the past, present, and future. Habitualized emotions are usually induced based on the cognitive appraisal. Among them, cognitively mediated and habitual emotions are more important in life and learning scenarios, while the control-value theory of achievement emotions mainly solves cognitively mediated emotions.

The theory also defines two types of cognition: (a) control-related cognitions and (b) value cognitions. Control-related cognition means the subjective appraisal of the relationship between variables. An important type of cognition is the individual's self-evaluation (e.g., self-concept ability, self-efficacy expectancy). Actions, situations, and results have intrinsic value and external values to individuals. The intrinsic value is its inherent attribute, and the external value refers to achieving other values.

2.2. Academic Emotions Category

The classification of academic emotions is discussed from two aspects: cognitive factors and somatic factors. The cognitive factors level is mainly based on the cognitive-motivational model and control-value theory. For the control-value theory, academic emotions are regarded as a student's emotions experienced in academic contexts such as class-related, learning-related, and test-related environments and are described by subjective control and perceived value by learners (Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011). Pekrun first developed academic emotions questionnaire (AEQ), a self-report method to measure students' enjoyment, hope, pride, relief, anger, anxiety, shame, hopelessness, and boredom. It divides academic emotions into nine categories and explains them positively and negatively. The authors of Pekrun & Stephens (2010) proposed a three-dimensional includes object focus, valence, and activation category of achievement emotions by using Pekrun's (2006) control value of achievement emotions as a theoretical framework. The study distinguishes between activity-related and outcome-related achievement emotions, divided into positive emotions, negative activation emotions, and negative deactivating emotions. Based on the valence dimension, the study integrated the activation dimension and developed Achievement Emotions Questionnaire (AEQ), which grouped achievement emotion into positive activating/deactivating, and negative activating/deactivating dimensions from 9 academic emotions (Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011). In recent studies, the students' feeling of confusion has been widely observed, and deemed that students' cognitive imbalances and confusion in the learning environment can create opportunities for deep learning of difficult conceptual content (D'Mello and Graesser, 2012). Furthermore, Pekrun's (2017) study measures emotions during epistemic activities with positive emotions, confusion, and negative emotions. Some of academic emotions category structures are represented in Table 1.

In recent studies, under the guidance of computer vision and deep learning, the category of academic emotions is mainly somatic factors, because the explicit features of individuals are easier to be captured by computers and accurately recognition. For example, Ashwin & Guddeti (2020) propose an affective database using the students' facial expressions, hand gestures, and body gestures to collect frustration, confusion, engaged, boredom, sleepy, and neutral. The study also identifies academic emotions through multimodal data, including engaged, confusion, frustration, and boredom. The study also used the Pleasure-Arousal-Dominance (PAD) emotion model to construct a multimodal database of students'

spontaneous emotion in emotional classroom (Wei, Sun, He, & Yu, 2017). Some of academic emotions category by using computers are presented in Table 1.

Table 1. Academic emotions category.

Academic emotions	Dimensional	References
enjoyment, hope, pride, relief, anger, anxiety, shame, hopelessness, boredom.	Positive, negative	Pekrun, Goetz, & Titz (2002)
enjoyment, hope, pride, relief, anger, anxiety, shame, hopelessness, boredom.	Object focus, valence, activation	Pekrun & Stephens (2010)
enjoyment, hope, pride, relief, anger, anxiety, shame, hopelessness, boredom.	Positive activating, positive deactivating, negative activating, negative deactivating	Pekrun, Goetz, Frenzel, Barchfeld, & Perry (2011)
surprise, curiosity, enjoyment, confusion, anxiety, frustration, boredom	Positive, confusion, negative	Pekrun, Vogl, Muis, & Sinatra (2017)
enjoyment, relax, calm, pride, boredom, anger, restlessness, shame, anxiety, frustration	Positive activity-related, negative activity-related, positive outcome-related, Negative outcome-related	Xu & Gong (2011)
enjoyment, anger, boredom	Positive activating, negative activating, negative deactivating	Moeles, Slemp, Oades & Pekrun (2019)
surprise, enjoyment, curiosity, confusion, frustration, anxiety, boredom	Epistemic emotions	Chievrier, Muis, Trevors, Pekrun & Sinatra (2019)
enjoyment, curiosity/interest, anxiety, anger/frustration, confusion, boredom	Positive activating, negative activating, negative, deactivating	Loderer, Pekrun & Lester (2020)
frustration, confusion, engaged, boredom, sleepy, neutral	Facial expressions, hand gestures, body postures	Ashwin & Guddeti (2020)
engaged, confusion, frustration, boredom	facial, heart rate, acoustic modalities	Peng & Nagao (2021)
Boredom, confuse, focus, frustrated, yawning, sleepy	facial expression	Pabba & Kumar (2022)
Bored, tired, confused, concentrated, interested, happy, thoughtful, relaxed, assentient, in a daze	Pleasure-Arousal-Dominance	Wei, Sun, He, & Yu (2017)

3. Academic Emotions Recognition Methods

By summarizing the current research papers, the recognition methods of academic emotions mainly include: self-report method, log data acquisition, voice/text recognition, facial expression recognition, physiological signals, behavior recognition, and multimodal approach. Each recognition method has its advantages and disadvantages. This paper summarizes the use of the above seven methods to measure academic emotions, and basic emotions are not discussed in this paper.

The self-report method is the simplest one. By using this method, researchers ask their feelings of academic emotions by using a questionnaire. The academic Emotions Questionnaire (AEQ), Achievement Emotions Questionnaire, and Test Emotions Questionnaire are developed by Pekrun (2002,

2004, 2011) to measure students' emotions during learning activities. However, it requires a lot of time and human resources and intrusive to E-learning scenarios. By obtaining real-time log data of students' learning process such as his/her time of answering questions, number of mistakes, written sentences, left mouse clicks and so on (Kardan & Einayypour, 2008; Lopatovska, 2009). Voice/text recognition identifies academic emotions through students' acoustic, linguistic, context information, and hybrid features cues (Mannepalli, Sastry, Suman, 2017). According to the results of psychological and medical research, some physiological signals such as brain signals, heart and skin electrical signals can reflect human emotional states (Zhou & Li, 2019; Alzoubi & D'Mello, 2012). However, this method requires students to wear specified devices, which is intrusive to the learning process and expensive and complicated to operate. There is also the recognition of student emotions through gestures, such as hand gestures, 3D body models (Caridakis, Moutselos, & Maglogiannis, 2013; Hayek, Nacouzi, & Kassem, 2014). With the fast growth of computer vision and deep learning, facial expression recognition is increasingly used in emotion classification because of its low cost, real-time acquisition of students' facial expressions, and its ease of use on a large scale. Finally, researchers also measure students' academic emotions by fusing multimodal data, such as facial expressions, vital signals, body postures, eye movements, etc. The academic emotions recognition methods are showed in Table 2.

The seven recognition methods mentioned above are used for student emotion recognition in various learning scenarios. However, in learning scenarios where the classroom environment is the mainstream, emotion recognition still faces huge challenges, such as complex scenes, low image quality, and facial occlusion. Facing these challenges, scholars have addressed these problems through technology optimization. Tingting Liu (2021) proposed a new infrared facial expression recognition method with multi-label distribution learning for understanding non-verbal behaviors in the classroom to address restricted scenarios and low image quality. Yujian Chen and Shiguang Liu (2020) proposed a 20-layer “VGG + residual” CNN network based on the improved VGG16 network, and adapt a hybrid feature strategy to parallelize the Gabor filter with the above CNN. LIU Shuaishi (2014) proposed a facial expression recognition method based on Weber Local Descriptor (WLD) histogram feature and decision fusion to solve the problem which the key information of facial expression missed under partial occlusion.

Table 2. Academic emotions recognition methods

Method	Features	References
Self-report method	Academic Emotions Questionnaire	Pekrun, Goetz & Titz (2002)
	Test Emotions Questionnaire	Pekrun, Goetz & Perry (2004)
	Achievement Emotions Questionnaire	Pekrun, Goetz, Frenzel, Barchfeld, & Perry (2011)
	Academic Emotions Questionnaire for College Student	Xu & Gong (2011)
Log data acquisition	mean time of solving of problems and the number of mistakes	Kardan & Einayypour (2008)
	scrolls up, left mouse clicks	Lopatovska (2009)
Voice/text recognition	acoustic, linguistic, context information, hybrid features	Mannepalli, Sastry, Suman (2017)

	word vector composition	Feng, Qiu & Guo (2019)
	semantic analysis	Toti, Capuano & Campos (2021)
	text classification model	Liu, Peng & Yang (2022)
Facial expression recognition	facial expression	Dukic & Krzic (2022)
	facial expression	Tang, Zhou & Zheng (2019)
	facial expression	Pabba & Kumar (2022)
Physiological signals	EEG	Zhou & Li (2019)
	ECG, EMG, GSR	Alzoubi & D'Mello (2012)
Behavior recognition	hand gestures (glove-based)	Hayek, Nacouzi, & Kassem (2014)
	hand gestures (vision-based) (3D body models)	Caridakis, Moutselos, & Maglogiannis (2013)
Multimodal approach	facial expressions, eye movements, head postures, body movements	Wei, Sun, He, & Yu (2017)
	facial expressions, hand gestures, body postures	Ashwin & Guddeti (2020)
	facial, heart rate, acoustic	Peng & Nagao (2021)

4. Empirical Study on Academic Emotions Analysis

In recent studies, the empirical research on the analysis of academic emotions is mainly from two aspects: qualitative analysis and quantitative research. Researchers' study of academic emotions began with qualitative analysis. For example, Pekrun (2002) developed Academic Emotions Questionnaire (AEQ), referred to in Table2 of the self-report emotion recognition method and found that: academic emotions are significantly related to students' motivation, cognitive resources, learning strategies, self-regulation, and academic achievement. Furthermore, the relationship between learners' positive academic emotions (enjoyment, pride) and academic performance is discussed (Villavicencio, Bernardo, 2013). The results showed that enjoyment and pride positively associated with academic achievement. Students who reported high levels of enjoyment and pride had a significant correlation between self-regulation and academic achievement, while students who thought they had low pride did not correlate with academic achievement. When students think they belong to a low level of enjoyment, it negatively correlates with academic achievement. Liu et al (2022) explored the predictive relationship between discussion posts in MOOC forums and students' academic performance by text analysis and classification of emotional engagement (positive, negative, and confusion) and cognitive engagement. The results show positive academic emotions can promote students' cognition and predict academic performance. Compared with cognitive engagement, emotional engagement has a more significant predictive effect on students' academic performance. The study also discussed the frequency of students' academic emotions in the active learning environment and its relationship with gender differences (Dukic & Krzic, 2022).

From the perspective of quantitative research, the researchers mainly analyzed academic emotions from the level of students' academic emotional changes. Pabba & Kumar (2022) draw a feedback curve of students' overall classroom engagement level by identifying academic emotions to help teachers understand students' learning status and adjust teaching strategies in time. The authors of Zeng et al (2021) propose EmotionCues, a visual analytics system to analyze classroom videos from the perspective of

emotion summary easily and detailed analysis, which integrates group academic emotional changes, individual academic emotional portrait and the characteristics of the relationship between group emotional changes. In addition, some studies explore the evolution of students' classroom emotions through a combination of qualitative analysis and quantitative research methods, and analyze the significant relationship between students' classroom emotions and their department, genders, the class time, location of the computer in classroom, and lecture type.

5. Conclusion and Discussion

In recent years, academic emotions have become an active research area. This paper collects and collates the basic theories and models of academic emotions, the identification methods of academic emotions, and the empirical research on academic emotions analysis. Combined with the current research status, there are still limitations for academic emotion research, including the following three parts:

1) Academic emotion category is not uniform. The current research on the category of academic emotions is mainly divided into two dimensions: educational psychology and computer vision. From the perspective of education psychology, Pekrun's (2002) academic emotions are usually divided into: enjoyment, hope, pride, relief, anger, anxiety, shame, hopelessness, and boredom. On the research level of computer vision and deep learning, scholars pay more attention to the academic emotion recognition of students' explicit behavior, and generally divide academic emotions into: frustration, confusion, engaged, boredom, sleepy, and neutral (Ashwin & Guddeti, 2020; Peng & Nagao, 2021; Pabba & Kumar, 2022). Therefore, the category of academic emotions has not yet constructed a consistent framework.

2) Academic emotion measurement difficult to popularize on a large scale. This paper introduces the existing academic emotion recognition methods, includes self-report method, log data acquisition, voice/text recognition, facial expression recognition, physiological signals, behavior recognition, and multimodal approach. Each method has advantages and disadvantages. Among them, the most widely used in current research is to identify academic emotions through students' facial expressions. However, this method faces many challenges: such as low resolution, occlusion, and complex backgrounds. Moreover, the current academic emotion recognition methods are difficult to be used for real-time feedback and analysis in large-scale classrooms.

3) An empirical study of academic emotion analysis needs further analysis. The research on academic emotions analysis is mainly carried out from qualitative and quantitative research. The existing research focuses more on qualitatively analyzing the relationship between students' academic emotions and learning motivation, cognitive resources, self-regulation and academic performance. Moreover, the analysis of the changes of academic emotions mainly focuses on the rules and characteristics of students' emotions in the classroom. There is no further in-depth exploration of this. However, there are still gaps in the evolution and analysis of students' academic emotional intervention in research.

In summary, in the future research on students' academic emotions, it is urgent to construct a recognized classification model of academic emotions, optimize the identification method of academic emotions, and strengthen its practicability in large-scale classrooms. Finally, it is hoped to discuss further the intervention of students' academic emotions and their related analysis in quantitative research. For example, in the classroom, we use the real-time analysis tool of students' academic emotions to provide real-time feedback for teachers in classroom teaching. Teachers adjust teaching strategies according to the feedback results, and timely intervene in students' learning process to promote and optimize teaching.

Funding

This work was supported in part by the National Natural Science Foundation of China under Grant 62277029, the National Collaborative Innovation Experimental Base Construction Project for Teacher Development of Central China Normal University under Grant CCNUTEIII-2021-19, the Humanities and Social Sciences of China MOE under Grants 20YJC880100, 22YJC880061, the Knowledge Innovation Project of Wuhan under Grant 2022010801010274，and the Fundamental Research Funds for the Central Universities under Grant CCNU22JC011.

References

Pekrun, R. (1991). Prüfungsangst und Schulleistung: Eine Längsschnittanalyse. *Zeitschrift für pädagogische Psychologie*, *5*(2), 99-109.

Pekrun, R. (2019). Expectancy–value theory of anxiety: Overview and implications. *Anxiety: Recent developments in cognitive, psychophysiological, and health research*, 23-42.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational psychologist*, *37*(2), 91-105.

Pekrun, R., Vogl, E., Muis, K. R., & Sinatra, G. M. (2017). Measuring emotions during epistemic activities: The epistemically-related emotion scales. *Cognition and Emotion*, *31*(6), 1268-1276.

Lopatovska, I., & Arapakis, I. (2011). Theories, methods and current research on emotions in library and information science, information retrieval and human–computer interaction. *Information Processing & Management*, *47*(4), 575-592.

Pekrun, R. (2000). A social-cognitive, control-value theory of achievement emotions.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational psychology review*, *18*, 315-341.

Pekrun, R., & Stephens, E. J. (2010). Achievement emotions: A control-value approach. *Social and Personality Psychology Compass*, *4*(4), 238-255.

Lehman, B., D'Mello, S., & Graesser, A. (2012). Confusion and complex learning during interactions with computer learning environments. *The Internet and Higher Education*, *15*(3), 184-194.

Ashwin, T. S., & Guddeti, R. M. R. (2020). Affective database for e-learning and classroom environments using Indian students' faces, hand gestures and body postures. *Future Generation Computer Systems*, *108*, 334-348.

Wei, Q., Sun, B., He, J., & Yu, L. (2017). BNU-LSVED 2.0: Spontaneous multimodal student affect database with multi-dimensional labels. *Signal Processing: Image Communication*, *59*, 168-181.

Xu, X. C., & Gong, S. Y. (2011). Development of an academic emotions questionnaire for college student. *Chinese Journal of Clinical Psychology*.

Camacho-Morles, J., Slemp, G. R., Oades, L. G., Pekrun, R., & Morrish, L. (2019). Relative incidence and origins of achievement emotions in computer-based collaborative problem-solving: A control-value approach. *Computers in Human Behavior*, *98*, 41-49.

Chevrier, M., Muis, K. R., Trevors, G. J., Pekrun, R., & Sinatra, G. M. (2019). Exploring the antecedents and consequences of epistemic emotions. *Learning and instruction*, *63*, 101209.

Loderer, K., Pekrun, R., & Lester, J. C. (2020). Beyond cold technology: A systematic review and meta-analysis on emotions in technology-based learning environments. *Learning and instruction*, *70*, 101162.

Peng, S., & Nagao, K. (2021). Recognition of students' mental states in discussion based on multimodal data and its application to educational support. *IEEE Access*, *9*, 18235-18250.

Pekrun, R., Goetz, T., Perry, R. P., Kramer, K., Hochstadt, M., & Molfenter, S. (2004). Beyond test anxiety: Development and validation of the Test Emotions Questionnaire (TEQ). *Anxiety, Stress & Coping*, *17*(3), 287-316.

Kardan, A., & Einavypour, Y. (2008). Involving learner's emotional behaviors in learning process as a temporary learner model. In *Proceeding of International Conference on Virtual Learning (ICVL), Bucurest, Romania*.

Lopatovska, I. (2009). *Emotional aspects of the online information retrieval process*. Rutgers The State University of New Jersey, School of Graduate Studies.

Mannepalli, K., Sastry, P. N., & Suman, M. (2017). A novel adaptive fractional deep belief networks for speaker emotion recognition. *Alexandria Engineering Journal*, *56*(4), 485-497.

Zhou, Y., Xu, T., Li, S., & Shi, R. (2019). Beyond engagement: an EEG-based methodology for assessing user's confusion in an educational game. *Universal Access in the Information Society*, *18*, 551-563.

AlZoubi, O., D'Mello, S. K., & Calvo, R. A. (2012). Detecting naturalistic expressions of nonbasic affect using physiological signals. *IEEE Transactions on affective computing*, *3*(3), 298-310.

El Hayek, H., Nacouzi, J., Kassem, A., Hamad, M., & El-Murr, S. (2014, April). Sign to letter translator system using a hand glove. In *The Third International Conference on e-Technologies and Networks for Development (ICeND2014)* (pp. 146-150). IEEE.

Dukić, D., & Sovic Krzic, A. (2022). Real-time facial expression recognition using deep learning with application in the active classroom environment. *Electronics*, *11*(8), 1240.

Tang, J., Zhou, X., & Zheng, J. (2019, February). Design of Intelligent classroom facial recognition based on Deep Learning. In *Journal of Physics: Conference Series* (Vol. 1168, No. 2, p. 022043). IOP Publishing.

Feng, X., Qiu, L., & Guo, X. (2019). Recognition method for academic emotions of students' feedback texts based on LSTM model. *Open Educ. Res*, *25*, 116-122.

Liu, S., Liu, S., Liu, Z., Peng, X., & Yang, Z. (2022). Automated detection of emotional and cognitive engagement in MOOC discussions to predict learning achievement. *Computers & Education*, *181*, 104461.

Villavicencio, F. T., & Bernardo, A. B. (2013). Positive academic emotions moderate the relationship between self-regulation and academic achievement. *British Journal of Educational Psychology*, *83*(2), 329-340.

Zeng, H., Shu, X., Wang, Y., Wang, Y., Zhang, L., Pong, T. C., & Qu, H. (2020). Emotioncues: Emotion-oriented visual summarization of classroom videos. *IEEE transactions on visualization and computer graphics*, *27*(7), 3168-3181.

Liu, S., Zhang, Y., & Liu, K. (2014, July). Facial expression recognition under partial occlusion based on Weber Local Descriptor histogram and decision fusion. In *Proceedings of the 33rd Chinese Control Conference* (pp. 4664-4668). IEEE.

Chen, Y., & Liu, S. (2020, December). Deep Partial Occlusion Facial Expression Recognition via Improved CNN. In *Advances in Visual Computing: 15th International Symposium, ISVC 2020, San Diego, CA, USA, October 5–7, 2020, Proceedings, Part I* (pp. 451-462). Cham: Springer International Publishing.

Liu, T., Wang, J., Yang, B., & Wang, X. (2021). Facial expression recognition method with multi-label distribution learning for non-verbal behavior understanding in the classroom. *Infrared Physics & Technology*, *112*, 103594.

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

Developing Students' Interest in STEM Based on Cultural Relics

Yumeng Li[1*]
[1] Heze Mudan Economic Development District Experimental School
* 442672085@qq.com

Abstract: *This project aims at increasing interest in STEM related subjects of high school students. As students show more interest in Chinese History, this project chooses cultural relics as learning materials to help students explore knowledge, apply knowledge, and create something new. 150 students from grade 4 to grade 6, 28 history teachers and 28 physics teachers from junior high schools, 9 science teachers from primary schools participated in this project. Results show that the History+STEM Learning Mode can arouse more interest in STEM.*
Keywords: STEM education, cultural relic, learning interest

1. Background

In the 21st century, as we live in a information-based and highly technological society, it's better for students to develop STEM capability. STEM capability refers to integrating knowledge and 21st skills to solve authentic problems innovatively. To prepare students for meeting the needs of STEM capability and whole-person development, from 2016, Hong Kong Education Bureau started to promote STEM education, which aims to build strong foundations for STEM literacy and prepare the STEM workforce for the future (Education Bureau, 2016). To popularize STEM education, the Education Bureau widely provided funding, professional training, learning and teaching resources (Council, 2018). STEM related subjects like Combined Science and Integrated Science were provided in senior high schools.

However, as shown in Table 1, school candidates who chose Combined Science and Integrated Science in Hong Kong Diploma of Secondary Education Examination(HKDSE) became fewer and fewer (HKEAA, 2022). On the contrary, more and more candidates chose liberal arts subjects like Chinese History. Su (2009) reported that most students choose subjects based on interest. So maybe the current STEM education can't arouse students' interest in STEM. From this point, this project chose cultural relic as learning material to investigate whether the History + STEM Learning Mode can arouse more interest in STEM.

Table 1. Percentage of entered school candidates of each subject in HKDSE

Subject / Year	Combined Science (%)	Integrated Science (%)	Chinese History (%)
2016	2.481	0.292	11.122
2017	1.739	0.256	11.390
2018	1.493	0.199	11.524
2019	1.240	0.172	12.545
2020	0.937	0.207	12.351

2021	0.682	0.165	13.198
2022	0.618	0.168	13.379

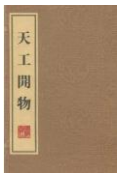
2. History + STEM Learning Mode

2.1. Introduction of History + STEM Learning Mode

Combine history and STEM means: Use history knowledge(e.g. cultural relics) as learning material to learn STEM related knowledge or improve STEM literacy. As shown in Table 2, there have been several STEM projects chose cultural relics as learning materials. They can be divided into seven main categories besed on the application of cultural relics and shortcomings of projects. From the shortcomings we can say few projects could fully make use of cultural relics.

Table 2. Introduction and analysis of previous projects

No.	Application of cultural relics and shortcomings of projects	Cultural relics
1	Arouse students' attention at the beginning of lessons (Liu & Li, 2022) No further exploration of cultural relics	shadow play, dragon boat, lost-wax process   
2	Students use cultural relics as experiment tools (Liu & Li, 2022) No application of knowledge	shadow play 
3	Students follow instructions to try the making process (Li, 2015) Lack knowledge learning and innovation	tie-dye 
4	Make students become interested in cultural relics (Fan, 2018) Lack feedback on students' findings or answers	chime 
5	Make students be proud of our nation (Wang, 2021) Almost have nothing to do with exlporation, knowledge learning and application	the great four inventions of ancient China 

6	Students try to find their shortcomings and develop them (Sui, Li, Sh, & Ch, 2015) Hard to make students be proud of cultural relics	tie-dye, papermaking, movable type 
7	Teach in classical Chinese (Dai & Zheng, 2021) No changes in learning process	ancient science and technology books 

2.2. Introduction of Oil-saving Lamp

This project chose oil-saving lamp as learning material (Figutr 1). It consumes less oil than normal oil lamp because of an empty interlayer (Figure 2). Add cold water into interlayer through the hole, then the temperature of oil will drop because of heat conduction between oil and water. Lower temperature makes the evaporation of oil slow down. Though water will get hot, the water evaporation will constantly lower water temperature and continue to cool the oil down and slow down its evaporation (Zhang, 2018).

After hundreds years, the hole became bigger and bigger. Finally the empty interlayer was opened as Figure 3 to expand evaporation area. Both area and temperature affect evaporation.

Figure 1.

Figure 2.

Figure 3.

2.3. Details of STEM Activities

This project used a time travel story as context because story can attract more attention, make students be curious about the lesson, and help remember what they learned for a longer time (Zhu, 2016). As teleplays about time travel are very popular, the context of this project is a travel in ancient times. As ancient students, they need to use oil lamp to read book. To save oil, they find oil-saving lamp and study the knowledge of it.

Now the STEM education is mainly conducted in three forms: learning integrated knowledge, application or invention of technology, creativity and innovation. In this project, these three fields are all involved. The oil-saving lamp integrates knowledge about heat conduction and evaporation. After learning knowledge, students learned their application in high-tech. These examples let students know that innovation can be the application of easy knowledge. Last, students were encouraged to share examples and apply easy knowledge to create something or improve something.Details of this project are as Table 3:

Table 3. Details of STEM Activities

Activities	STEM related content	Further application or innovation

1. Go back to ancient times	Pre-test about students' interest in STEM; Introduction of oil-saving lamp	
2. How does oil disappear?	Concept of evaporation	
3.1 Why does oil-saving lamp save oil?	Temperature affects evaporation	Temperature-controlled work; Application of easy knowledge
3.2 When cold water gets hot	Heat energy is absorbed in evaporation	Pottery refrigerator; Cooling rod of Qinghai-Tibet Railway; Application of easy knowledge
3.3 The development of oil-saving lamp	Evaporation area; Engineering design process	Redesign hanger
3.4 Do you agree with official introduction?	What other factors affect evaporation? Knowledge integration	
4.1 Other cultural relics	Heat convection; gyroscope	Aerospace
4.2 What will you do in the future?	Post-test about students' interest in STEM	

2.4. Participants and Tests

2.4.1. Students

150 students from grade 4 to grade 6 participated in this project and 101 of them finished questionares.

Queationare consists of pre-test in form of choice questions and post-test in form of choice questions and open questions (Table 4). The context for pre-test is they have gone back to ancient times; the context for post-test is they have the opportunity to go back to modern society. Choice questions can directly show students' preference between history and STEM. Options such as *I want to learn technologies* and *I want to explore knowledge behind relic* show students are interested in STEM; options such as *I like history* and *I want to visit celebrities* indicate students are interested in history. The amount of options about history and STEM are the same to avoid bias. As for open questions, they were encouraged to point out the content that they thought was interesting, boring, or hard to understand. The comparison between pre-test and post-test shows whether their interest in STEM can be improved by history+STEM Learning Mode, especially those who prefer history at first.

Table 4. Pre-test and Post-test for Students

Tests	Options	Preference
Pre-test	I can't wait to learn history.	History
	I want to visit celebrities.	
	I wan to learn technologies that fail to be passed down.	STEM
	I can apply science knowledge and daily materials to make something new.	
Post-test	I will stay here to learn history.	History
	I enjoy learning history.	
	I will introduce ancient history to more people.	
	I will continue studying knowledge behind cultural relics.	STEM
	I will stay here to explore technologies that fail to be passed down.	
	I will learn from cultural relics to make modern products.	

2.4.2. Teachers

Teachers were invited to give professional comments. 28 history teachers and 28 physics teachers from junior high schools, 9 science teachers from primary schools finished questionares. Some teachers gave further suggestions in interview based on students' questionares and daily learning performance.

Questionares for teachers are different (Table 5, Table 6). Apart from the same questions, question in Table 5 *Which topic in ancient Chinese history attracts more students?* is to investigate more topics that can be used as learning materials in STEM lesssons. The last three questions in Table 6 is to improve learning efficiency.

Table 5. Questionare for History Teachers

How many students attend lesson actively?
What's your opinion about the History + STEM learning mode?
Which topic in ancient Chinese history attracts more students?

Table 6. Questionare for Physics Teachers and Science Teachers

How many students attend lesson actively?
What's your opinion about the History + STEM learning mode?
What are the most attractive ways to learn?
What are the most effective ways to learn?
What are the most effective ways to help students apply knowledge?

3. Results and Discussion

3.1. Analyses of Students' Pre-test and Post-test

Among 101 questionares, only 16 questionares show more interest in history in pre-test, the other 85 questionares show more interest in STEM. For the 16 questionares, the percentage of STEM related options is just 19%. In post-test, the percentage rises to 66%.

Among the other 85 questionares, 65 were assigned by science teachers, 20 were assigned by Chinese teachers. In view of the potential influence from science teachers, it's necessary to analyze the 20 questionares separately. In pre-test, the percentage of STEM related options is 55%. After activities, it rises to 72%.

These changes show that the History+STEM Learning Mode can arouse more interest in STEM.

3.2. Results of Open Questions

33 students answered open questions. 10 of them think learning STEM based on cultural relics is boring and hard to understand and they prefer authentic context or learning the knowledge of high-tech. The other 23 students show positive attitude towards History+STEM Learning Mode.

So we can conclude that the History+STEM Learning Mode can make students who prefer history be more interested in STEM.

3.3. Analyses of Teachers' Opinions

The first comparison is about students’ engagement in history lessons and physics lessons (Table 7). It’s obvious that students perform more actively in history lessons so using history knowledge as learning marerials may help improve interest or engagement in learning STEM.

Table 7. Students’ Engagement in History Lessons and Physics Lessons

Students' Engagement	History Lessons(%)	Physics Lessons(%)
More than half students are active	75.9	57.1
Half students are active	6.9	25
Only several students are active	3.4	7.1
More than half students can concentrate on the lesson	34.5	42.9
More than half students seems bored	10.3	0

Then is teachers' opinions on the History+STEM Learning Mode (Table 8). From the first three options we can see science teachers of primary schools show more interest in the learning mode. The rest of opinions show that the integrated learning will be easier to conduct in primary schools. We may explain the results from three aspects. First, STEM activities and integrated learning(e.g. general education) are more popular in primary schools and usually all teaching tasks are conducted by one teacher, so primary teachers have been familiar with the integrated mode and they don't need to collaborate with other teachers. Second, primary students mainly learn knowledge related to daily life so it's easy to find common point among different disciplines. Third, one major learning objective for primary students is to arouse interest, so teachers are willing to try different learning mode.

Table 8. Teachers' Opinions on the History+STEM Learning Mode

Opinions	Physics Teachers of Junior High Schools (%)	Science Teachers of Primary Schools (%)
Students' perspectives on the relationship between disciplines, the subject registration and career selection will change	67.9	77.8
This mode is new so students will be curious and perform actively (especially for students prefer history)	25	55.6
Students may have a try after activities	21.4	55.6
There is little in common among history and STEM so it's difficult to provide suitable and enough learning resources	17.9	11.1
Learning objectives are not clear so students will feel confused and difficult to prepare for exams	17.9	11.1
Each discipline has its own system so it's better to learn separately	21.4	11.1
It's hard to collaborate with other teachers	28.6	0
Others	3.6	0

Then from Table 9 we can know that most students are interested in stories. Besides, in interview, one teacher suggested to learn by role-play. So the History+STEM Learning Mode can try to involve more stories rather than just focusing on cultural relics.

Table 9. Students' Engagement in Different Topics

Topics in Ancient Chinese History	Students' Engagement (%)
Political system	10.7
Economics	3.6
Stories about celebrities	89.3
Technology and architecture	21.4

The foundation and extinction of dynasties	39.3
Legend	57.1
Culture and art	50
Diplomacy	21.4
War	35.7

Last, Table 10 shows that teachers think students benefit the most from games and experiments. So STEM projects should include more activities.

Table 10. Science Teacher' Opinions on Different Learning Methods

Learning Methods	Most Interesting (%)	Most Effective (%)	Most helpful in Facilitating Applying Knowledge (%)
Read science stories	44.4	0	11.1
Watch videos	77.8	0	22.2
Play games and do experiments	100	100	88.9
Read encyclopaedia	22.2	0	11.1
Study models	66.7	0	55.6
Visit science and technology museums	66.7	0	44.4

3.4. Reflection on Teachers' Interviews

3.4.1. I never thought about using unfamiliar context

Almost all STEM activities have three features: authentic context, integrated content, solving authentic problems. Teachers always select examples that students are familiar with to help students understand knowledge. But a science teacher told me that even the context was their daily life, many students didn't show any interest and always act as audience. But in their post-test, several *audience* wrote down which section they couldn't understand. The teacher felt amazed about students' change. She said *I never thought about using unfamiliar context, maybe the reason why students are not interested in my lessons is because they are so familiar with daily life context that they won't be curious anymore.* From this perspective, we should try new context to arouse interest.

3.4.2. I should design different activities for different students

Usually, students of the same grade attend the same activities, but questionares show that activities that fit students' preferences are more effective. One teacher said *I thought students won't prefer science or liberal art until grade 10 because knowledge is very easy from grade 1 to grade 9. But now it seems that interest has nothing to do with difficulty so I should design different activities for different students.*

3.4.3. It's necessary for students that prefer liberal art to attend STEM education

This Chinese teacher told a story. In their school, during summer, cleaners spray water on ground after noon everyday. Once a students asked him why. He told students *to clean dust*. But in the sharing activity of section 3.2, this students wrote *Spray water on ground. The evaporation of water can absorb heat energy, then we feel cool*. The Chinese teacher deeply regretted what he answered. He said *I majored in liberal art, so I seldom analyze from the perspective of science. On the contrary, I always rely on experience. But our society develops so fast that our experience is not enough to analyze and solve problems. Besides, the thinking style is as important as knowledge. So it's necessary for students that prefer liberal art to attend STEM education.*

3.4.4. Creativity can be a little change or a new method

Creativity competitions are very popular in primary schools. Most competitions require students to

make a product. To win the competition, students always apply programming or high-tech tools such as 3D printing, but the products may not be really creative. In this project, students were encouraged to apply easy knowledge to change something and they shared many interesting and creative answers. One creative and practical answer is *When there is just a little ink in ink bottle, it's hard to take it out. Then I pour it into bottle lid. V=S·h. V is certain, the area of lid is smaller than that of bottle base, so ink will be higher*. This students didn't create any product, he created a new method with very easy knowledge. His teacher said *We focus too much on high-tech tools and try to make products more and more complex, but creativity can be a little change or a new method.*

4. Conclusion

STEM education is necessary for all students and the learning mode should cater for students' interest. This project proves that the History+STEM Learning Mode can arouse students' interest and it's easy to popularize in primary schools to improve students' engagement in STEM.

5. Limitation and Further Development

5.1. Lack participants

The History+STEM Learning Mode is initially designed for Hong Kong students that are more interested in history, but because of COVID-19, all participants were from mainland China and only 16 students were more interested in history in pre-test. So more students are welcome to improve reliability and validity.

5.2. Difficult content

The knowledge of this project is for students of grade 7, but participants were from grade 4 to grade 6. This may influence their understanding and attitudes towards STEM. Even so, the results show that this learning mode is effective.

5.3. No further interview with students

For further development, I consulted teachers about attractive learning materials and effective learning methods. But I ignored that students are the center of activities so I should also get feedback from students. Besides, for those *audience* who became active in this project, we need to know the reasons behind so as to change more students.

References

Dai, J. Q., & Zheng, L. P. (2021). Application of Tian Gong Kai Wu in Chemistry Classroom. *Chinese Journal of Chemical Education, 42*(5), 50-53.

Education Bureau (2016). *Report on promotion of STEM education: Unleashing potential in innovation.* Hong Kong, China: Education Bureau.

Fan, L. (2018). *Research on the Integration of Traditional Science and Technology into Kindergarten Curriculum*(PhD dissertation, Southwest University).

Hong Kong Examinations and Assessment Authority. (2022). *2013 - 2022 HKDSE Statistics of candidates' results in each of the Category A subjects.* Retrieved from https://www.hkeaa.edu.hk/DocLibrary/HKDSE/Exam_Report/Examination_Statistics/dseexamstat22_7.pdf

Legislative Council(2018). *LCQ22: Promotion of education on subjects related to Science, Technology, Engineering and Mathematics*. Retrieved from

https://www.info.gov.hk/gia/general/201811/28/P2018112800583.htm?fontSize=1

Li, Y. (2015). Intangible Cultural Heritage Inheritance Education in Nantong Vocational College of Science and Technology. *Vocational and Technical Education, 36*(35),1.

Liu, J., & Li, Q. Y. (2022). Use Local Museum Collections to Develop Physics Teaching Resources in Middle School. *Journal of Physics Teaching, 40(8), 69-72.*

Su, S. S. (2009). Research on the Influence Factors of Senior High School Students' Intention of the Division of the Arts and Science. *CURRICULUM, TEACHING MATERIAL AND METHOD, 2009*(10), 14-16.

Sui, J. Zh., Li, C. Y., Sh, J., & Ch, Y. L. (2015). Ancient Science and Technology---Case Study of Teenagers' Science and Technology Activities. *Museum-School Collaboration in the Context of Global Science Education Reform -- Proceedings of the 7th Symposium on Museum-School Collaboration Science Education*(pp198-203). Science Popularization Press.

Vasileios Symeonidis, & Johanna F. Schwarz. (2016). *Phenomenon-Based Teaching and Learning through the Pedagogical Lenses of Phenomenology: The Recent Curriculum Reform in Finland*. Forum Oświatowe, 28(2(56)), Forum oświatowe, 2016-12-01, Vol.28 (2(56)).

Wang, N. (2021). The Exploration of High School Science and Technology History Teaching From the Perspective of STS Education -- Taking "The History of the Creation and Utilization of Four Great Inventions" as an Example. *Teaching Reference of Middle School History,* (20), 4-7.

Zhang, Y. K. (2018). *The study of the Qiong kiln Oil-saving lamps in the Tang and Song Dynasties*(Master dissertation, Jiangnan University).

Zhu, Y. W. (2016). "Stories"Endow Science Museums with "Power"---Thoughts on "the Power of Stories" of "Dimensions". *Journal of Natural Science Museum Research* (01), 73-81.

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

A Study on the Influence Mechanism of Data Literacy Based on HLM Model

Hui, Zhang [1*]
[1] East China Normal University
[*] 51214108043@stu.ecnu.edu.cn

Abstract: *Data are reshaping the way of humans' thinking and behavior, and the data-driven living and working environment requires a new generation of citizens to be data literate. However, although research on data literacy cultivation continues to grow, the factors and mechanisms that influence data literacy have not been thoroughly explored. This study uses a hierarchical linear modeling (HLM) approach based on the Educational Productivity Model and the Program for International Student Assessment (PISA 2018) to analyze the factors that influence students' data literacy from multiple dimensions and perspectives. The findings indicate that students' cognitive strategies of comprehension and memory, generalization, and assessing credibility, students' resilience and goal-oriented non-cognitive dimensions, family's economic and socio-cultural level and parents' emotional support, school context, and teacher level all have significant effects on data literacy.*
Keywords: data literacy, PISA, hierarchical linear model, secondary education

1. Introduction

Data literacy refers to the process of extracting meaningful information from data, such as knowing how to read or visualize data, and identifying understandable pattern relationships in data (Kotu & Deshpande, 2019). In recent years, many scholars in the fields of library science, education, and statistics have tried to give the connotation of data literacy, and domestic scholars mostly take the perspective of structural attributes, and Meng & Li (2014) first defined the connotation of scientific data literacy, which is divided into data awareness, basic data knowledge and skills, and discovery analysis to solve problems. Huang & Li (2016) emphasizes the importance of data ethics in data literacy, arguing that data literacy includes three aspects: data awareness, data competence and data ethics. Foreign scholars, on the other hand, take a more elemental skills perspective, with Vahey et al. (2012) defining data literacy as understanding, collecting, and visualizing qualitative or quantitative data to explain supporting arguments.

Data is changing the way people think and act, and the data-driven living and working environment makes data literacy a must for the next generation of qualified citizens. Data literacy education can improve students' data skills and critical thinking to meet the development needs of the big data era. To compete for a strategic position in the data age, governments are implementing policies to support data literacy education. in 2012, the leading U.S. passed Big Data Research and Development Initiative to increase funding for data science and improve the nation's ability to derive knowledge and insights from massive amounts of data (Alley-Young, 2017). In 2015, the Chinese government released the Action Plan for Promoting Big Data Development to promote data literacy and education and to raise the overall level of data awareness in society.

The development of data literacy is a complex process, and in order to effectively promote students' data literacy development, the intrinsic and extrinsic factors that influence its development need to be clarified. However, there is a paucity of research on data literacy influencing mechanisms and most implement some sort of instructional intervention for learners to qualitatively analyze data literacy development pathways, such as a design study conducted by Kahn (2020) on the topic of family geography biographies that qualitatively analyzed student interaction videos to understand how the research design influenced students to place family and self in the data and the role of family members in writing the storyline.

In 1978, Walberg first proposed the theory of educational productivity, emphasizing that differences in learners' academic performance arise from the synergistic effect of individual ability and external environment, and that educational productivity should consider students' ability motivation, the quality and quantity of teaching in schools, and the social and family environment (Walberg, 1980). Based on the psychological model of educational productivity and combined with the PISA 2018 test questionnaire, this study proposes a multidimensional assessment model of the factors influencing data literacy, as shown in Figure 1.

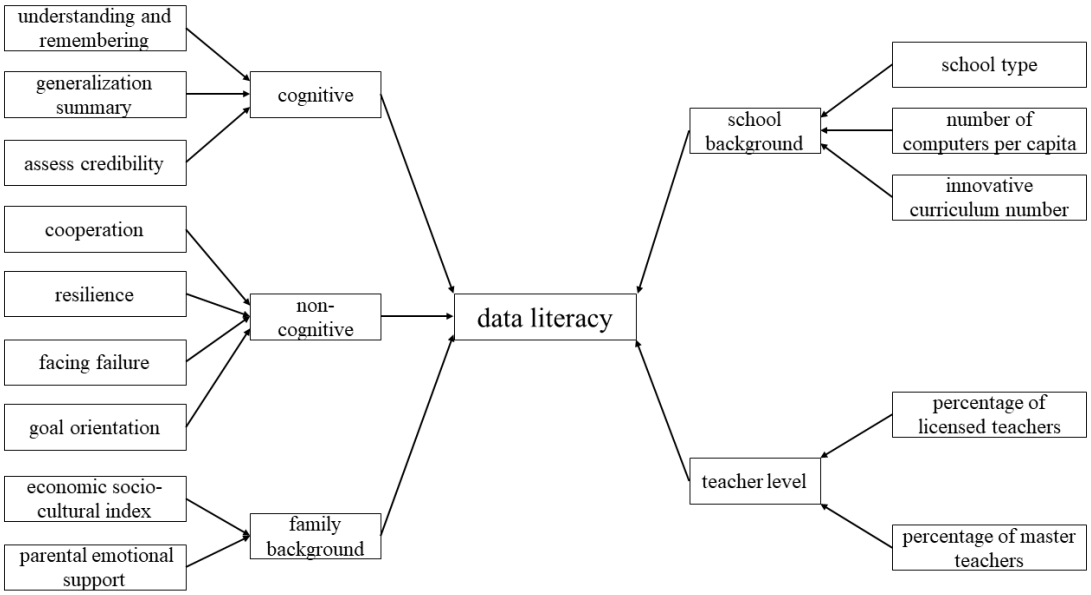


Figure 1. A model for influencing student data literacy.

Accordingly, this study poses the following research questions:

(1) Do student-level factors have an impact on data literacy? Is the influence effect significant?

(2) Do school-level factors have an impact on data literacy? Is the impact effect significant?

2. Methodology

2.1 Data

The data used for the study were derived from selected data from the 2018 Program for International Student Assessment (PISA), conducted by the Organization for Economic Cooperation and Development (OECD), for secondary school students aged 15. The PISA test, which began in 2000 and is administered every three years, is one of the most influential student assessment programs in the world. 79 invited countries participated in 2018, surveying students, parents, schools and teachers for background and trait information, and examining students' reading literacy, mathematical literacy and scientific literacy. Four

Chinese provinces and cities, Beijing, Shanghai, Zhejiang, and Jiangsu, were represented in the assessment, with 361 schools, 17,634 teachers, and 12,058 students completing the questionnaire. Because this study involved missing values for both the independent and dependent variables data literacy, 11,190 student data and 355 school data remained after excluding the missing data.

2.2 Variables

The independent variables for this study were selected from a total of nine variables in the PISA student dataset for cognitive dimensions (understanding and remembering, generalization summary, assess credibility), non-cognitive dimensions (cooperation, resilience, facing failure, goal orientation), and family situation (economic socio-cultural index, parental emotional support), and from the school dataset for school context characteristics (school type, number of computers per capita, innovative curriculum number) and teacher level (percentage of licensed teachers, percentage of master teachers) for a total of 5 variables in the school dataset. Data literacy is a latent trait of learners, and the traditional classical test method uses the percentage of correct answers as the student's ability level, which relies heavily on the sample of test questions, so PISA tests students' data literacy based on item response theory (IRT), which describes the probability $\beta(\theta)$ of a student answering a test question correctly to estimate the student's ability level θ. The student data set provides 10 data literacy ability level estimates, and the study took their mean as the dependent variable data literacy proficiency level. A detailed description of each variable in this study is provided in the table 1.

Table 1. Variable details.

	Category	Variable	Description	Mean	Standard deviation
Student level	Student meta-cognitive	UNDREM	Meta-cognition: understanding and remembering	0.21	0.993
		METASUM	Meta-cognition: summarizing	-0.11	0.962
		METASPAM	Meta-cognition: assess credibility	0.10	0.964
	Family situation	ESCS	Index of economic, social and cultural status	-0.35	1.084
		EMOSUPS	Parents' emotional support perceived by student	0.02	0.932
	Student non-cognitive	PERCOOP	Perception of cooperation at school	0.24	1.006
		GFOFAIL	General fear of failure	0.01	0.874
		RESILIENCE	Resilience	-0.07	0.957
		MASTGOAL	Mastery goal orientation	0.07	0.907
School level	School background	SCHLTYPE	School Ownership 1—private school, 2—public school	1.86	0.345
		RATCMP2	Proportion of available computers that are connected to the Internet	0.97	0.095
		CREACTIV	Creative extra-curricular	2.54	0.767

		activities		
Teacher level	PROAT5AM	Index proportion of all teachers ISCED LEVEL 5A Master	0.14	0.131
	PROATCE	Index proportion of all teachers fully certified	0.96	0.120

2.3 Data Analysis

Given that the data set is structurally nested, i.e., the individual student level is nested within the school level, this study uses a multilayer linear model (HLM) to analyze the effects of student- and school-level variables on data literacy. While traditional regression analyses ignore the shared variance between levels of data, multilayer linear models based on ordinary least squares regression (OLS) can accurately estimate the slope of low-level data and its effect on high-level outcomes, and this method has been widely used in education, economics, and health care (Woltman et al., 2012). This analysis was performed using HLM 8.2 software to build a multilayer linear model, and the following three steps were performed separately according to the HLM user manual (Raudenbush, 2004).

2.3.1 Zero-model

To determine whether a multilayer linear model analysis was necessary for this dataset, model-1 called zero-model was created with data literacy as the dependent variable, which did not contain any independent variables. The zero-model is as follows:

Level-1 Model: $Data\_literacy_{ij} = \beta_{0j} + r_{ij}$

Level-2 Model: $\beta_{0j} = \gamma_{00} + u_{0j}$

where $Data\_literacy_{ij}$ represents the data literacy of student i, j stands for his or her school; γ_{00} represents the intercept of the model; and r_{ij}. u_{0j} represent the respective residuals at the two levels.

The intra-group correlation coefficient (ICC = 0.55718) of the model was calculated, indicating that 55.7% of the variance in student data literacy originated from school factors. The study showed that when ICC > 0.138 indicates a high degree of within-group correlation and between-group variance, i.e., students' data literacy varies significantly across schools, thus requiring a multi-layer linear model analysis.

Table 1. Estimation of variance components.

	Standard deviation	Variance component
between-group	50.74995	2575.55771
within-group	56.92717	3240.70239

2.3.2 Student level model

Based on the model-1, the nine independent variables of the student hierarchy, including the student cognitive dimension, the family background dimension, and the student non-cognitive dimension, were included in level-1 to build model-2, as follows:

Level-1 Model: $Data\_literacy_{ij} = \beta_{0j} + \beta_{1j}*(UNDREM_{ij}) + \beta_{2j}*(METASUM_{ij}) + \beta_{3j}*(METASPAM_{ij}) + \beta_{4j}*(ESCS_{ij}) + \beta_{5j}*(EMOSUPS_{ij}) + \beta_{6j}*(PERCOOP_{ij}) + \beta_{7j}*(GFOFAIL_{ij}) + \beta_{8j}*(RESILIEN_{ij}) + \beta_{9j}*(MASTGOAL_{ij}) + r_{ij}$

Level-2 Model: $\beta_{0j} = \gamma_{00} + u_{0j}$

……

$\beta_{8j} = \gamma_{80} + u_{8j}$

$\beta_{9j} = \gamma_{90} + u_{9j}$

where β_{aj} represent the fixed slopes of predictors at the student level.

2.3.3 Full model

Based on the model-2, 5 independent variables at the school level were included in level-2 to build model-3, as shown in the following, to explore the factors affecting students' data literacy. The γ_{0b} represent the fixed slopes of predictors at the school level.

Level-1 Model: $Data\_literacy_{ij} = \beta_{0j} + \beta_{1j}*(UNDREM_{ij}) + \beta_{2j}*(METASUM_{ij}) + \beta_{3j}*(METASPAM_{ij}) + \beta_{4j}*(ESCS_{ij}) + \beta_{5j}*(EMOSUPS_{ij}) + \beta_{6j}*(PERCOOP_{ij}) + \beta_{7j}*(GFOFAIL_{ij}) + \beta_{8j}*(RESILIEN_{ij}) + \beta_{9j}*(MASTGOAL_{ij}) + r_{ij}$

Level-2 Model: $\beta_{0j} = \gamma_{00} + \gamma_{01}*(SCHLTYPE_j) + \gamma_{02}*(RATCMP2_j) + \gamma_{03}*(CREACTIV_j) + \gamma_{04}*(PROAT5AM_j) + \gamma_{05}*(PROATCE_j) + u_{0j}$

$\beta_{1j} = \gamma_{10} + u_{1j}$

……

$\beta_{9j} = \gamma_{90} + u_{9j}$

3. Results

3.1 Impact of Student-level Variables on Data Literacy

In terms of cognitive dimensions of students, the correlation coefficient between comprehension and memory and data literacy was 3.96 ($p<0.001$), and the metacognitive strategies of comprehension and memory significantly and positively influenced students' data literacy levels, i.e., the stronger students' ability to use comprehension and memory strategies, the greater the contribution to their data literacy. The correlation coefficient between generalization ability and data literacy was 6.84 ($p<0.001$), and the metacognitive strategy of summarization also positively influenced data literacy, the more students were able to abstract general properties of things from concrete content, the higher their data literacy. Compared to the first two, assessment credibility had the highest correlation with data literacy with a correlation coefficient of 13.33 ($p<0.001$), i.e., when faced with uncertain information, the more aware students are to remain skeptical and figure out how to discern the accuracy of information, the more they are able to promote data literacy development. Thus, the cognitive dimensions of comprehension and memory, generalization and assessment credibility factors all had a positive effect on data literacy, with assessment credibility having the highest impact.

Table 2. Cognitive dimension.

Independent variable	Coefficient	Standard deviation	T-value	p-value
UNDREM	3.96	0.534	7.408	<0.001
METASUM	6.84	0.615	11.124	<0.001
METASPAM	13.33	0.587	22.711	<0.001

In terms of the non-cognitive dimensions of students, students' expressed willingness to cooperate was correlated with data literacy with a coefficient of 0.26, but not significantly. Similarly, students' fear of failure in learning was correlated with data literacy with a coefficient of -0.03 but was not a significant predictor. The correlation coefficient between students' learning resilience and data literacy was 3.66 ($p<0.001$), meaning that for each unit increase in the learner's ability to recover from a difficult setback, their data literacy increased by 3.66 points. Goal orientation also significantly and positively predicted data literacy, increasing it by 2.23 points for each unit of improvement. Thus, for the non-cognitive

dimensions, willingness to cooperate and fear of failure had no significant effect on data literacy, and learning resilience and goal orientation significantly and positively predicted data literacy.

Table 3. Non-cognitive dimension.

Independent variable	Coefficient	Standard deviation	T-value	p-value
PERCOOP	0.26	0.536	0.494	0.622
GFOFAIL	-0.03	0.565	-0.053	0.958
RESILIENCE	3.66	0.561	6.517	<0.001
MASTGOAL	2.23	0.589	3.780	<0.001

In terms of family background, the correlation coefficient between family economic and socio-cultural status index and data literacy was 4.24 ($p<0.001$), indicating that family hard and soft strengths can positively predict students' data literacy. Meanwhile, the correlation coefficient between parents' emotional support to students and data literacy was 2.59 ($p<0.001$), and although it was lower than the former, it was also able to positively influence the development of data literacy. Thus, both the family economic and socio-cultural status index of the family background dimension and parental emotional support had a highly significant positive effect on data literacy.

Table 4. Family dimension.

Independent variable	Coefficient	Standard deviation	T-value	p-value
ESCS	4.24	0.665	6.374	<0.001
EMOSUPS	2.59	0.572	4.529	<0.001

3.2 Impact of School-level Variables on Data Literacy

In the school context dimension, the correlation coefficient between school type and data literacy was -4.95, meaning that private schools were more able to promote the development of high data literacy than public schools, but this predictive effect was not significant. The correlation coefficient between number of computers per school and data literacy was 24.26, again with no significant effect. The correlation coefficient between innovative courses and data literacy was high and positively significant, with the higher the number of innovative courses offered by the school, the higher the data literacy of the students. In a word, school background dimension, school type and number of computers per capita have no significant effect on data literacy and innovative courses positively predict student data literacy.

Table 5. School background.

Independent variable	Coefficient	Standard deviation	T-value	p-value
SCHLTYPE	-4.95	3.573	-1.384	0.167
RATCMP2	24.26	26.137	0.928	0.354
CREACTIV	13.09	3.486	3.756	<0.001

For the teacher level dimension, the percentage of teachers with a master's degree was positively and significantly correlated with student data literacy ($c = 145.13$, $p < 0.001$), meaning that for every 10% increase in the percentage of teachers with a master's degree in the school, student data literacy performance increased by 14.51 points. The percentage of teacher credential holders also positively predicted student data literacy competency ($c = 121.69$, $p < 0.001$), with a 12.17-point increase in student data literacy performance for each 10% increase in the percentage of teachers holding a teaching

credential. Thus, both the master's degree and the percentage of credential holders in the teacher level dimension positively impacted student data literacy.

Table 6. Teacher level.

Independent variable	Coefficient	Standard deviation	T-value	p-value
PROAT5AM	145.13	25.527	5.685	<0.001
PROATCE	121.69	25.747	4.727	<0.001

4. Discussion and Conclusion

Based on PISA 2018 test data from four Chinese provinces and cities, 355 schools, and 11,190 students, this study analyzed the factors affecting data literacy, including student-level cognitive, non-cognitive, and family background factors and school-level school context and teacher level factors, using a multi-layer linear modeling approach based on a theoretical framework of educational productivity.

4.1 Student-level Variables on Data Literacy

Students' acquisition and agreement with metacognitive strategies for understanding and remembering, summarizing and generalizing, and assessing plausibility positively influenced data literacy. When confronted with data, students actively use a variety of knowledge to construct an understanding of the data to help them perform subsequent data manipulation, which is consistent with the study by Jiang et al. (2022) who found that students construct an understanding of data based on social and cultural context at the beginning of data modeling priming. At the same time, students with high levels of summarization and generalization skills are able to distill key points from complex data information and avoid the distraction of noisy data. In the face of a complex data society, it is essential to acquire cognitive strategies for assessing trustworthiness, as in the study of Van Wart et al. (2020), students were encouraged to apply critical perspectives to data literacy concepts and practices.

Students' resilience and goal orientation were significant predictors of data literacy, while a sense of collaboration and fear of failure were not. Data inquiry is a complex process in which learners are vulnerable to failure, so the ability to remain motivated to explore in the face of adversity can help them achieve ultimate success. Berhenke et al. (2011) reached a similar conclusion in his study, in which student persistence in learning was strongly associated with the development of mathematical ability. Data are created to pursue human goals and objectives, and students with a strong goal orientation are driven by intrinsic motivation to explore the information contained in the data (Hardy et al., 2020). The lack of a significant role for a sense of collaboration in data literacy seems to contradict existing research, although collaboration is the process of multiple people working together to achieve a goal, and data literacy emphasizes collaboration as more of a group atmosphere than just an individual sense of collaboration (Vance, 2021).

Families with high ESCS scores tend to have a variety of modern digital devices, as Lee & Dubovi (2020) found in their study of families with children with diabetes, where less well-off families used static blood glucose meters to measure and then record blood glucose data in notebooks, while middle-class or highly knowledgeable families used dynamic blood glucose meters and data analysis software to observe changes in the child over a certain period of time. Thus, students from families with different economic and socio-cultural levels experience differential informal data collection, management and storage in their family life, which affects their development of data literacy development. At the same time, research has shown that perceptions of parental support influence learners' learning resilience (Liu

et al., 2022), which in turn positively predicts data literacy, and thus parental emotional support can contribute to students' data literacy development.

4.2 School-level Variables on Data Literacy

The number of innovative course offerings in the school context had a positive effect on data literacy, while school type and number of computers per capita had no significant effect. The availability of innovative courses fosters creative thinking among students to come up with new insights into data problems with existing patterns of thinking and to use existing knowledge to improve or create new data models (Thompson & Arastoopour Irgens, 2022). The non-significant difference in data literacy between students by school type, i.e., public or private, implies that the issue of educational equity in data literacy education in China has not yet become apparent. The insignificant effect of the number of computers per capita on data literacy may be due to the fact that the current information technology curriculum at the secondary school level is still low in terms of data literacy and should be given more weight in the future to enhance international competitiveness in the data age.

Both teacher level of education and percentage of licensure significantly and positively affect data literacy competencies. Previous research has shown that teachers' education and credentialing necessarily effects student achievement (Kola & Sunday, 2015). Teachers who have undergone higher education at the master's level have improved their thinking and are more likely to accept and experiment with new teaching methods and ideas. Data literacy, a field that has only emerged in recent years, requires teachers to go beyond the content of classroom subject matter and create curricula that integrate course content and data literacy (Matuk et al., 2022) in order to bring students up to date.

5. Conclusion

Based on the educational productivity model, this study analyzed the PISA 2018 data from four Chinese provinces and cities using a multi-layer linear model to uncover the influencing mechanism of data literacy. The value of this study is to break through the traditional regression research method to further uncover the influencing factors of data literacy and provide some insights for future data literacy cultivation. However, there are some limitations in this study; the Chinese data of PISA2018 were only collected from developed provinces and cities, and the explanatory power of data literacy in the central and western regions needs to be further verified. Therefore, a larger study can be conducted in the future to explore the heterogeneity of factors affecting data literacy in developed and backward regions.

References

Alley-Young, G. (2017). White House Big Data Initiative. In *Encyclopedia of Big Data* (pp. 1–5). SpringerLink.

Berhenke, A., Miller, A. L., Brown, E., Seifer, R., & Dickstein, S. (2011). Observed emotional and behavioral indicators of motivation predict school readiness in Head Start graduates. *Early Childhood Research Quarterly*, *26*(4), 430–441. https://doi.org/10.1016/j.ecresq.2011.04.001

Hardy, L., Dixon, C., & Hsi, S. (2020). From Data Collectors to Data Producers: Shifting Students' Relationship to Data. *Journal of the Learning Sciences*, *29*(1), 104–126. https://doi.org/10.1080/10508406.2019.1678164

Huang, R., & Li, B. (2016). Data literacy education:expansion of information literacy instruction in the big data era. *Documentation,Information & Knowledge*, *01*, 21–29.

Jiang, S., Nocera, A., Tatar, C., Yoder, M. M., Chao, J., Wiedemann, K., Finzer, W., & Rosé, C. P. (2022). An empirical analysis of high school students' practices of modelling with unstructured data. *British Journal of Educational Technology*, *53*(5), 1114–1133. https://doi.org/10.1111/bjet.13253

Kahn, J. (2020). Learning at the intersection of self and society: The family geobiography as a context for data science education. *Journal of the Learning Sciences*, *29*(1), 57–80. https://doi.org/10.1080/10508406.2019.1693377

Kola, A. J., & Sunday, O. S. (2015). A review of teachers' qualifications and its implication on students' academic achievement in Nigerian schools. *International Journal of Educational Research and Information Science*, *2*(2), 10–15.

Kotu, V., & Deshpande, B. (2019). *Data science: Concepts and practice* (Second edition). Elsevier/Morgan Kaufmann Publishers.

Lee, V. R., & Dubovi, I. (2020). At Home With Data: Family Engagements With Data Involved in Type 1 Diabetes Management. *Journal of the Learning Sciences*, *29*(1), 11–31. https://doi.org/10.1080/10508406.2019.1666011

Liu, Y., Guan, X., & Gu, X. (2022). Shall we really know about perseverance?A malleable model from the viewing of teachers'based on grounded theory. *Global Education*, *51*(02), 39–58.

Matuk, C., DesPortes, K., Amato, A., Vacca, R., Silander, M., Woods, P. J., & Tes, M. (2022). Tensions and synergies in arts-integrated data literacy instruction: Reflections on four classroom implementations. *British Journal of Educational Technology*, *53*(5), 1159–1178. https://doi.org/10.1111/bjet.13257

Meng, X., & Li, A. (2014). Scientific data literacy education in overseas academic libraries. *Journal of Academic Libraries*, *32*(03), 11–16.

Raudenbush, S. W. (2004). *HLM 6: Hierarchical linear and nonlinear modeling*. Scientific Software International.

Thompson, J., & Arastoopour Irgens, G. (2022). Data Detectives: A Data Science Program for Middle Grade Learners. *Journal of Statistics and Data Science Education*, *30*(1), 29–38. https://doi.org/10.1080/26939169.2022.2034489

Vahey, P., Rafanan, K., Patton, C., Swan, K., van 't Hooft, M., Kratcoski, A., & Stanford, T. (2012). A cross-disciplinary approach to teaching data literacy and proportionality. *Educational Studies in Mathematics*, *81*(2), 179–205. https://doi.org/10.1007/s10649-012-9392-z

Van Wart, S., Lanouette, K., & Parikh, T. S. (2020). Scripts and Counterscripts in Community-Based Data Science: Participatory Digital Mapping and the Pursuit of a Third Space. *Journal of the Learning Sciences*, *29*(1), 127–153. https://doi.org/10.1080/10508406.2019.1693378

Vance, E. A. (2021). Using team-based learning to teach data science. *Journal of Statistics and Data Science Education*, *29*(3), 277–296. https://doi.org/10.1080/26939169.2021.1971587

Walberg, H. J. (1980). *A psychological theory of educational productivity.*

Woltman, H., Feldstain, A., MacKay, J. C., & Rocchi, M. (2012). An introduction to hierarchical linear modeling. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 52–69.

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

Examining Social Presence and Students' Motivational Beliefs in Large-Scale Online Learning: An Exploratory Approach

Wen-Li Chang[1*], Jerry Chih-Yuan Sun[2]

[1.2]Institute of Education, National Yang Ming Chiao Tung University, Hsinchu, Taiwan

[*] wendychangnycu@nycu.edu.tw

Abstract: *The present study is intended to explore students' motivational changes, under the correlational influence of socializers' beliefs and behavior perceived. Expectancy-value theory provides a framework for identifying and examining the proposed serial influencers in the developmental process (i.e., self-perception of others that affects self-schema, expectation of success and task value, as well as academic choice). The participants are college students enrolled in a large-scale online course and therefore engaged in content-based language learning tasks that collect their reflective feedback. Content analysis is performed on the collected data for deepened understanding of the phenomenon. Tentative findings lend support to the intersecting factors, both psychological and environmental, as well as to their role of necessary building blocks for a motivational learning environment.*

Keywords: motivational belief, environmental determinant, socially-driven self-schema, large-scale online learning, expectancy-value theory

1. Introduction

Learning in an open society type of digital environment is expected to introduce beneficial aspects to language learners' exploration of knowledge and skills. Relevant studies have lent strong support to the positive tie between the so-called highly socially oriented environment and foreign language students' learning achievements (e.g., Mondahl & Razmerita, 2014). Examining students' own perception, considerable studies have additionally strengthened the substantial benefits accruing toward their attitudes and beliefs when the social element of a web-based language learning context is highly involved (e.g., Kitchakarn, 2016). A closer look at students' individual thinking and learning process, however, is comparatively rare as opposed to a heightened research focus on their learning outcomes in a web-based social context, examined either with or without student perspectives.

An open online course at a school-wide massive scale provides an essential ground and is therefore adopted for further exploration of foreign language (FL) students' motivation for learning, as defined by Clark et al. (2006). In such a complex learning environment that requires adequate prior knowledge and cognitive operations, it is assumed that students' motivation becomes a significant predictor of learning. The expectancy-value model of motivation is drawn on for a theoretical image that illustrates the hypothesized interrelationships of the various factors and determinants. The present study continues this thread of discussion with a special focus on the interactions between environmental factors (a social context) and psychological factors (personal perception of others and oneself), as well as on the resulting impact brought upon their motivational beliefs (values and expectations), and correlationally, on the ultimate academic choice. Specific research questions to be addressed are as follows.

(1) How does the presence of class socializers (peers) affect online students' perception of the social context?
(2) What categories of self-schema (perceived ability in relation to peer performance and to self-performance in other domains) do the students reflect on in their peer-directed reflective comments?
(3) How does the perception level of social (peer) presence correlate with the students' reflection on self-schema and motivational beliefs (expectancy and task values)?
(4) How does different social (peer) perception levels affect the students' academic choices to persist in (language) learning?

2. Literature Review

2.1. Expectancy-Value Theory (EVT)

Modern EVTs, evolving from Atkinson's (1964) expectancy-value model, strengthen the research support for individual motivational development that interacts with the social context. This clearly illustrates the distinct efforts prevailing among contemporary perspectives on motivation in education, as pointed out by Schunk et al. (2014), that motivation, being an extremely complex phenomenon, is determined by a multitude of personal, social, and contextual factors, and that at an individual level, one's motivation involves cognitions, that is, a person's mental action and process of understanding through thoughts, beliefs, goals, and self-representations. Most importantly, motivation is not an interchangeable term with achievement outcomes, the definition of which, though, can be as broad as to include learning, performance, and self-regulation, but future motivation is considerably strengthened by these outcomes, in that students' notion of personal learning progress guarantees their being motivated to continue. Table 1 summarizes the fundamental building constructs of modern EVTs.

Table 1. Modern EVT constructs explained (adapted from Eccles & Wigfield, 2002).

Expectancy (E) construct: Beliefs about ability as predictor of performance			
Key theorist	Definition	Core value	Similar terms
Eccles et al. (1983): Expectancies for success	Individuals' broad beliefs about how well they will do on upcoming tasks	Focus on task, outcome expectations	A different level of beliefs in ability, i.e., personal competence in a given domain
Bandura (1997)	Personal efficacy expectations	Focus on self, personal efficacy expectations	
Value (V) construct: Beliefs about benefits as predictor of involvement decision			
Key theorist	Definition		Similar terms
Eccles et al. (1983): Task-value	Four components of (1) attainment value (personal importance of doing well on the task), (2) intrinsic value (enjoyment from doing the task or subjective interest in the subject), (3) utility value (relevance of a task to current or future goals), and (4) cost (negative aspects of engaging in the task)		Intrinsic value, intrinsic motivation, interest and flow; utility value and extrinsic motivation

For the present study, the underlying theoretical framework, as shown in Figure 1, is adapted from Eccles and Wigfield (2002)'s expectancy-value theory, highlighting one of the chains that contain the following items: (a) Socializers' beliefs and behaviors, (b) perception of socializers' beliefs and behaviors, (c) general self-schema, (d) expectation of success and subjective task value, and (e) achievement-related choices. Within a targeted intra-individual process of motivation development in a language learning context, it is assumed that under the direct influence of expectancies and values are students' performance, persistence, and task choice (precisely specified in the present study for intentions to continue learning). The theorists propose that expectancies and values are assumed to fall under the influence of individuals' goals and self-schema, among others. These internal and psychological variables, during the same motivational development, are influenced by individuals' perceptions of other people's attitudes and expectations for them (Eccles & Wigfield, 2002).

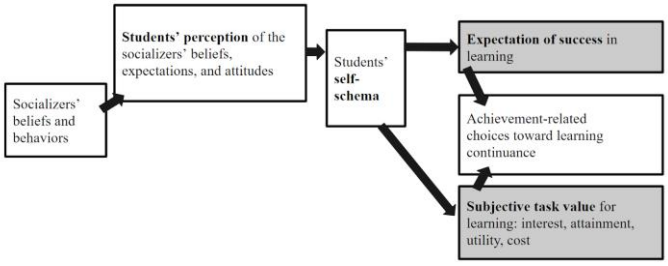


Figure 1. Theoretical framework (adapted from Eccles & Wigfield, 2002).

2.2. Socially-Driven Self-Schema as an Antecedent of Expectancy and Value Beliefs

In an educational context highlighted by Eccles et al. (1983, 2002), an individual's developmental process of motivational beliefs that determine their achievement-related choice is, in itself, under the influence of both environmental and psychological determinants. It is within a given sociocultural context that an individual forms goal-related or task-specific perceptions and interpretations, namely, perceptions of competence, perceptions of the difficulty of different tasks, individual's goals and self-schema. Such social cognitive variables, in other words, are never to be overlooked in terms of their close ties with the environment where socializers (instructors and peers) are involved. In consistency with the theoretical assumptions is Wigfield's (1994) explanation of the expectancy-value model, highlighting that an individual's goals and task specific beliefs are influenced by one's interpretations of prior performance and perceptions of significant socializer's attitudes and expectations. One's self-schema is defined as competence beliefs, as reflected in a self-perceived level (e.g., *How good at math are you?*) and in relation to the whole class (namely, other peer socializers) (e.g., *If you were to order all the students in your math class from the worst to the best in math, where would you put yourself?*).

Wigfield et al. (2009) stressed the distinction between expectancies for success and ability beliefs; while the former (expectancies for success) is defined as students' beliefs about how well they will do on upcoming tasks, regardless of the time span (either immediate or longer term future), the latter (ability beliefs) is defined as student's perceptions of their individual competence level demonstrated at the current time point at a given activity. In other words, the scholars assumed that ability beliefs should be conceptually differently understood from expectancies for success, with ability beliefs focused on present ability and expectancies focused on the future. By this clear distinction, an additional aspect in measuring ability beliefs is introduced in their study as a relational scale, as having been applied to the pre-existing aspects, and further includes perception of overall academic performance, *e.g., Some kids are better in one subject than in another. For example, you might be better in math than in reading. Compared to most of your other school subjects, how good are you in math? (a lot worse in math than in other subjects - a lot better in math than in other subjects.* Upon such a broadened scope of one's ability and competence beliefs, the social cognitive role is still heightened in relation to the subsequent motivational beliefs generated (expectancy and value), as well as to the antecedent personal previous experiences and a variety of socialization influences.

Previous studies focused on self-schema or perceived ability in particular have shown its influence on students' achievement-related choices through the preceding impacts on expectancy and value beliefs. By comparing two educational stages, Wigfield et al. (2009) illustrated the developmental change in children's and adolescents' ability beliefs, expectancies for success, and subjective values, the manifestation of which bears a strong relation to their task performance and academic choice. The scholars further noted that at least through early adolescence, the level of negativity in ability-related beliefs and values grows higher with age: students at the later educational stage show decreasing beliefs in their task competence and place comparatively low values on certain tasks. The negative changes in achievement-related beliefs and values are explained in association with the co-occurring socialization process in which evaluative feedback is better understood, and social comparison with peers has been made more salient in school environments and therefore becomes more likely. Regarding the relations of competence beliefs and values to academic performance and choice, the scholars concluded that a strong prediction is confirmed from students' ability beliefs and expectancies for success to subsequent score performance, and from subjective task values to their motivated learning behavior or to continuance intentions towards learning.

The afore-mentioned scholarly contributions are consistent with Eccles et al.'s (1983) proposed connection between self-schema and task values that occurs in primary school environments where learning tasks are assigned and opportunities arise, along with the increasing attainment values in particular, for students engaged to confirm or

disconfirm noticeable aspects of their self-beliefs (actual or ideal) in specific domain competence and even in broadened generalizations of self-image. Relevant studies have continued the research effort by further exploring advanced educational stages that include high school and college levels, among which a heightened focus is shared and placed on the language learning context, as well as on the predicted motivational development that influences their persistence and achievement in the process, as summarized in Table 2. The development of individual students' expectancy and value beliefs is closely associated with their socially-driven self-schema, therefore being redefined and positioned as not only a crucial antecedent in the process but an effective motivation predictor. In the context of second language (L2) learning, self-beliefs are commonly included as one of the predictors of students' motivated language learning behavior, desired language achievement, and/or their continued learning intention, despite the broadened scope of practice entailing self-confirmation of their possession (or non-possession) of ideal attributes that range widely over different facets of an L2-specific self (e.g., Lamb, 2012), a metacognitive self (e.g., Xu, 2017), and a general academic self (e.g., Nagle, 2021), evolving at the same time, from a secondary educational level to a higher education context.

Table 2. Relevant studies on predicting motivational development in school-based language learning contexts.

EVT framework	Student background	Level and target language	Motivation predictor	Learning behavior motivated	Target proficiency achieved	Learning intention continued
Nagle (2021)	USA	College Spanish	Learning experience	+	−	−
			Expectancy of success, academic self-concept	−	+	−
			Task values	+	−	+
Xu (2017)	China	College English	Metacognitive self		(As mediator)	
			Anxiety	−	+	−
			Expectancy	−	+	−
			Interest	−	+	−
Lamb (2012): Motivational Self System	Indonesia	Junior high English	Ideal L2 self	−	−	−
			Learning experience	+	+	−

2.3. Social (and Environmental) Determinants in Motivational Beliefs and Achievement Choices

In an attempt to broaden the conceptualizations of achievement values, Wigfield and Eccles (1992) emphasized the various antecedents in the developmental process that correspond to the different task characteristics, as well as to each individual's dynamic needs and values growing under socialization influences. This continued the research effort of Eccles et al. (1983), who proposed four primary antecedents of young learners' achievement values for varied tasks: self-schema and goal, relative task cost or benefit perceived at one task over others, prior task performing experience, and their perception of beliefs and values of their parents, teachers, and peers. Further attempts are made, as Wigfield et al. (2009) highlighted, in relation to later models, including Eccles and Wigfield's (2002) EVT model, for a richer definition of the expectancy and value components, and for a wider range of psychological, social, and cultural determinants that are hypothesized for and tested in real-world achievement situations rather than laboratory settings.

From a widened perspective towards the environmental determinants, a school-based educational context has necessitated relevant studies on the socialization of achievement choices by moving beyond parental aspirations and introducing the influence of socializers at school, namely, instructors and peers. Extending Wigfield and Eccles' (1992) suggestion for examining teacher influence on young learners' values that are attached to different school subjects but change over time, Eccles and Wigfield (2020) systematically reviewed relevant research attempts regarding the role of schools in situated EVT. The general facilitating and supportive role of school context is highlighted, with focus on the positive impacts of classroom aspects and characteristics on students' self-perceived confidence in their academic

ability and subjective task values, namely, their feelings of competence, connectedness, and autonomy, as well as on their enhanced learning engagement (e.g. Eccles, 2012; Wang & Eccles, 2012; Wigfield et al., 2015). It is further noted that students' changing school/classroom experiences are also highly associated with both declines and increases in their academic self-schema, as well as their expectancy and value beliefs across the school years, though primarily set in U.S. schools (e.g., Eccles & Roeser, 2010, Wigfield et al., 2012).

The refined, situated EVT is justified by Eccles and Wigfield (2020) with heightened focus on school contexts, meanwhile bridging a common gap in utilizing the theoretical model to the full extent. The middle part of the model is often neglected, which leads to comparatively rare research focus on socialization influences, namely, how individual students perceive class socializers' beliefs, expectations, and attitudes toward their participation and performance, thereby affecting how they understand and interpret their own competence and proficiency, as opposed to devoted attention to the relation of students' expectancy-related and task-value beliefs to their achievement performance, persistence, and academic choice. In response to the scholars' call for research attempts to examine peers as socializers in a learning context and, in the meantime, to deepen understanding of the crucial role of EVT in L2 learning, as suggested by Loh (2019), the present study is intended to delve into the niche by exploring a web-based language learning context where peer-based sources of self-perceptions and interpretations are collected and analyzed for their influences on subsequent development of motivational beliefs and achievement choices.

3. Methods and Procedure

The research questions are answered mainly based on a content analysis approach to peer-directed written comments that are collected from students (a convenient sample from a large-scale online class) (N = 293). A triangulating attempt is made by referring to the students' end-of-the-semester personal reflection. Figure 2 summarizes the sequential research steps of this study.

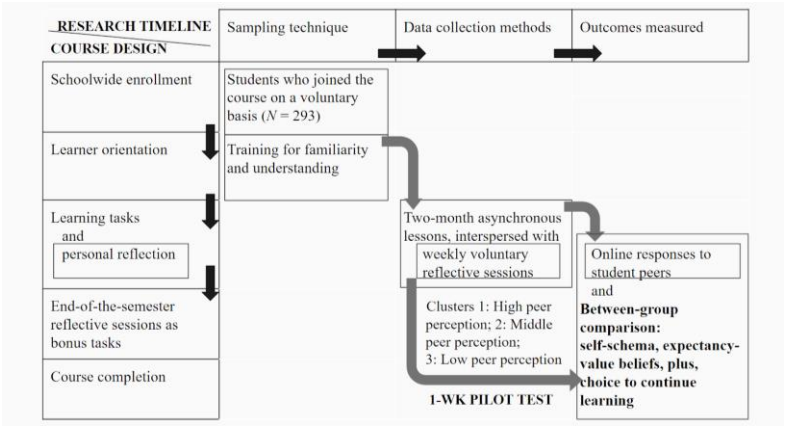


Figure 2. Experimental procedure.

All lessons are provided asynchronously using Google Classroom, a web-based online learning platform. The students participate in this fully online course at their own learning pace, which allows them to complete each learning task even after a scheduled due date. The students' learning path is meanwhile being traced and recorded in terms of their contributed output, participation level, and general completion rate. Within the overall two-month semester, one thematic learning task is assigned to the whole class on a weekly basis. Namely, the students are presented with a new learning topic every week, on which relevant video contents, intended to elicit student feedback, are announced on the class platform that integrates a YouTube-based learning website (BEST BANANA) and a Google Play-supported learning application (Notstupid), as illustrated in Figure 3. The website provides modified input to the students, primarily in the form of automatically generated captions in the same English language as videos, whereas the mobile

application extends the assigned learning tasks by capturing difficult vocabulary items for the students' personal review purposes.

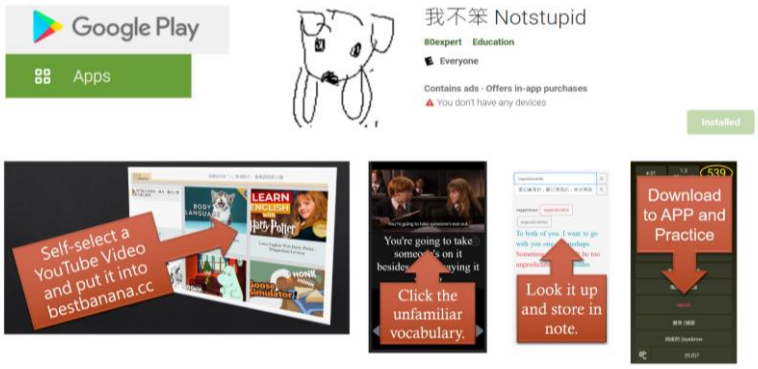


Figure 3. Snapshot of digital support for the class.

This study is focused on the website-enhanced video learning context where an extended learning opportunity is presented as a bonus task. The students, on a voluntary basis, respond to their classmates' postings on learning topics from the previous weeks by writing minimum one brief reflective comment in Mandarin Chinese, the students' mother tongue language. The written comments are collected by the class platform in a specifically designated session that requires the volunteer students to submit their self-captured snapshots of their own comments (namely, responses to others) that could come under different thematic learning tasks from early class activities. (See Figure 4 for an example.) A cluster analysis approach is adopted for classifying the collected comments (the student commenters) according to the perceived presence of their peer socializers (high/middle/low peer perception). Following the underlying EVT theoretical framework, the present pilot study is supported by a subsequent content analysis approach that explores the students' written comments for further identifying the distinctively reflected categories and interacting relationships of self-schema, their expectancy and value beliefs, and their ultimate academic choices to continue learning the target language. Inter-coder reliability is enhanced by engaging two experienced language teaching experts in coder training prior to data analysis. At the stage of coding and analyzing data, the experts/coders refer to a predetermined coding scheme developed according to Wigfield & Eccles (2000), who proposed major items for consideration: ability beliefs items, expectancy items, and usefulness, importance, and interest items.

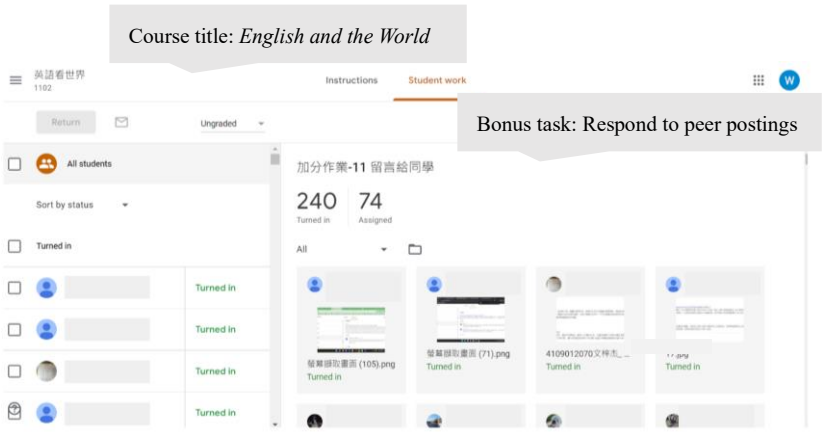


Figure 4. Snapshot of class webpage.

4. Tentative Findings

Regarding online students' perception of the social context under the influence of peers as class soilalizers (Research Question 1), cluster analysis is performed for grouping the students' peer-directed reflective comment inputs according to the overall frequency of their referring to peers in words (e.g., nǐ or nǐ/-men/you, tóng-

xué/classmate, yī-qǐ/joined). The students are thus expected to form groups of different peer perception levels. Figures 5 and 6 are sample source texts for clustering: lower and higher peer perception, respectively.

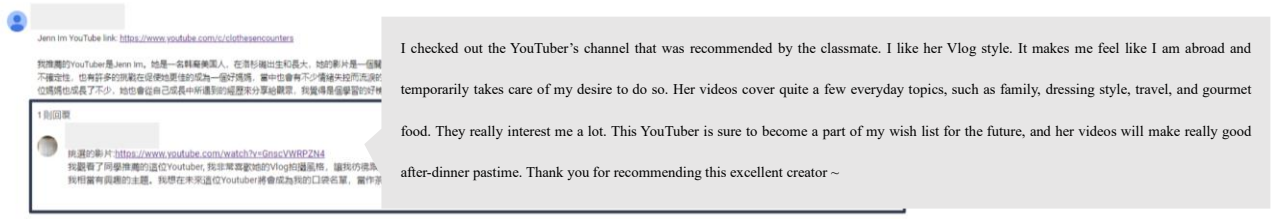


Figure 5. Source text for cluster analysis: Higher peer perception.

Figure 6. Source text for cluster analysis: Lower peer perception.

The clustered student comments are later coded against a predetermined but tentative scheme for further categorization into different types of self-schema represented: perceived ability, both present and future, in relation to peer performance and to self-performance in other domains (Research Question 2). Figure 7 provides a sample collection of input words and phrases (e.g., xīn-shǒu/newbie) from a student with lower peer perception, namely, the lack of direct reference to the class peers. Comparison with pairing results between self-reference (e.g., wǒ/I, wǒ-de/my) and perceived ability from the higher peer perception cluster is expected to identify peer influence on self-representations that are generated along their socially-oriented mental processing.

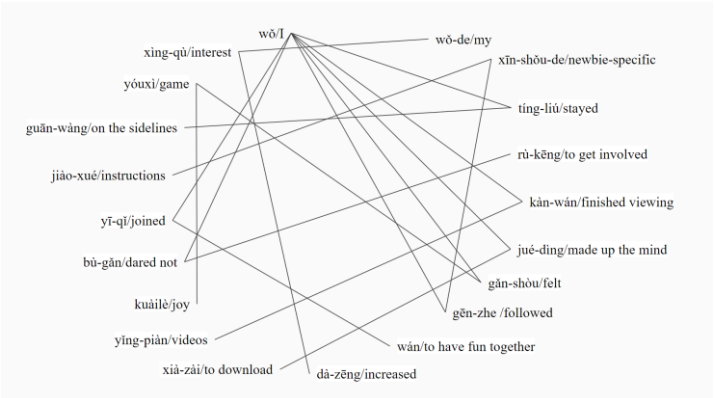


Figure 7. Pair of co-occurring words drawn from a source text from lower peer perception.

As suggested in the underlying EVT model, individual students' self-schema (perceived ability) influences their motivational beliefs. The observed self-representations in words (e.g., xīn-shǒu-de/newbie-specific), from both lower and higher peer perception clusters, are therefore expected to introduce the connected expectancy and/or value components (e.g., gēn-zhe xīn-shǒu-de jiào-xué /follow newbie-specific instructions). This content-based relational analysis of immediate word choices following personal attempts to identify and represent oneself is to lend support to the strong ties and interactions between the major EVT constructs, namely, both the environmental and psychological factors (Research Question 3).

The end-of-the-semester, whole-class reflection is also to be coded for detailed descriptions of their future actions, as reported by the students themselves in different clusters: lower and higher peer perception. Whether the students persist in their academic choices about language learning (Research Question 4) is to be examined as a necessary revisit to the underlying EVT model, as illustrated in Figure 8, for possible refinements that may move beyond predicted improvements in English proficiency and include as broadly as motivated video-based or mobile learning behavior and continued learning purposes for additional language, for advanced online learner support, and for new friends in a web-based social context.

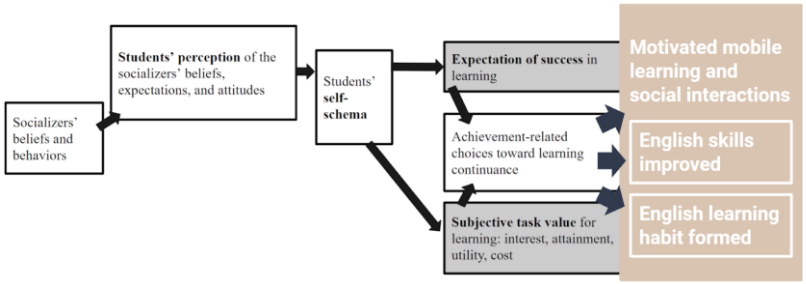


Figure 8. Theoretical framework revisited.

5. Conclusion

A closer look at individual students' expectancy-value development is for understanding and enhancing the building blocks of motivational design and, in a broader sense, the overall instructional and learning environment design, as emphasized by Keller (2009). In the mean-time when the present study attempts to foster the existing interactions between environmental and psychological factors in students' learning motivation and therefore to encourage massive pedagogical implications in an eLearning context, future directions are provided to facilitate thorough practice of the situated EVT framework. The socialization influences of school environments are sure to be worth further exploration at an equally large scale where media channels and sociocultural determinants are under careful consideration.

Acknowledgements

My greatest appreciation is dedicated to all participants in this research project, namely, the research team members from NYCU Interactive Learning Technology and Motivation (ILTM), ROC Taiwan.

References

Atkinson, J.W. (1964). An Introduction to Motivation. Princeton, N.J.: Van Nostrand.

Bandura, A. (1997). Self-efficacy: The exercise of control. New York: Freeman

Clark, R. E., Howard, K., & Early, S. (2006). Motivational challenges experienced in highly complex learning environments. Handling Complexity in Learning Environments: Theory and Research, 27-43.

Eccles, J. S. (2012). Supporting America's children and adolescents. Macalester International, 29, 1-23.

Eccles, J., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., et al. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), Achievement and achievement motivation (pp. 75-146). San Francisco, CA: W. H. Freeman.

Eccles, J. S. & Roeser, R. W. (2010). An ecological view of schools and development. In J. L Meece and J. S. Eccles (Eds.), Handbook of Research on Schools, Schooling, and Human Development, (pp. 6-21). New York: Routledge.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. Annual Review of Psychology, 53(1), 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153.

Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. Contemporary Educational Psychology, 61, 101859.

Keller, J. M. (2009). Motivational design for learning and performance: The ARCS model approach. Springer.

Kitchakarn, O. (2016). How students perceived social media as a learning tool in enhancing their language learning performance. Turkish Online Journal of Educational Technology-TOJET, 15(4), 53-60.

Lamb, M. (2012). A self system perspective on young adolescents' motivation to learn English in urban and rural settings. Language learning, 62(4), 997-1023.

Loh, E. K. (2019). What we know about expectancy-value theory, and how it helps to design a sustained motivating learning environment. System, 86, 102-119.

Mondahl, M., & Razmerita, L. (2014). Social media, collaboration and social learning a case-study of Foreign language learning. Electronic Journal of E-learning, 12(4), pp. 339-352.

Nagle, C. (2021). Using Expectancy Value Theory to understand motivation, persistence, and achievement in university-level foreign language learning. Foreign Language Annals, 54, 1238–1256. https://doi.org/10.1111/flan.12569

Schunk, D. H., Meece, J. L., & Pintrich, P. R. (2014). Motivation in education: Theory, research, and applications (4th ed.). Pearson Education.

Wang, M. T., & Eccles, J. S. (2012). Adolescent behavioral, emotional, and cognitive engagement trajectories in school and their differential relations to educational success. Journal of Research on Adolescence, 22(1), 31-39.

Wigfield, A., & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. Developmental Review, 12(3), 265-310.

Wigfield, A.. (1994). Expectancy-value theory of achievement motivation: A developmental perspective. Educational Psychology Review, 6(1), 49–78. https://doi.org/10.1007/bf02209024

Wigfield, A., Cambria, J., & Eccles, J. S. (2012). Motivation in education. In R. C. Ryan (Ed.), Oxford handbook of motivation (pp.463-478). New York: Oxford University Press.

Wigfield, A, Eccles, J. S., Fredricks, J., Simpkins, Roeser, R., & Schiefele, U. (2015). Development of achievement, motivation and engagement. In R. Lerner (Series ed.) and M. Lamb (Vol. Ed.), Handbook of child psychology and developmental science (7th ed., vol. 3, pp. 657-700). New York: Wiley.

Wigfield, A., Tonks, S., & Klauda, S. L. (2009). Expectancy-Value Theory. In Handbook of Motivation at School (pp. 69-90). Routledge.

Xu, J. (2017). The mediating effect of listening metacognitive awareness between test-taking motivation and listening test score: An expectancy-value theory approach. Frontiers in psychology, 8, 2201.

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

Mining and Visualizing Teachers' Knowledge in Online Discussion Based on Topic Model and Epistemic Network Analysis

Si Zhang[1], Tongyu Guo[2*], Yuting Zhang[3], Xinyue He[4], Yiyao Wang[5]
[12345] Faculty of Artificial Intelligence in Education, Central China Normal University
[*]1550488644@qq.com

Abstract: *By using the topic model, this study mined 13 topics implicit in the data of teachers' online discussion and then labeled these topics according to the 7 factors of the Technological Pedagogical and Content Knowledge framework, so as to determine the types of teachers' knowledge in online discussion. This study further applied the Epistemic Network Analysis (ENA) to explore the characteristics of teachers' knowledge network. The results showed that teachers paid attention to both the teaching and learning of students. The epistemic network characteristics of teachers in different score groups and at different ages were different.*

Keywords: learning communities, computer-mediated communication, evaluation methodologies, interactive learning environments

1. Introduction

In order to improve the quality of online training, education researchers often collect teachers' online discussion data, and then use content analysis, statistical analysis, social network analysis, and qualitative case study to analyze these data to understand behavior, cognition and social network characteristics of teachers in the online learning community (Burhan-Horasanlı & Ortaçtepe, 2016; Lee & Brett, 2015; Liu, 2012; Rienties & Kinchin, 2014). Although these studies are helpful to propose the online training strategy, there are some problems. On the one hand, these research methods are time consuming and laborious, and only a small volume of data can be analyzed. On the other hand, conclusions obtained through these data analysis methods may be subject to the subjective factors of educational researchers. Therefore, it is necessary to adopt data-driven method and integrate education data mining and learning analysis technologies to realize automatic analysis and evaluation the characteristics of teachers' knowledge in the OPLC.

The purposes of this study are 1) to explore the types and distribution of teachers' knowledge by using the Latent Dirichlet Allocation (LDA) topic model; and 2) to explore the epistemic network characteristics of teachers with different scores and ages by using Epistemic Network Analysis (ENA). Mining teachers' online discussion data to understand their knowledge types and network characteristics, mainly because: (1) Participation in online discussion is the main way for the professional development of teachers in the OPLC. (2) LDA is an unsupervised machine learning method, which is very suitable for mining topics from large volumes of text data. Based on the identified topics, ENA can be used to explore the relationship between teachers' knowledge types and the development characteristics. (3) On the basis of understanding teachers' knowledge types and development characteristics, education researchers and teacher training managers can design more targeted teacher training content and training strategies to improve the quality of teacher professional development.

2. Literature review

2.1. Teachers' online professional development and knowledge

There are many researches on teachers' knowledge, among which the most typical one is the framework of Technological Pedagogical Content Knowledge (TPACK). It is generally believed that the TPACK framework consists of seven factors, which are the Technological Knowledge (TK), the Pedagogical Knowledge (PK), the Content Knowledge (CK), the Technological Pedagogical Knowledge (TPK), the Technological Content Knowledge (TCK), the Pedagogical Content Knowledge (PCK), and the TPACK. Referring to Cox and Graham (2009)'s definition, this study operationally defined the seven factors of TPACK framework in teachers' online discussion data and the topics mined by LDA were compared with the seven factors of TPACK to determine the types and distribution of teachers' knowledge in online discussion.

2.2. Learning analytics

Learning analysis and evaluation is an important guarantee for students' learning performance in education. Common methods used in learning analysis include statistical analysis, content analysis, and social network analysis.

Epistemic network analysis (ENA) takes epistemic elements as nodes and co-occurrence of epistemic elements in a time window as a link, so as to explore the epistemic network characteristics of individual or collectivity. There are three core concepts in the ENA: Unit of analysis, conversation, and codes. Unit of analysis is the object for which the researchers want to understand the interactions between the codes. Conversation is the units in which the researchers measure the codes co-occurrence. The codes are a set of concepts and the interactions between these concepts are the researchers want to understand. The data encoded by ENA can be Boolean data (0 or 1), integer data, and fractional number, where the fractional number can be used for LDA topic model. Typical applications of ENA include cognitive games and assessment of engineering design thinking (Arastoopour, Shaffer, Swiecki, Ruis, & Chesler, 2016; Hatfield, 2015).

2.3. Educational data mining

Educational data mining (EDM) refers to the "development, research, and application of computerized methods to detect patterns in large collections of educational data that would otherwise be hard or impossible to analyze due to the enormous volume of data" (Romero & Ventura, 2013).

As an unsupervised machine learning method, LDA is very suitable for the analysis of large volume text data. By manually labeling the topics of LDA mining, deep information of teachers' online discussion can be analyzed. Further combined with the ENA, the development characteristics of teachers' knowledge can be explored.

Based on the literature review, the research problems proposed in this paper are as follows:

(1) What are the overall characteristics of teachers' knowledge in online discussion?

(2) What are the types and evolutionary characteristics of teachers' knowledge in online discussion?

(3) What are the differences between the epistemic network characteristics of teachers in high-score and low-score groups?

(4) What are the differences between the epistemic network characteristics of teachers in different age groups?

3. Methodology

3.1.Research design

This study collected teachers' discussion data in the OPLC and used the LDA topic model to explore topics and the ENA to explore development characteristics of teachers' knowledge. The research process includes five main phases.

Phase 1: This study used a web crawler called the Octopus to collect online discussion data of teachers in an OPLC.

Phase 2: This study preprocessed the text data stored in the database, including word segmentation, removing stop words, calculating the TF-IDF (Term Frequency – Inverse Document Frequency).

Phase 3: This study used the LDA to explore the hidden topics of the text data and got the topics and words distributions.

Phase 4: 500 comments were randomly selected from the text data set for manual label.

Phase 5: After the comments were classified, they were organized according to the data format of ENA.

3.2. Participants

The OPLC in this study was established on a teachers' online professional learning platform (http://guopei.crtvu.edu.cn/cms/). The criteria for selecting the OPLC were: 1) it was an OPLC for teachers of main subject, 2) teachers in this OPLC had completed the learning task as planned, and 3) the demographic characteristics of teachers in this OPLC were close to the population of training teachers.

A math teachers' OPLC was finally selected as the object of this study. Teachers in this OPLC first watched an online video case. In this video case, a teacher organized games and interactive teaching activities to present the topic-specific subject content by using the special functions in the electronic whiteboard, such as searchlight and curtain. Then, all the teachers in this OPLC online discussed how to design, implement and evaluate the application of information technology, especially the special functions of electronic whiteboard, in the teaching of mathematics. The OPLC consisted of 91 teachers, among whom 42 were male teachers (46.2%).

3.3. Online discussion activity

Participation in online asynchronous discussion is the main task of teachers in the OPLC. The workflow of the online discussion activity is shown in Figure 1.

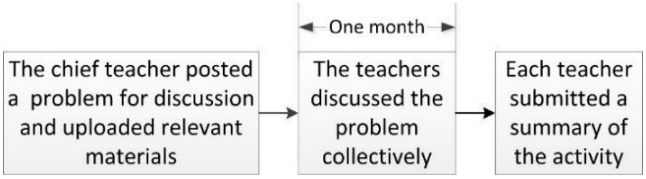


Figure 1. The workflow of the online discussion activity

3.4. Data collection and preprocess

This study used a web crawler tool, the Octopus, to collect discussion data from OPLC, and a total of 1,960 comments were collected. After the preliminary data processing, the remaining number of comments was 1,872. Teachers' demographic information (name, age, etc.) and training scores were also collected.

After the comments was sorted out, the "jiebaR" word segmentation tool of R software package was used to pre-process the text, including Chinese word segmentation, deleting stop-words and keyword extraction.

3.5. LDA topic model

Firstly, we need to determine the optimal number of topics before applying the LDA. 1,872 comments were used as input data, and the package “ldatuning” in R language was used to determine the optimal number of topics. The number of topics were set to a sequence {0,1,2..., 50}, and then the FindTopicsNumber function with the "Griffiths2004" as the judgment standard and the "VEM" as the LDA method was run. The result was shown in Figure 2. As the number of topics changed, the function peaked at the position of "13". Therefore, "13" was determined to be the optimal number of topics for LDA model.

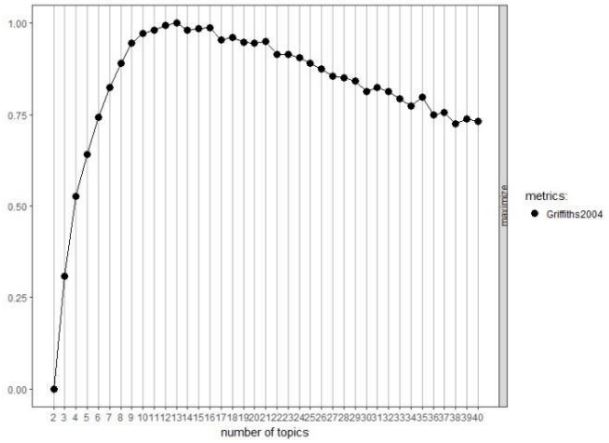


Figure 2. The optimal number of topics

After the optimal number of topics was determined, all the 1,872 comments were preprocessed by using the "slam" package in R language, and then the "topicmodels" package was used for topic mining to explore the topics-words information implied in the 1,872 comments. As an unsupervised machine learning method, LDA automatically determined the topic of each comment.

3.6. Topic labeling

When the LDA was finished running, each comment was assigned a topic. The next step was to determine the type of knowledge that each topic corresponded to. 500 comments were randomly selected from 1,872 comments for manual label. Two researchers familiar with both teaching contents and methods of mathematics courses and definition of seven factors in the TPACK framework encoded the 500 comments. The intercoder reliability coefficient was calculated and the value was 0.81 (Cohen’s Kappa), which showed a good reliability (Fleiss, 1981).

3.7. Epistemic network analysis

After all the comments were labeled according to the TPACK factors, then these comments were sorted according to ENA's data format. The "rENA" package in R language was used to explore the epistemic network characteristics of teachers in high- and low- scores groups and different age groups.

This study used the seven factors of TPACK framework as the coding unit, the serial number of

comments as the conversion window (the window's size was 4), and teachers' training score as the unit of analysis, to explore the epistemic network characteristics of teachers in high- and low- scores groups. Then, using the age groups as the analysis units, this study explored the epistemic network characteristics of teachers in different age groups. Among them, the value 2 represented the teachers aged 20 to 29, the value 3 represents the teachers aged 30 to 39, the value 4 represents the teachers aged 40 to 49, and the value 5 represents the teachers aged 50 to 59.

4. Results

4.1. What are the overall characteristics of teachers' knowledge in online discussion?

The analysis of high frequency phrases was mainly used to explore the overall characteristics of teachers' knowledge. Teachers' comment contained a total of 3,241 key phrases. After sorting these key phrases, the 9 phrases with the highest number of occurrences were listed, as shown in Table 1. In the teachers' comments, the phrase that appeared most frequently was "Student". The frequencies of the phrases "Teaching" and "Learning" ranked second and fourth, indicating that teachers not only cared about classroom teaching, but also students' learning. In addition, the frequency of the phrase "Information technology" ranked the third, indicating that teachers thought actively about the applications of information technology in mathematics teaching.

Table 1. The most frequent words appeared in teachers' online discussion

No.	Words	Freq	No.	Words	Freq	No.	Words	Freq
1	Student	3338	4	Learning	1583	7	Knowledge	573
2	Teaching	2056	5	Math	753	8	Class	567
3	Information technology	1674	6	Teacher	633	9	Multimedia	548

4.2. What are the types and evolutionary characteristics of teachers' knowledge in online discussion?

13 topics appeared in different proportions in teachers' online discussion. Three topics with the highest proportion and words under these topics were shown in Table 2. In topic 1, this topic was related to the pedagogical knowledge, and teachers paid attention to the application of teaching methods such as " Inquiry " and " Group collaboration ". Topic 5, 6 were all related to the pedagogical content knowledge. In topic 5, teachers paid attention to the way of teaching and learning of mathematical knowledge. In Topic 6, teachers paid attention to improving students' interest in learning mathematics knowledge and cultivating students' innovative thinking.

Table 2. Topics and their core vocabularies

Topic 5（0.245）		Topic 1 (0.139)		Topic 6 (0.104)	
Key words	Probability	Key words	Probability	Key words	Probability
Student	0.092	Learning	0.071	Student	0.091
Teacher	0.081	Teaching method	0.049	Cultivate	0.072
Learning	0.066	Resource	0.044	Learning	0.064
Method	0.057	Process	0.031	Knowledge	0.053
Teaching	0.053	Inquiry	0.026	Innovation	0.049

After the 13 topics were labeled, the corresponding relationship between the 7 factors in the TPACK

framework and the 13 topics are shown in Table 3.

Table 3. Types of knowledge in teachers' online discussion

TPACK factors	Example	Corresponding topics
TK	For the special functions of electronic whiteboard, such as searchlight, spotlight, etc., many teachers still cannot use them well. Most teachers can only use the text markers and courseware playback functions in the electronic whiteboard.	Topic 11
PK	The use of the game teaching method stimulates students' learning interest.	Topic 1, 2, 12
CK	Two equal sides in an isosceles triangle are called loins, while the two unparallel sides in a trapezoid are called loins.	Topic 7
PCK	In a paper-cutting activity, if students cut along the height of the isosceles trapezoid, they will get the rectangular trapezoid.	Topic 3, 5, 6, 8
TCK	The searchlight function in the electronic whiteboard is used to display trapezoids in daily life.	Topic 4
TPK	The searchlight function in the electronic whiteboard can amplify details and highlight key points.	Topic 10
TPACK	The searchlight function in the electronic whiteboard is used to compare the differences between triangles, quadrilaterals, and trapezoids, thus highlighting the features of trapezoids.	Topic 9, 13

4.3.What are the differences between the epistemic network characteristics of teachers in high-score and low-score groups?

In Figure 3, the red graph represented the epistemic network structure of teachers in the high-score group, and the blue one represented the epistemic network structure of teachers in the low-score group. In the epistemic network of teachers in the high-score group, there were more connections between PK and PCK, which indicated that the alternation of general pedagogy knowledge and pedagogical content knowledge occurred more frequently in the comments of high-score group. In the epistemic network of teachers in the low-score group, there were more connections between PK and CK, PK and TCK, PK and TK, which indicated that the knowledge types of teachers in low-score group were rich, but the integration between these knowledge types was poor.

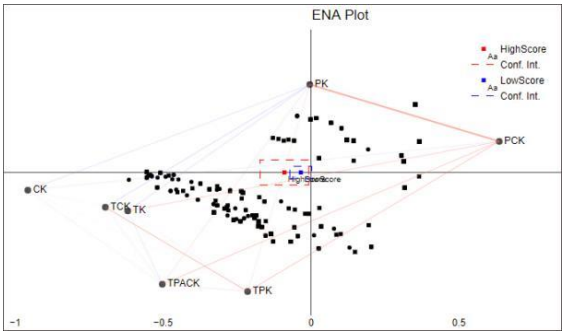


Figure 3. Epistemic network characters of high- and low-score groups teachers

4.4.What are the differences between the epistemic network characteristics of teachers in different age

groups?

In Figure 4, the red graph represents the epistemic network structure of teachers with age 5, and the blue graph represents the epistemic network structure of teachers with age 4. In the epistemic network of teachers with age 5, there were more connections between PK and PCK, TK and PCK, which indicated that teachers with age 5 tended to integrate technological knowledge, general pedagogical knowledge into subject content knowledge. That was, teachers with age 5 were more willing to explore the ways of applying information technology knowledge and general pedagogical knowledge to present subject content knowledge. In the epistemic network of teachers with age 4, there were more connections between PK and TCK, PCK and TCK, TCK and TPACK, which indicated that teachers with age 4 had different paths to develop TPACK than the teachers with age 5. Teachers with age 4 tended to integrate pedagogical knowledge into technological content knowledge. That was, on the basis of the determination of the topic-specific representations that utilize emerging technologies, teachers with age 4 were more willing to explore the application of subject-specific and content-specific teaching methods.

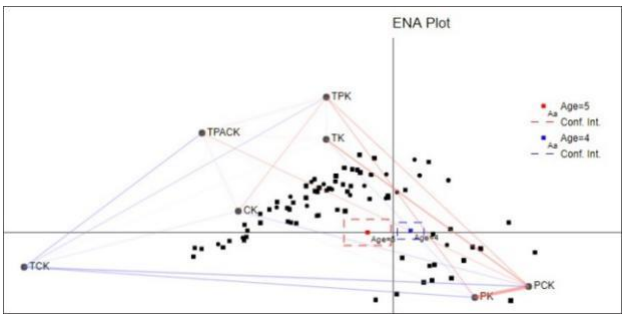


Figure 4. Epistemic network characters of teachers in different age groups (Age=5,4)

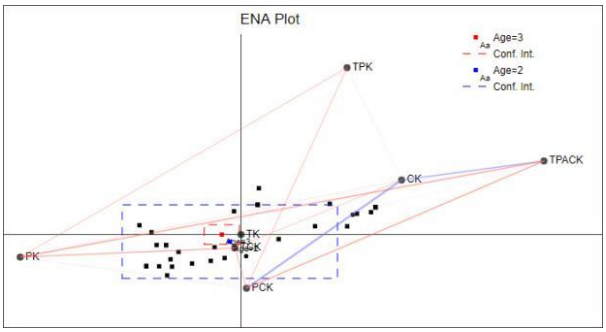


Figure 5. Epistemic network characters of teachers in different age groups (Age=3,2)

In Figure 5, the red graph represents the epistemic network structure of teachers with age 3, and the blue graph represents the epistemic network structure of teachers with age 2. In the epistemic network of teachers with age 3, there were more connections between PK and TPACK, PCK and TPACK, which indicated that teachers with age 3 tended to integrate technological knowledge into pedagogical content knowledge. In the epistemic network of teachers with age 2, there were more connections between CK and PCK, CK and TPACK, which indicated that teachers with age 2 had different paths to develop TPACK than the teachers with age 3. On the basis of the determination of topic-specific subject content, teachers with age 2 were more willing to explore the application of information technology and subject-specific teaching method to present the topic-specific subject content.

5. Discussion

In mathematics teaching, teachers realized the prominent role of information technology. However, the special functions of the electronic whiteboard had not become frequent words in teachers' comments (Hsieh & Tsai, 2017). Referring to Cox (2009)'s definition of technological knowledge, teachers' technological knowledge in this study was defined as knowledge of using special functions in electronic whiteboard, so there were relatively few comments related to technical knowledge (Jen, Yeh, Hsu, Wu, & Chen, 2016). Two possible reasons may be related to this phenomenon. First, although teachers had received training on the use of electronic whiteboard, they gradually forgot how to use the special functions because these functions were rarely used in teachers' classroom teaching. Second, without the support of technicians, teachers will stop using the special functions in electronic whiteboard once they encounter technical problems.

The epistemic network of teachers in the high-score group was mainly composed of pedagogical knowledge and pedagogical content knowledge, which indicated that teachers in high-score group paid great attention to subject-specific and topic-specific teaching methods and subject teaching content, and the connection between technological knowledge and other types of knowledge was not very close. The epistemic network of teachers in the low-score group was dominated by the single knowledges. Although they try to apply information technology to their teaching activity, their knowledge was not systematic. The epistemic network characteristics of teachers at different ages were different. Teachers aged 20-29 emphasized the application of information technology and subject-specific teaching method to present the subject content. Teachers aged 30-39 emphasized the integration of technical knowledge into pedagogical content knowledge in the development of the TPACK. Teachers aged 40-49 emphasized the integration of technological knowledge, pedagogical knowledge and the subject content knowledge in the development of TPACK. Teachers aged 50-59 emphasized the integration of technical knowledge and general pedagogical knowledge into subject content knowledge. Due to the differences in the development paths of TPACK among teachers of different ages, the teachers' training institutions should design reasonable courses for teachers of different ages that conformed to their TPACK development path (Jang & Tsai, 2012).

6. Conclusion, limitations and future study

This study collected teachers' online discussion data and used the LDA topic model to explore the topics of teachers' online discussion and the ENA to explore the teachers' epistemic network characteristics. The conclusion of this study provides a useful reference for improving the quality of teacher training and realizing the automatic analysis and evaluation of teachers' knowledge in the OPLC. However, there are two main limitations in this study. Firstly, the text data volume analyzed in this study was small, and the advantage of LDA unsupervised machine learning algorithm cannot be brought into full play. Secondly, the topics mined by LDA were labeled by manual method, which was prone to subjective deviation. In the following research, the data volume will be expanded and more online discussion text data will be collected for LDA analysis and modeling. In addition, the method of automatic topic labeling will be studied.

Reference

Arastoopour, G., Shaffer, D.W., Swiecki, Z., Ruis, A.R., & Chesler, N.C. (2016). *Teaching and assessing engineering design thinking with virtual internships and epistemic network analysis. International Journal of Engineering Education, 32(3B)*, 1492–1501

Burhan-Horasanlı, E., & Ortaçtepe, D. (2016). *Reflective practice-oriented online discussions: A study*

of EFL teachers' reflection-on, in and for-action. Teaching and Teacher Education, 59, 372-382.

Cox, S., & Graham, C. R. (2009). *Diagramming TPACK in practice: Using an elaborated model of the TPACK framework to analyze and depict teacher knowledge. TechTrends, 53(5)*, 60-69.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions (2 edition)*. John Wiley, New York.

Hatfield, D. (2015). *The right kind of telling an analysis of feedback and learning in a journalism epistemic game. Inter-national Journal of Gaming and Computer-Mediated Simulations, 7(2)*, 1–23.

Hsieh, W.-M., & Tsai, C.-C. (2017). *Taiwanese high school teachers' conceptions of mobile learning. Computers & Education, 115*, 82-95.

Jang, S.-J., & Tsai, M.-F. (2012). *Exploring the TPACK of Taiwanese elementary mathematics and science teachers with respect to use of interactive whiteboards. Computers & Education, 59*, 327-338.

Jen, T.-H., Yeh, Y. -F., Hsu, Y. -S., Wu, H. -K., & Chen, K. -M. (2016). *Science teachers' TPACK-Practical: Standard-setting using an evidence-based approach. Computers & Education, 95*, 45-62.

Lee, K., & Brett, C. (2015). *Dialogic understanding of teachers' online transformative learning: A qualitative case study of teacher discussions in a graduate-level online course. Teaching and Teacher Education, 46*, 72-83.

Liu, M. -H. (2012). *Discussing teaching video-cases online: Perspectives of preservice and in-service EFL teachers in Taiwan. Computers & Education, 59*, 120-133.

Prestridge, S. (2010). *ICT professional development for teachers in online forums: Analyzing the role of discussion. Teaching and Teacher Education, 26*, 252-258.

Rienties, B., Kinchin, I. (2014). *Understanding (in)formal learning in an academic development program: A social net-work perspective. Teaching and Teacher Education, 39*, 123-135.

Romero, C., & Ventura, S. (2013). *Data mining in education. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 3(1)*, 12-27.

Wu, P. F, Yu, S. Q., & Wang, D. (2018). *Using a learner-topic model for mining learner interests in open learning environment. Educational Technology and Society, 21(2)*, 192-204.

Xu, H., Zhang, F., & Wang, W. (2015). *Implicit feature identification in Chinese reviews using explicit topic mining model. Knowledge-Based Systems, 76*, 166-175

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

The Role of Peer Feedback on the Quality of Students' Computer-Supported Collaborative Argumentation

Wenli Chen[1], Eng Eng Ng[1*], Xinyi Li[2], Siew Cheng Aileen Chai[3] and Qianru Lyu[1]
[1] *National Institute of Education, Nanyang Technological University, Singapore*
[2]*Beijing Normal University, China*
[3]*Ministry of Education, Singapore*
* caryn.ng@nie.edu.sg

Abstract: *The importance of peer feedback in collaborative argumentation has been well-established. However, little is known about the extent to which peer feedback is associated with the quality of collaborative argumentation. Particularly, there is limited evidence for how specific types of feedback is related to argumentation quality. This study investigated peer feedback against four dimensions of collaborative argumentation quality (clarity, multiple perspectives, selection of evidence, and elaboration and depth). Collaborative argumentation quality was also compared against peer feedback types (appropriateness, specificity, and elaboration). In this design-based research (DBR), a class of 40 secondary Grade Three students in Singapore participated in three cycles of argumentation and peer feedback activities using the AppleTree online learning environment, each cycle consisting of five collaborative learning phases scripted by the Spiral Model of Collaborative Knowledge Improvement (SMCKI): Individual ideation, group synergy, peer critique, group refinement, and individual achievement. Scaffolds of sentence openers and reflections were added in Cycles 2 and 3. Quantitative analyses comparisons of argumentation and per feedback quality across three cycles revealed that except for the multiple perspectives dimension of argumentation quality, students performed significantly better in forming their argumentations and giving peer feedback. Additionally, the quality of argumentation improved significantly over the three cycles when accounting for peer feedback types as correlates, and vice versa.*

Keywords: collaborative argumentation, peer feedback, computer-supported collaborative learning

1. Introduction

Argumentation is a necessary skill that fosters higher-order cognitive abilities as students dissent, evaluate, expand, and synergize claims, evidence, and counterarguments to reach conclusions (Harney et al., 2017; Haro et al., 2020; Latifi et al., 2020, 2021; Noroozi et al., 2022). Such higher-order thinking skills involve knowledge construction, concept integration, critical thinking, and reflective judgments (Harney et al., 2017; King, 2002; Lu & Zhang, 2013). Quality argumentation also affords the opportunity to construct knowledge from multiple perspectives and evaluate each other's arguments which help students realize their areas of weakness, leading them to refine their ideas (Clark et al., 2007; Mayweg-Paus et al., 2021). Additionally, engaging in in-depth explanations during argumentation work enables students to process information at a deeper level which fosters a better mastery of the concepts learnt, generation of interrelationships among concepts, and integration of new concepts with prior knowledge (Asterhan & Schwarz, 2007) .

Well-constructed argumentation is crucial to peer feedback/assessment/review in Computer-Supported Collaborative Learning (CSCL) research. Essentially, feedback refers to the student-student, student-teacher, or teacher-teacher interactions, reception of feedback, and decision to implement feedback (Carless et al., 2011; Guasch et al., 2013; Nicol & Macfarlane-Dick, 2006). While it is well established that peer feedback improved student collaborative work (Guasch et al., 2013), there are limited studies that can elucidate the contributions of peer feedback on specific dimensions of argumentation quality in computer-supported collaborative argumentation (CSCA). This DBR study sheds light on how specific argumentation types and quality relate to peer feedback support (Guasch et al., 2013; Noroozi et al., 2022).

2. Literature Review

Collaborative argumentation is a process that enhances knowledge building and content understanding by working with others to form individual arguments and weigh the evidence of others' arguments without any social pressures to conform to a single position at the outset (Golanics & Nussbaum, 2008; Nussbaum, 2008). The term "collaborative" implies a dialogic interaction. The socio-cognitive theory tied with Piaget's constructivism model states that learning new concepts is strengthened when discussions and reflections addressing conflicting arguments occur (Mugny & Doise, 1978; Noroozi et al., 2020; Nussbaum, 2008). Tools such as the rubrics may be used to aid students (Lu & Zhang, 2013).

Peer feedback allows students to actively self-regulate learning independently of teachers' input and class size, thus, enabling feedback provision in a more timely manner (Latifi et al., 2020). While peer feedback may differ on the amount of interactions (low level of interaction if feedback is given through engaging in the main activity to high level of dialogic interaction between peers), it is regarded as a social and interpersonal process (Deiglmayr, 2018). Similar to argumentation, support for students is provided in the form of prompts, rubrics, and scaffolding in order to help students utilize and implement peer feedback as an effective pedagogical tool (Deiglmayr, 2018; Harney et al., 2017).

The online workspace affords unique opportunities for collaborative argumentation and peer feedback. Students complete, store, and submit their argumentation work without being held back by time and space constraints (Tsai, 2009). Moreover, students give and receive feedback anonymously, modify, continuously reflect on the feedback given by their peers, and improve their work (Noroozi et al., 2016; Tsai, 2009) while teachers create and manage student groupings based on similar demographic profiles (Lu & Zhang, 2013). Although students may anonymously provide more honest and emotionally-provoking critiques of their peers' work online, teachers can monitor students' work and interactions.

2.1 Theoretical Framework

Toulmin's 1958 model of argumentation posits that an argument consists of claims, evidence, warrants, backing, qualifiers, and rebuttals (Lu & Zhang, 2013). This study focuses on argumentation in the form of claims backed by evidence and rebuttals against the claims (Bouyias & Demetriadis, 2012). A claim is an assertion or conclusion about the topic being argued. Evidence supports the claim and rebuttals refute the claim. Sound argumentation is the basis for providing good quality feedback due to the cognitive/logical reasoning skills associated with quality argumentation (Noroozi et al., 2022). A critical and reflective review of peers' work requires identifying strengths and weaknesses of others' work against the assessment criteria, persuading others of areas for improvement and offering

suggestions to revise future work (Deiglmayr, 2018). Alvarez et al. (2011)'s model indicates that feedback can be understood in terms of corrective feedback: whether the responses adequately address the assignment question; epistemic feedback: in-depth clarifications or explanations to the responses, and suggestive feedback: improvements and further elaborations on the topic. Students may not feel secure about the feedback they received if their peers are unable to provide "quality" feedback (Guasch et al., 2013). Challenging and critiquing peers' arguments require the ability to reason, think critically about opposing arguments, and amend one's arguments upon reflection on peer feedback i.e., using metacognition, thus facilitating deep learning (Loll & Pinkwart, 2013). Research has evidenced that good quality peer feedback potentially produced better quality argumentation. However, the mechanisms underlying quality peer feedback and conceptualizations of argumentation are little understood. For example, Lu & Zhang (2013) conceptualized argumentation quality into evidence, claims, reasoning, and knowledge application, and found that peer feedback only significantly improved the evidence dimension, implying that other aspects of argumentation are still being honed.

While past studies have established that peer feedback improved student collaborative writing (Guasch et al., 2013), there are not many studies within CSCA that can shed light on whether peer feedback contributed to specific dimensions of argumentation quality and vice versa. In addition, the contribution of peer feedback scaffolding in aiding argumentation work remains to be explored. Collaborative argumentation goes beyond sharing ideas and thoughts found in collaborative writing to forming claims, and providing supporting evidence and counterarguments. Additionally, communicating on an online collaborative workspace has unique attributes (Lu & Zhang, 2013; Tsai, 2009; Tseng & Tsai, 2007). Many previous studies on peer feedback and collaborative argumentation have emphasized higher cognitive skills and group work in higher education as these skills are highly sought after by employers hiring university graduates (Deiglmayr, 2018; Guasch et al., 2013; Latifi et al., 2021; Noroozi et al., 2022).

Therefore, this study (1) compared the differences between collaborative argumentation quality dimensions (clarity, multiple perspectives, selection of evidence, and elaboration and depth) and quality of peer feedback types (appropriateness, specificity, and elaboration) received across cycles; (2) examined argumentation quality against peer feedback types; and (3) tested whether peer feedback quality was related to dimensions of argumentation quality.

3. Method

3.1 Study Design and Participants

We chose DBR as our methodology because it allows us to work closely with teachers and students in a real-life classroom setting to transform practices in knowledge outcomes, and generate artefacts and student reflections in a socio-technical context (Anderson & Shattuck, 2012; Prieto et al., 2020). DBR was utilized to assess the quality of collaborative argumentation and peer feedback in three cycles whereby Cycle 1 served as the baseline, Cycle 2 introduced scaffolds of sentence openers, and Cycle 3 provided an opportunity for students to reflect. Each cycle lasted for 1.5 hours. A class of 40 Grade Three students were sampled from a co-educational government secondary school in Singapore. Participants collaborated with one another in 10 randomly assigned groups of 4 students to address the extent in which they provided evidence for their claims and rebuttals/counterarguments surrounding the following Artificial Intelligence (AI)-related topics: (1) 'AI will make transportation better.', (2) 'AI will make our lives at home more enjoyable.', and (3) 'AI will make our jobs easier.' Participation in this study was

strictly voluntary and a signed written informed consent was obtained from each participant prior to the data collection from August to October 2022.

3.2 Measures

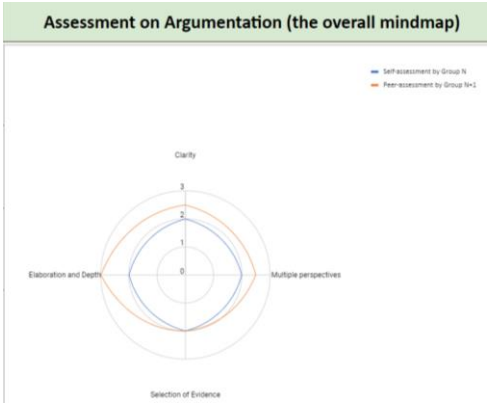
We collected argumentation graphs and peer critique data from the participants. We analyzed the quality of argumentation and type of feedback using content analysis. The coding framework on student-generated argumentation were adapted from Felton & Kuhn (2001)'s Utterance Types for Coding Argumentive Dialogue scheme which consists of these dimensions: Clarify, adopt a position/counterargument(s), additional evidence to justify one's claims, and elaboration. The four dimensions in our coding framework are (1) clarity: Well-formed central idea and purpose without any repetition; (2) multiple perspectives: Synthesis of varied and individual perspectives; (3) selection of evidence: Sound claims supporting a wide variety of perspectives; and (4) depth and elaboration: Claims with detailed explanations demonstrating deep insight and line of thought. These four dimensions are scored on a scale of 0 = Poor, 1 = Needs work, 2 = Good, and 3 = Excellent. Quality of argumentation has been independently coded by two researchers who calculated the average score (Kappa = .91) of these four dimensions in the argumentation rubrics.

During Phase 3 (Peer Critique), groups were paired up and instructed to provide feedback to the other group. The feedback types are praise, problem, suggestion, and clarify (Nelson & Schunn, 2009), of which the latter three feedback types are further divided into and adapted from Gielen et al. (2010)'s model of quality peer feedback: (1) appropriateness: Relevant feedback is provided based on the rubrics and criteria of the given task (Inter-rater reliability (IRR): $r = .73$, $p < .01$); (2) specificity: Clear solutions to problem areas and weaknesses (IRR: $r = .063$, $p < .01$); and (3) elaboration: Explanation on why the problem occurred and/or how the solution addresses the problem (IRR: $r = .49$, $p < .05$).

3.3 Procedure and Data Analysis Strategy

The collaborative argumentation lessons were divided into five phases based on the Spiral Model of Collaborative Knowledge Improvement (SMCKI) (First Author et al., 2021) – (1) Individual ideation: Students work individually to write their claims and evidence; (2) group synergy: Individual claims are merged into a single claim; (3) peer critique: Groups are paired up to comment on each other's argumentation graph (see Figure 1); (4) refinement: Students view and reflect on the comments given to them by the other group, making necessary amendments; and (5) individual achievement: Students reflect on the argumentation activity. In line with Piaget's social cognitive theory, SMCKI aims to facilitate knowledge creation at the individual, group, and class levels. 60 collaborative argumentation graphs and peer feedback were generated via the AppleTree collaborative learning platform (Figure 1) and the peer critique space, Radarmap Peer Critique Tool (Figure 2). We further collected peer feedback using sentence openers and reflections in the Peer Critique Tool in Cycles 2 and 3 respectively. An example of a peer feedback with the sentence opener, "Identify Strengths: We like…" is followed by the student's response: "…*your points on people losing their jobs and AI being more efficient than them.*" An example of a reflection is, "We agree… *that our points are great, but we need to break up points to make it (them) easier to absorb.*"

Figure 1. Sample of argumentation graph (First Author, 2022).



Assessment			Rubrics			
Dimensions/Feedback	Self-assessment by Group N	Peer-assessment by Group N+1	0	1	2	3
Clarity	2	2.5	The central idea and clarity of purpose are absent.	The central idea is basically expressed but it is vague. There are many repetitive ideas. Readers have difficulties understanding the argumentation.	The central idea and the clarity of purpose are generally evident. A few repetitive ideas. Readers occasionally have confusions.	The central idea and the clarity of purpose are well developed; no repetitive ideas, very compelling statements. Readers have a clear understanding of the argumentation.
Multiple perspectives	2	2.5	No consideration of multiple perspectives; No stakeholder considered	Insufficiently explore multiple perspectives Consider limited type of stakeholders	Display multiple perspectives though not consistently done Consider many type of stakeholders but not systematically.	Synthesizing multiple perspectives while valuing contributions from these perspectives Consider variety of stakeholders systematically
Selection of Evidence	2	2	No evidence at all or evidence provided is not credible or relevant	Limited evidence provided. Evidence provided is somewhat credible and relevant but not supporting the multiple perspectives	Much evidence has been provided. Evidence provided is credible and relevant and supporting multiple perspectives, though not consistently done	A rich variety of credible and relevant evidence which consistently supports multiple perspectives
Elaboration and Depth	2	3	Irrelevant ideas, or ideas with no elaboration	Little or superficial development of claims and evidence; lack of explanation and elaboration	Shows development of claims and evidence, with explanation and elaboration; demonstrate in-depth thinking process such as synthesizing, reflecting, evaluating	Shows an extensive development of claims and evidence with with thorough explanation and elaboration, Demonstrate in-depth thinking processes such as synthesizing, reflecting, evaluating
Average	2	2.5				

Figure 2. Sample of the Radarmap Peer Critique Tool rubrics and peer assessment.

We used IBM SPSS version 28 to analyze the data. To answer our first research hypothesis, we ran one-way repeated measures analyses of variance (ANOVA) to examine the differences among collaborative argumentation, argumentation quality at Phase 5 and quality of peer feedback received across cycles. In addition, we conducted paired sample t-tests to compare between the quality of argumentation at Phase 1 and 5. To address our second and third research hypotheses, we performed one-way analysis of covariance (ANCOVA) to compare the quality of collaborative argumentation and peer feedback across three cycles while controlling for specific corresponding correlates of the variables.

4. Results

One-way repeated measures ANOVAs revealed significant differences in quality of argumentation between test cycles, $F(1, 9) = 24.29$, $p < .001$, Phase 5 quality of argumentation, $F(1, 9) = 26.56$, $p < .001$, and quality of peer feedback $F(1, 9) = 23.91$, $p < .001$. Post-hoc related samples t-tests using Bonferroni's adjustment were run to compare the differences across cycles. The findings revealed that students performed significantly better in all aspects during Cycle 3 as compared to Cycle 1 ($p < .001$). However, students did not perform significantly better or worse in collaborative argumentation and Phase 5 quality of argumentation when comparing Cycle 2 with Cycles 1 and 3 ($p > .05$). Meanwhile, students appeared to receive better quality feedback from their peers when comparing Cycles 1 and 2 against Cycle 3 ($p < .001$). However, students did not receive significantly different feedback in Cycle 2 than in Cycle 1 ($p > .05$). Paired samples t-tests revealed that students performed better in collaborative argumentation quality and its dimensions at Phase 5 than at Phase 1 (Table 1): quality of argumentation $t(9) = -5.30$, $p < .001$; clarity $t(9) = -2.76$, $p < .05$; selection of evidence $t(9) = -2.76$, $p < .05$; and elaboration and depth

$t(9) = -3.53$, $p < .01$. However, students did not perform significantly better at Phase 5 than at Phase 1 for multiple perspectives $t(9) = -1.68$, $p > .05$.

Table 1. Means and standard deviations for the quality of argumentation and its dimensions.

	Phase 1		Phase 5	
	M	*SD*	*M*	*SD*
Argumentation Quality	8.18	2.10	8.77	2.08
Clarity	1.90	0.66	2.05	0.63
Multiple Perspectives	2.25	0.61	2.32	0.56
Selection of Evidence	2.02	0.52	2.18	0.59
Elaboration and Depth	2.02	0.58	2.22	0.55

One-way ANCOVAs revealed significant differences in argumentation quality between the three cycles when controlling for appropriateness [$F(2,26) = 4.86$, $p < .05$], specificity [$F(2,26) = 5.29$, $p < .05$], and elaboration [$F(2,26) = 3.98$, $p < .05$]. Post-hoc Bonferroni tests showed significant differences between Cycles 1 and 3 ($p < .05$) when controlling for appropriateness, specificity, and elaboration. Similarly, there were significant differences in quality of peer feedback between the three cycles when controlling for clarity [$F(2,26) = 7.32$, $p < .01$], multiple perspectives [$F(2,26) = 11.71$, $p < .001$], selection of evidence [$F(2,26) = 9.99$, $p < .001$], and elaboration and depth [$F(2,26) = 10.34$, $p < .001$]. Post-hoc Bonferroni tests also showed significant differences between Cycles 1 and 3 ($p < .01$), and Cycles 2 and 3 ($p < .05$) when controlling for clarity, and between Cycles 1 and 3 ($p < .001$) and Cycles 2 and 3 ($p < .05$) when controlling for multiple perspectives, selection of evidence, and elaboration and depth. Comparing the estimated marginal means (Table 2) showed that the quality of argumentation and peer feedback increased on Cycles 2 and 3 compared to Cycle 1.

Table 2. Estimated marginal means for the quality of argumentation and peer feedback.

Cycles	1	2	3
Argumentation quality when controlling for…			
Appropriateness	7.29	8.94	10.07
Specificity	6.76	9.15	10.39
Elaboration	7.40	8.97	9.93
Quality of peer feedback when controlling for…			
Clarity	4.82	5.46	6.46
Multiple Perspectives	4.70	5.51	6.53
Selection of Evidence	4.63	5.50	6.61
Elaboration and Depth	4.73	5.52	6.49

5. Discussion

Students performed better in the quality of argumentation and peer feedback during Cycle 3 than in Cycles 1 and 2. Students may have needed some time to adapt to the online platform during Cycle 1. Students also improved in argumentation dimensions at Phases 1 and 5 – except for multiple perspectives. Nonetheless, students were provided sentence openers and reflections to function as scaffolds during

Cycles 2 and 3. Prior research maintained that critical reasoning tasks in collaborative argumentation promoted clarity, multiple perspectives, selection of evidence, and elaboration and depth (Felton & Kuhn, 2001). There was also a gradual increase in students' ability to give quality feedback to others using sentence openers and reflections from Cycle 1 to 2 and Cycle 2 to 3. In addition, quality argumentation was linked with quality peer feedback and vice versa, while controlling for specific correlates of each corresponding variable. Argumentation quality and peer feedback improved significantly with each subsequent cycle.

As expected, argumentation quality improved peer feedback (Guasch et al., 2013). Justifying "good" arguments and reasoning with evidence-based claims predicted the ability to give constructive feedback (Deiglmayr, 2018; Lu & Zhang, 2013; Noroozi, 2022). Peer feedback helps students to collate the gap between their work and the assessment criteria, bridging the gap by making changes to their work (Haro et al., 2019). However, as peer feedback – like argumentation – involves higher order cognitive skills, giving feedback does not immediately result in improved argumentation (Guasch et al., 2013; Kollar & Fischer, 2010). We did not find significant improvements in multiple perspectives argumentation dimension across cycles. Students develop multiple perspectives on a topic and converge to discuss opposing views, reflecting on the claims and evidence to resolve a conflict (Asterhan & Schwarz, 2007; Nussbaum, 2008). Toulmin did not regard the interactive and social contexts of argumentation between opponents, implying that the students in our study may have used technology to seek for information, and evaluate their claims and evidence without dissenting alternative views (Mayweg-Paus et al., 2021; Pechorro et al., 2017). In the recent times, prior knowledge appears to enhance recall of arguments (Schmidt et al., 2017). As prior knowledge was used during the individual ideation Phase 1 (First Author, 2021), students' knowledge in different domains may have contributed to forming arguments from multiple perspectives.

The present study must be interpreted in light of some limitations. Testing repeatedly presents a maturity effect whereby students improved their performance over time. Although prior studies have suggested that gender differences played a role in collaborative argumentation and peer review (Noroozi et al., 2020, 2022), the nature of our student groupings did not allow us to examine gender differences in our study. Additionally, we did not investigate excluded mediating factors to explain the lack of significant increase in the multiple perspectives dimension of argumentation. As the combination of "expert" teacher and peer feedback showed most potential in improving argumentation (Guasch et al., 2013), future studies may replicate the current study findings with students and teachers.

6. Conclusion

The current DBR research showed that students increasingly performed better in the quality of argumentation and peer feedback at later cycles. Except for multiple perspectives, students improved all other specific dimensions of argumentation. There was also a gradual increase in students' ability to give quality feedback to others using sentence openers and reflections as scaffolds with each subsequent cycle. This study implicates further research and interventions incorporating gender differences, mediating variables, and instructors providing sentence openers and reflections.

References

Alvarez, I., Espasa, A., & Guasch, T. (2011). The value of feedback in improving collaborative writing assignments in an online learning environment. *Studies in Higher Education*, *37*(4), 387–400.

Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research? *American Educational Research Association*, *41*(1), 16–25.

Asterhan, C. S., & Schwarz, B. B. (2007). The effects of monological and dialogical argumentation on concept learning in evolutionary theory. *Journal of Educational Psychology*, *99*(3), 626–639.

Bouyias, Y., & Demetriadis, S. (2012). Peer-monitoring vs. micro-script fading for enhancing knowledge acquisition when learning in computer-supported argumentation environments. *Computers and Education*, *59*(2), 236–249.

Carless, D., Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in Higher Education*, *36*(4), 395–407.

Clark, D. B., Stegmann, K., Weinberger, A., Menekse, M., & Erkens, G. (2007). Technology-enhanced learning environments to support students' argumentation. In S. Erduran & M. Jiménez-Aleixandre (Eds.), *Argumentation in Science Education* (pp. 217–243). Springer.

Deiglmayr, A. (2018). Instructional scaffolds for learning from formative peer assessment: effects of core task, peer feedback, and dialogue. *European Journal of Psychology of Education*, *33*(1), 185–198.

Felton, M., & Kuhn, D. (2001). The development of argumentative discourse skill. *Discourse Processes*, *32*(2), 135–153.

Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and Instruction*, *20*(4), 304–315.

Golanics, J. D., & Nussbaum, E. M. (2008). Enhancing collaborative online argumentation through question elaboration and goal instructions. *Journal of Computer Assisted Learning*, *24*, 167–180.

Guasch, T., Espasa, A., Alvarez, I. M., & Kirschner, P. A. (2013). Effects of feedback on collaborative writing in an online learning environment. *Distance Education*, *34*(3), 324–338.

Harney, O. M., Hogan, M. J., & Quinn, S. (2017). Investigating the effects of peer to peer prompts on collaborative argumentation, consensus and perceived efficacy in collaborative learning. *International Journal of Computer-Supported Collaborative Learning*, *12*(3), 307–336.

Haro, A. V., Noroozi, O., Biemans, H. J. A., & Mulder, M. (2020). Argumentation competence: Students' argumentation knowledge, behavior and attitude and their relationships with domain-specific knowledge acquisition. *Journal of Constructivist Psychology*, *35*(1), 1–23.

Haro, A. V., Noroozi, O., Biemans, H. J. A., & Mulder, M. (2019). The effects of an online learning environment with worked examples and peer feedback on students' argumentative essay writing and domain-specific knowledge acquisition in the field of biotechnology. *Journal of Biological Education*, *53*(4), 390–398.

King, A. (2002). Structuring peer interaction to promote high-level cognitive processing. *Theory Into Practice*, *41*(1), 33–39.

Kollar, I., & Fischer, F. (2010). Peer assessment as collaborative learning: A cognitive perspective. *Learning and Instruction*, *20*(4), 344–348.

Latifi, S., Noroozi, O., Hatami, J., & Biemans, H. J. A. (2021). How does online peer feedback improve argumentative essay writing and learning? *Innovations in Education and Teaching International*, *58*(2), 195–206.

Latifi, S., Noroozi, O., & Talaee, E. (2020). Worked example or scripting? Fostering students' online argumentative peer feedback, essay writing and learning. *Interactive Learning Environments*, 1–15.

Loll, F., & Pinkwart, N. (2013). LASAD: Flexible representations for computer-based collaborative argumentation. *International Journal of Human Computer Studies*, *71*(1), 91–109.

Lu, J., & Zhang, Z. (2013). Assessing and supporting argumentation with online rubrics. *International Education Studies*, *6*(7), 66–77.

Mayweg-Paus, E., Zimmermann, M., L., N. T., & Pinkwart, N. (2021). A review of technologies for collaborative online information seeking: On the contribution of collaborative argumentation. *Education and Information Technologies, 26*.

Mugny, G., & Doise, W. (1978). Socio-cognitive conflict and structure of individual and collective performances. *European Journal of Social Psychology*, *8*(2), 181–192.

Nelson, M. M., & Schunn, C. D. (2009). The nature of feedback: How different types of peer feedback affect writing performance. *Instructional Science, 37*(4), 375–401.

Nicol, D., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218.

Noroozi, O., Biemans, H. J. A., & Mulder, M. (2016). Relations between scripted online peer feedback processes and quality of written argumentative essay. *Internet and Higher Education*, *31*(1), 20–31.

Noroozi, O. (2022). The role of students' epistemic beliefs for their argumentation performance in higher education. *Innovations in Education and Teaching International*, 1–12.

Noroozi, O., Banihashem, S. K., Kerman, N. T., Parvaneh A. K., M., Babayi, M., Ashrafi, H., & Biemans, H. J. A. (2022). Gender differences in students' argumentative essay writing, peer review performance and uptake in online learning environments. *Interactive Learning Environments*, 1–15.

Noroozi, O., Hatami, J., Bayat, A., van Ginkel, S., Biemans, H. J. A., & Mulder, M. (2020). Students' online argumentative peer feedback, essay writing, and content learning: Does gender matter? *Interactive Learning Environments*, *28*(6), 698–712.

Noroozi, O., Weinberger, A., Biemans, H. J. A., Mulder, M., & Chizari, M. (2012). Argumentation-Based Computer Supported Collaborative Learning (ABCSCL): A synthesis of 15 years of research. *Educational Research Review*, *7*(2), 79–106.

Nussbaum, M. E. (2008). Collaborative discourse, argumentation, and learning: Preface and literature review. *Contemporary Educational Psychology*, *33*(3), 345–359.

Pechorro, P., Hawes, S. W., Gonçalves, R. A., & Ray, J. V. (2017). Psychometric properties of the inventory of callous-unemotional traits short version (ICU-12) among detained female juvenile offenders and community youths. *Psychology, Crime and Law*, *23*(3), 221–239.

Prieto, L. P., Magnuson, P., Dillenbourg, P., & Saar, M. (2020). Reflection for action: Designing tools to support teacher reflection on everyday evidence. *Technology, Pedagogy and Education*, *29*(3), 279-295.

Schmidt, K. H., Rothgangel, M., & Grube, D. (2017). Does prior domain-specific content knowledge influence students' recall of arguments surrounding interdisciplinary topics? *Journal of Adolescence*, *61*, 96–106.

Tsai, C. C. (2009). Internet-based peer assessment in high school settings. In L. W. H. Tan & R. Subramaniam (Eds.), *Handbook of research on new media literacy at the K-12 level: Issues and challenges* (pp. 743–754). Information Science Reference.

Tseng, S., & Tsai, C. C. (2007). On-line peer assessment and the role of the peer feedback: A study of high school computer course. *Computers & Education*, *49*(4), 1161–1174.

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

Research review on the application of eye-movement technique in Education-Based on applied bibliometrics and content analysis

Yibei Shen [1], Jiajie Song [1], Chien-Liang Lin [1*], Yu-Sheng Su [2*],
[1] College of Science & Technology Ningbo University
[2] National Taiwan Ocean University, Taiwan
[*] lin.chienliang@gmail.com, ntoucsiesu@mail.ntou.edu.tw

Abstract: In recent years, eye-movement experiments have been applied to a wide range of fields, especially in education, where eye-tracking can be used to study education and learning processes. In the context of the COVID-19 pandemic, more and more countries around the world are trying to experiment with new teaching methods, making eye-movement technology increasingly important in analyzing the effectiveness of higher education as well. Experts and scholars in the field of education have also applied eye-movement technology in the past few years to explore the effectiveness and application of teaching and learning in more diversified analyses. Although eye-movement experiments have been widely accepted by academia, there is no relevant research at this stage in related pedagogy literature and trend analysis in terms of eye-movement experiments. Therefore, based on bibliometrics, this study investigates the application of eye movements in educational behavioral analysis during the period 2009-2021. The relationship of journals, authors, organizations, countries, paper years, keywords and citations is analyzed in knowledge graphs. The following conclusion is ultimately reached: researchers are the key to the progress of research on related topics. At the same time, prolific authors have contributed greatly to the advancement of the relevant educational fields. In addition, the highly cited authors are largely from highly cited countries and research institutions, with the U.S. taking a significant lead in this area. In a bibliometric approach, the study attempts to identify the integration of eye-movement and education research and to propose some new views and research directions for follow-up studies.

Keywords: eye-movement technique, eye-tracking, Neuroscience, bibliometric study, content analysis

1. Introduction

In recent years, eye-tracking experiments have been widely applied in various fields, including education (Coskun & Cagiltay, 2021; Rets & Rogaten, 2021), marketing (Segijn et al., 2021), information systems (Brinton Anderson et al., 2016) and management. Especially in education, eye movement tracking can be used in research on education and learning process. There has been a continuous increase in related papers while the related research has been performed at a much deeper level. Eye movement tracking has made great contributions to the field of educational neuroscience. Previous studies mostly emphasized the benefits of eye movement tracking apps to the cognition of both explorers and learners (da Silva Soares et al. 2021). Some cases in point include understanding the reading process of learners by tracking their eye movement based on conceptual graph learning strategy (Liu, 2014); applying eye movement experiments to improve the teaching design based on computer-assisted learning and testing environment (Jarodzka et al. 2017); or tracking the visual attention of learners in the multimedia

demonstrations in a physical classroom (Yang et al. 2013). Some studies analyzed the visual cognition of teachers towards classroom events through eye movement experiments (van den Bogert, 2014). To study the application and progress of eye movement experiments in the field of education at a deeper level, more scholars have widely discussed related research topics in recent years.

Under the impact of the COVID-19 pandemic, universities across the world have tried to adopt various new teaching methods, which also raises the importance of education (Lin et al., 2021; Jin et al., 2021; Mo et al., 2021). Scholars who are expert in the field of education have also started widespread discussions on the research trends and applications in the last few years, such as participation of students in flipped math class (Lo & Hew, 2021), surgical education in the field of medical care (Merali et al., 2019), and behavioral analysis of subjects in teaching and learning environments (Goldberg; Schwerter et al., 2021). In the past, bibliometrics have proven a popular research method. However, the literature review on eye movement experiments only includes frontier analysis on the application of eye movement in linguistic studies (Aryadoust & Ang, 2021) and application of eye movement tracking technique in the field of medical care (Zammarchi & Conversano, 2021). There has been a lack of related studies or discussions on incorporating eye movement in education. As a result, this study proposed an analysis on the application of eye movement for behavioral analysis in the field of education between 2009 and 2021 based on bibliometrics. Therefore, based on bibliometrics, this study reviewed related literature within the predetermined period in order to extract the key information of the research papers in the field of education, including journals, authors, organizations, countries, year of papers, keywords, and citations. By forming a knowledge network graph, this study provided meaningful reference for related scholars and practitioners.

2. Data Sources

This paper is based on the database from Social Science Citation Index (SSCI) and Science Citation Index (SCI) indexed by Web of Science (WOS) and focuses on the following keywords as the criteria for data screening, including “eye tracking”, or “education”, “movement” or “education” and “eye tracking”, or “education”. The data was retrieved from 2009 to 2021 as the journal of “Computer & Education” set up a special issue themed on “Learning with ICT: New perspectives on help seeking and information searching” in 2009. This issue mainly explores the related applications of eye movement technique. The other reason that this paper is based on the number of published papers in 2009 is that there were almost no papers themed on the application of eye movement technique in the field of education prior to 2009. Therefore, the literature retrieval in this paper starts from 2009. There are still some issues in the papers we retrieved from WOS. For example, the titles and keywords of some papers are merely descriptions of certain terms in the papers. Thus, data cleansing is required to improve the sample quality and enhance the reliability of the bibliometrics analysis results (Cobe et al., 2011). However, data cleansing cannot make judgment through other analytical tools or keywords. Instead, data must be filtered in a manual manner. As a result, in reference to the practice of Lin et al. (2022), Su et al. (2020) the samples collected in this study were manually filtered. Specifically, in the filtering process, researchers preliminarily read the extracts of the papers and determined if the research topic and content of the papers are consistent with the topic through a multi-party check. If they are not correlated, the papers will be removed. Through above process, a total of 131 papers were removed as the research content contracts this study or the type of research is identical to that of this study, while a total of 221 papers are consistent.

3. Results

3.1. Publication Trends

The publishing period was set from 2009 to 2021 and the paper type was set as "article". Finally, a total of 221 papers were consistent with the topic. Figure 2 shows the number of published papers in the field of eye tracking and education every year from 2009 to 2021. Between 2009 and 2013, no more than five papers were published in this field each year, which shows that the research in this field was only preliminary and laid the basic framework. Between 2014 and 2017, the number of published papers in this field kept increasing. The number of published papers soared in 2019 and reached 59 in 2020, which may be closely related to the impact of COVID-19 pandemic in 2019. Under the impact of the pandemic, the patterns of teaching, learning, and working underwent great changes from in-person to remote. The methods and manners of educational teaching have changed at a rapid rate, which triggered many scholars to reflect on the issues in remote teaching and learning under the impact of the pandemic. From 2009 to 2021, the number of published papers has risen step by step while the cross field of eye movement experiments and education has become the hot topic and frontier in both the two fields. Figure 1 shows the number of published papers every year.

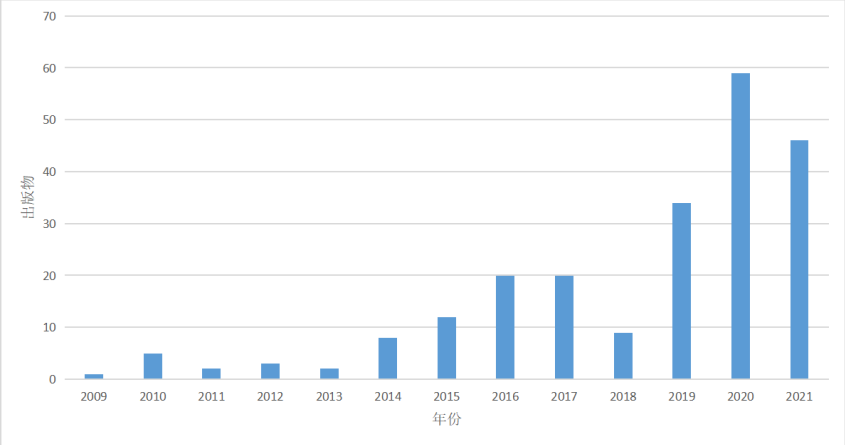


Figure 1. Number of publications per year from 2009 to 2021.

3.2. Descriptive analysis

This study statistically analyzed the number of published papers in different journals based on the 221 papers related to neuromarketing retrieved from WOS (in Table 1), in order to observe the progress of disciplinary knowledge structure of neuromarketing and provide guiding reference for future researchers on related paper submissions. Table 1 shows the number of published papers and citations of selected journals. Specifically, Computers & Education and Journal of Chemical Education have published the most papers, 10 papers published by each mostly after 2020. Computers & Education focuses on the cross application field of engineering technology and computer while Journal of Chemical Education focuses on the education of chemistry, including topics of chemistry; chemical education research; introductory chemistry; eye-tracking; learning strategies; education. In addition, the 10 papers published by Computers & Education have resulted in 331 citations, which are far more than that of other journals.

Among, Computers & Education is a top 3 journal in education technology field and leads the research progress and methods of technology-integrated education. The papers of Computers & Education mostly explore the impact of teachers' behaviors on students' learning in online education (Pi et al., 2020), visual attention of students when solving multi-choice questions (Tsai et al. 2012),

capability of students in selecting Internet information (Tsai, & Wu, 2021), impact of audio support on the multi-media learning process and results of students with reading or speaking difficulties (Knoop-van Campen et al. 2020), guiding role of video modeling on learning optimization (van Gog et al. 2014), effective realization of video annotations (Mu, 2010), and impact of tactile feedback on students' learning (Wiebe et al.2009). In addition, the papers published by Journal of Chemical Education, which ranks second, mostly study the changes or influences in the cognitive process (Tang & Pienta, 2012) or visual behaviors (Rodemor et al.,2020) of students towards related factors in the teaching process of chemistry, capability of high school students in solving problems with periodic table of elements (Tothova et al.2021), and benefit of simulation experiments on students' understanding of related concepts (Vanden Plas et.al, 2021).

Table 1. Number of published papers and citations of different journals.

Ranking	Journals	Documents	TC
1	Computers & Education	10	331
2	Journal of Chemical Education	10	148
3	Frontiers in Psychology	7	18
4	Journal of Computer Assisted Learning	7	68
5	Journal of Eye Movement Research	7	61
6	Sustainability	6	14
7	Anatomical Sciences Education	5	50
8	Computers in Human Behavior	5	72
9	Learning and Instruction	5	37
10	Instructional Science	4	23

TC: total citations

3.3. Corresponding co-author network.

Figure 2 shows the corresponding co-author network. As seen, the co-author network is formed into one cluster, including the papers published by Jarodzka H and Van Merrienboer JJG on computers in human behavior, instructional science , advances in health sciences education. These studies focus on application of eye movement in medical health and teaching education. They are performed via tools, such as eye trackers. Specifically, the studies include: exploring the influence of pause effect on learners' cognitive load through medical simulation (Lee JY et al, 2020), investigating students' self-adjustment in the selective process of learning tasks through eye movement tracking (Nugteren, ML et al, 2018), correlation among system observations, coverage and diagnostic performance through eye movement tracking in radiology (Kok, EM et al, 2016). The paper published by Jarodzka H and Van Gog T on Learning and Instruction investigates the ambiguity and priori knowledge of speech interpretation from eye movement modeling, and discusses its effective role in addressing geometry problems by EMME.

Figure 2. Analysis corresponding co-author network.

Regarding authors' citations, based on WOS, it is concluded that one paper has a higher citation (more than 100 citations, see Table 2). This paper focuses on exploring the relationship among learners' visual search patterns, information anxiety and OIPS task performance under the Internet search environment through eye movement tracking technique. It discusses the application of data science in education and derives the data results based on eye movement tracking technique and experimental design. The top three papers with the highest citations are Tsai (2012), Richstone (2010), van den Bogert (2014), respectively cited by 135 papers, 96 papers, and 68 papers. The paper by Richstone (2021), which ranks second, investigates the relationship between complex eye movement and pupil movement of non-expert surgeons and expert surgeons in simulated surgical environment through eye movement tracking. Specifically, linear discriminant analysis (LDA) and non-linear neural network analysis (NNA) are adopted. The paper by van den Bogert N (2014), which ranks third, explores teachers' visual cognition and detection of classroom events through eye movement tracking. It discusses related clues on how teachers quickly and accurately react in classroom. The paper by Kok et al. (2016), which ranks fourth, investigates the relationship among system observations, coverage, and diagnostic performance in radiology through eye movement tracking. The paper by Halszka et al. (2017), which ranks fifth, discusses the eye movement tracking in education science and summarizes the major educational theories which succeed in applying eye movement tracking based on literature review.

3.4. Analysis of co-country network

According to the Co-country Network, from 2009 to 2021, 397 research institutes from 38 countries/regions have performed related studies in this field, and 9 countries/regions have published at least 10 papers. The number of published papers of the United States is the highest. Scholars from the United States are involved in 27.15% of related published papers (60 papers) in this field, which are cited by 801 other papers. Germany and Netherlands are the second most productive countries/regions, each with 323 citations and 482 citations. Scholars from the People's Republic of China are involved in 9.95% of related published papers (22 papers) and ranks fourth, which are cited by 67 other papers. The next in order are Canada (17 papers), England (14 papers), Taiwan (12 papers), Spain (12 papers) and Sweden (11 papers).

3.5. Co-occurrence Analysis: Keywords Network

We drew a keyword co-occurrence graph on 221 papers (in Figure3) and visualized the major keywords which occur for no fewer than three times. According to the clustering results, when the node

is larger, the corresponding keyword occurs in more papers, which means that they can better represent the research hotspots; the node connection represents the intensity of correlation. When the connection is thicker, the two nodes co-occur more in the same paper; the node colors represent different clusters, namely research topic.

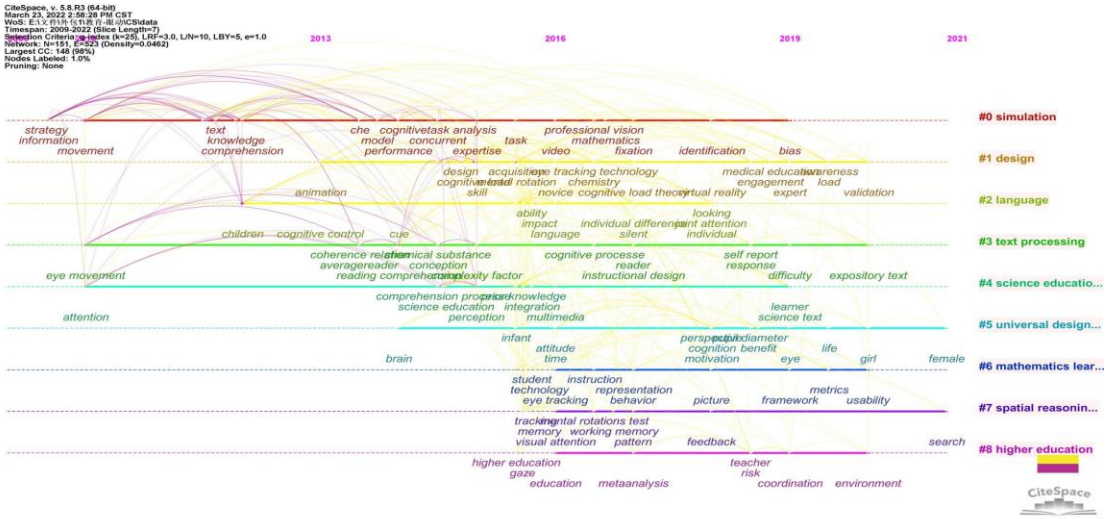


Figure 3. Timeline of the eye tracking cited network.

Cluster 0 reflects the hotspot keywords in simulation field, including movement, performance, comprehension, knowledge, information, strategy, model, expertise, text, fixation, video, task, online, and mathematics. These hotspot keywords reflect that Cluster 0 focuses on the simulation, visual attention, simulation performance evaluation tools, cognitive process, cognitive load, thinking hypothesis, education, model, fixation, and theoretical hypotheses to perform comprehension based eye movement experiments. For example, some studies investigate how designated experts (experienced in science education and conceptual mapping) and non-experts (pre-service teachers) build their conceptual map when considering their cognitive process (Dogusoy-Taylan & Cagiltay, 2014); some studies investigate the verification of ETG technology as an objective evaluation method (Shinnick, 2016); some studies discuss possible contributions of eyeball tracking in educational neuroscience field (Soares et.al, 2021). The researchers discussed the ET videos recorded when teachers solve mathematics questions for students and performed qualitative analysis on the results to determine the fixation strategy adopted by students. Papers of Cluster 0 focus on applying eye movement tracking to perform simulated cognitive attention experiments in the field of disciplinary education and neuroscience education under the guidance of mathematical model or theoretical hypothesis Cluster 1 reflects the hotspot keywords in the field of design, including skill, cognitive load, design, novice, animation, acquisition, expert, mental rotation, medical education, cognitive load theory, virtual reality, validation, complexity, and engagement. It is found through keyword search that this cluster mainly focuses on the building of virtual tools in different fields (Thomas & Franklin, 2011) to provide progressive help in this field. Cluster 2 reflects the hotspot keywords in the language field, including children, language, impact, ability, quality, fluency, vocabulary, recognition, and cue. The papers related to these keywords study the language learning ability of children, similarities and differences of low-skilled adult readers and child readers (Barnes & Kim, 2016), differences in the reading skill of students at different ages (Rooijackers, 2020), and impact of television interaction enhancement on children's visual attention and vocabulary learning (Nussenbaum & Amso, 2016). The paper titled "Low-skilled adult readers look like typically developing child readers: a comparison of reading skills and eye movement behavior" includes five keywords, which

are language, ability, quality, vocabulary, fluency, and is the most correlated in Cluster 2. 25 low-skilled adult readers and 25 first-graders were selected as research subjects in this paper to examine their similarities and differences in reading skill and eye movement. Cluster 3 reflects the hotspot keywords in text processing field, including eye movement, reading comprehension, difficulty, reader, instructional design, and complexity factor. Different from other clusters, Cluster 3 only focuses on text processing and addresses the following mainly issues: helping to reduce reading comprehension for readers whose second language is English, comparing the complexity factor of papers from different categories on reading comprehension through eye movement (Rets & Rogaten, 2021). Cluster 4 reflects the hotspot keywords in science education field, including attention, perception, multimedia, integration, science education, prior knowledge, and comprehension process. Its difference from the previous clusters lies in that these keywords focus on related topics of science education and the effect of science education is evaluated through eye movement experiments. For example, regarding the health harm of smoking, researchers analyze smokers and non-smokers through 14 types of health warning signs. Through eye movement experiment and questionnaire survey, the research findings conclude that effective warning on the cigarette package can effectively make people aware of the health harm from smoking (Gercek et al., 2016). Cluster 5 reflects the hotspot keywords in universal design for learning field and focuses on learning design. For example, the research of Cluster 5 studies students' learning outcomes in different learning environments, such as teachers' interactive style and familiarity, and discusses the differences in students' learning outcomes in curricular environment through different communication skills at different stages (Lui et al., 2021). The research of Cluster 5 can effectively help people to understand the differences in students' learning outcomes under different teaching designs. In science education course, researchers have discussed that the multi-media content and pictures are very essential for design of mathematics textbooks (Ogren et al., 2017). Some researchers have adopted eye movement technique to explore the performance of students from science major and non-science major in online science evidence-based reasoning tasks. It has been found that students from science majors perform far better than students from non-science major in terms of reasoning capability. In addition, students from science majors demonstrated significantly more eye movement during the first science evidence-based reasoning task (Tsai et al., 2019). On the whole, Cluster 5 includes methods which explore eye movement experiments in science education. Cluster 6 reflects the hotspot keywords in mathematics learning field. As shown by the keywords, Cluster 6 focuses on analysis of eye movement experiments and helps students learn at a faster and more efficient rate through effective training of geometry and graphics (Molina et al., 2018). Cluster 7 reflects the hotspot keywords in spatial reasoning field and focuses on the anatomy research. Different from general geometric math, anatomy tend to be more challenging in terms of both space and vision. Therefore, apart from general teaching, Cluster 7 further refines the research tracks and includes the application of eye movement technique in aiding the academic progress of related fields (Roach et al., 2017). Cluster 8 reflects the hotspot keywords in higher education field, including education, gaze, higher education, meta-analysis, environment, and coordination. Compared with Clusters 6 and 7, Cluster 8 is more diverse. Apart from teaching quality evaluation (Coskun & Cagiltay, 2021), Cluster 8 focuses on helping students better solve graphic tasks (Brückner et al., 2020).

4. Results and Discussion

Researchers is the key to the progress of research on related topics. 23 scholars have published more than 3 papers, accounting for 2.87% of all the scholars. The most productive scholar, Jarodzka H, has published 8 papers. Regarding citations, the 8 papers published by Jarodzka H have been cited by 214

papers, which is the highest among all the author retrieved. It can be seen that Jarodzka H has played a significant role in promoting the research progress of this field. Meanwhile, productive authors make great contributions to the progress of related education fields. Besides, the authors with high citations are basically from the research institutes and countries with high citations while the USA occupies a significantly leading position in this field. Moreover, concerning keywords, the research tracks of education fields include higher education and science education. In recent years, the research on learning behaviors has become widely popular. As CiteSpace can only present the keyword nodes in a qualitative manner, this study is subject to certain limitations. Quantitatively speaking, the value represented by the keyword nodes originates from their number of occurrences in the whole literature dataset. If CiteSpace can categorize keywords and authorship order based on their number of occurrences in specific literature, this study will be more convincing.

References

Aryadoust, V., & Ang, B. H. (2021). Exploring the frontiers of eye tracking research in language studies: a novel co-citation scientometric review. Computer Assisted Language Learning, 34(7), 898-933.

Barnes, A. E., & Kim, Y. S. (2016). Low-skilled adult readers look like typically developing child readers: a comparison of reading skills and eye movement behavior. Reading and Writing, 29(9), 1889-1914.

Brinton Anderson, B., Vance, A., Kirwan, C. B., Eargle, D., & Jenkins, J. L. (2016). How users perceive and respond to security messages: a NeuroIS research agenda and empirical study. European Journal of Information Systems, 25(4), 364-390.

Brückner, S., Schneider, J., Zlatkin-Troitschanskaia, O., & Drachsler, H. (2020). Epistemic Network Analyses of Economics Students' Graph Understanding: An Eye-Tracking Study. Sensors, 20(23), 6908.

Coskun, A., & Cagiltay, K. (2021). Investigation of classroom management skills by using eye-tracking technology. Education and Information Technologies, 26(3), 2501-2522.

Dogusoy-Taylan, B. and K. Cagiltay (2014). Cognitive analysis of experts' and novices' concept mapping processes: An eye tracking study. Computers in Human Behavior 36: 82-93.

Gercek, C., Ozcan, O., Ocak, N., Ferhat, C., Berberoglu, S., Cakir, E., . . . Cagiltay, K. (2016). An evaluation of the effects of combined health warnings on cigarette lepapackets through eye movements. Journal of Baltic Science Education, 15(6), 680-692.

Knoop-van Campen, C. A. N., Segers, E., & Verhoeven, L. (2020). Effects of audio support on multimedia learning processes and outcomes in students with dyslexia. Computers & Education, 150. doi:10.1016/j.compedu.2020.103858

Lo, C. K., & Hew, K. F. (2021). Student engagement in mathematics flipped classrooms: Implications of journal publications from 2011 to 2020. Frontiers in Psychology, 12.

Lui, M., Maharaj, A., Shimaly, R., Atcha, A., Ali, H., Carroll, S., & McEwen, R. (2021). User Experiences of Eye Gaze Classroom Technology for Children With Complex Communication Needs. Journal of Special Education Technology, 01626434211019399.

Molina, A. I., Navarro, Ó., Ortega, M., & Lacruz, M. (2018). Evaluating multimedia learning materials in primary education using eye tracking. Computer Standards & Interfaces, 59, 45-60.

Mu, X. M. (2010). Towards effective video annotation: An approach to automatically link notes with video content. Computers & Education, 55(4), 1752-1763.

Nussenbaum, K., & Amso, D. (2016). An Attentional Goldilocks Effect: An Optimal Amount of Social Interactivity Promotes Word Learning From Video. Journal of Cognition and Development, 17(1), 30-40.

Ogren, M., Nystrom, M., & Jarodzka, H. (2017). There's more to the multimedia effect than meets the eye: is seeing pictures believing? Instructional Science, 45(2), 263-287.

Pi, Z. L., Xu, K., Liu, C. X., & Yang, J. M. (2020). Instructor presence in video lectures: Eye gaze matters, but not body orientation. Computers & Education, 144.

Rets, I., & Rogaten, J. (2021). To simplify or not? Facilitating English L2 users' comprehension and processing of open educational resources in English using text simplification. Journal of Computer Assisted Learning, 37(3), 705-717.

Roach, V. A., Fraser, G. M., Kryklywy, J. H., Mitchell, D. G., & Wilson, T. D. (2017). Different perspectives: Spatial ability influences where individuals look on a timed spatial test. Anatomical sciences education, 10(3), 224-234.

Rooijackers, P., van Silfhout, G., Schuurs, U., Mulders, L., & van den Bergh, H. (2020). Reading texts and answering questions; a cross-sectional eye-tracking study among preacademic students. Pedagogische Studien, 97(4), 281-308.

Segijn, C. M., Voorveld, H. A., & Vakeel, K. A. (2021). The Role of Ad Sequence and Privacy Concerns in Personalized Advertising: An Eye-Tracking Study into Synced Advertising Effects. Journal of Advertising, 50(3), 320-329.

Thomas, N. P., Crow, S. R., & Franklin, L. L. (2011). Information literacy and information skills instruction: Applying research to practice in the 21st century school library. ABC-CLIO.

Tothova, M., Rusek, M., & Chytry, V. (2021). Students' Procedure When Solving Problem Tasks Based on the Periodic Table: An Eye-Tracking Study. Journal of Chemical Education, 98(6), 1831-1840.

Tsai, M. J., & Wu, A. H. (2021). Visual search patterns, information selection strategies, and information anxiety for online information problem solving. Computers & Education, 172. doi:10.1016/j.compedu.2021.104236

Tsai, M. J., Hou, H. T., Lai, M. L., Liu, W. Y., & Yang, F. Y. (2012). Visual attention for solving multiple-choice science problem: An eye-tracking analysis. Computers & Education, 58(1), 375-385. doi:10.1016/j.compedu.2011.07.012

Tsai, P. Y., Yang, T. T., She, H. C., & Chen, S. C. (2019). Leveraging College Students' Scientific Evidence-Based Reasoning Performance with Eye-Tracking-Supported Metacognition. Journal of Science Education and Technology, 28(6), 613-627.

van Gog, T., Verveer, I., & Verveer, L. (2014). Learning from video modeling examples: Effects of seeing the human model's face. Computers & Education, 72, 323-327.

Wiebe, E. N., Minogue, J., Jones, M. G., Cowley, J., & Krebs, D. (2009). Haptic feedback and students' learning about levers: Unraveling the effect of simulated touch. Computers & Education, 53(3), 667-676.

Zammarchi, G., & Conversano, C. (2021). Application of Eye Tracking Technology in Medicine: A Bibliometric Analysis. Vision, 5(4), 56.

Shih, J. L., Jiang, B., Lee, M. H., Yin, C. J., Sun, D. E., Lu, Y. (Eds.) (2023). Main Conference Proceedings (English Paper) of the 27th Global Chinese Conference on Computers in Education (GCCCE 2023). China: Beijing Normal University.

The Effects of Individual Preparation on Students' Collaborative Argumentation-Based Learning: A Exploratory Study in a Secondary School Classroom

Wenli Chen1, Junzhu Su1*, Qianru Lyu1, Xinyi Li2, Siew Cheng Aileen Chai1, Guo Su and Eng Eng Ng1

1 Nanyang Technological University, Singapore

2Beijing Normal University

* nie21.sj8277@e.ntu.edu.sg

Abstract: *Computer-supported collaboration is considered to be one effective strategy to improve students' argumentation skills. The integration of an individual preparation (IP) activity before collaboration activity has been applied in collaboration script studies. Few studies examined the role of IP activity in argumentation skill improvement via online collaboration in face-to-face classrooms. In the context of computer-supported collaborative argumentation (CSCA) activities with a graph-based argumentation platform, this study explored the effects of IP on students' collaborative argumentation-based learning in three conditions: 10-minute IP before collaboration condition (condition 1), 5-minute IP before collaboration condition (condition 2), and immediate collaboration condition without IP (control condition). Students (N= 39) from one class in one Singapore secondary school participated in three face-to-face computer-supported collaborative learning lessons. The results showed a higher quality of collaborative argumentation in the immediate collaboration condition than in the conditions with IP, which means that more time should be allocated to collaborative activities instead of IP in similar CSCA contexts. These findings contribute to the current understanding of the role of IP activity before collaboration on students' collaboration process as well as outcomes.*

Keywords: Collaborative Argumentation, Individual Preparation, Secondary School, CSCL

1. Introduction

Computer-Supported Collaborative Learning (CSCL) technologies that provide scripting or scaffolding support aim to maximise students' opportunities for learning and develop effective collaboration strategies (Fischer et al. 2007). However, technology does not guarantee successful collaboration in real classrooms (Okojie, 2020). One instructional approach to support high-quality collaborative learning is the collaboration scripts (Kobbe et al., 2007; Weinberger, 2008; Weinberger et al., 2009), which was also known as computer-supported collaboration scripts ((Dillenbourg, 2002; Rummel et al., 2009). To close the gaps between theory and practice about collaboration scripts in CSCL and to extend theories on learning (Barab and Squire 2004), the researchers used the design-based research approach (DBR) in this study to investigate the effect of IP in a CSCA by comparing students' learning outcomes in three 20-minute activities. DBR was defined as a research method or related methodology which used a blended form of design activities and research activities to produce design-relevant, empirically supported knowledge (Hoadley, 2002). DBR methods are an important cornerstone

in the methodological repertoire of the learning sciences, and they play a particularly important role in CSCL research and development (Kali, 2021). The activities include 10-minute individual ideation before collaboration condition (condition 1), 5-minute individual ideation before collaboration condition (condition 2), and immediate collaboration condition (control condition). 10-minute individual ideation was designed in lesson 1 to help students generate more perspectives and ideas in the argumentation context. When the researchers and the teacher found the learning outcomes were not as ideal as expected, the time duration of individual ideation was reduced in lesson 2 and removed in lesson 3.

2. Literature Review

To improve students' argumentation skills and content knowledge learning, CSCA was designed to support sharing, constructing, and representing arguments in multiple formats and was found to be effective in improving students' content knowledge and developing their argumentation skills (Hsu et al., 2015), critical and elaborative discussions (Scheuer et al., 2014), domain-specific knowledge (Stegmann et al., 2012), in-depth and productive argumentative learning (Kirschner et al., 2002). In addition, computer-supported learning environments help students collect evidence for argumentation and access data to structure their argumentation skills (Clark et al., 2008).

As one effective approach to support CSCA, collaboration scripts were defined as "sets of coordinated scaffolds that specify and sequence individual and collaborative learning activities and thereby aim to facilitate knowledge acquisition" (Stegmann et al., 2012, p. 301). Collaboration scripts could guide the learners through a sequence of interaction phases with designated activities and roles (O'Donnell, 1999) and promote particular conducive cognitive, social, and metacognitive processes (King et al., 2007). Students are provided with specific and explicit guidelines in terms of what, when, and by whom certain learning activities should be executed (Weinberger et al., 2007). The main components of collaboration scripts were identified: participants, activities, roles, resources, and groups (Kobbe et al., 2007). The mechanisms of collaboration scripts include how are activities, roles, and resources distributed across participants (task distribution), how are participants distributed across groups, and how are task distribution and groups distributed over a time sequencing (Kobbe et al., 2007). CSCL scripts engage learners in beneficial collaborative practices, eventually supporting the development of collaboration skills and domain learning (Vogel et al., 2021).

One of the CSCL scripts is IP before collaboration, which was defined as "providing learners with time to perform activities directed at processing the instructional material on their own before the collaboration" (Mende et al., 2021, p. 30). IP may allow students to activate relevant knowledge and to develop some initial ideas and conclusions with regard to the task before having to invest efforts in communicating and coordinating with others (Mende et al., 2021). Previous studies have explored the impact of IP before collaboration in various contexts (e.g. Farrokhnia et al., 2019; Van Boxtel et al., 2000; Erkens et al., 2016) and reveal mixed results of IP for collaboration: it sometimes has positive (Lam & Kapur, 2018; Weinberger, 2011), no, or even negative effects (Mende et al., 2020; Tsovaltzi, Judele, et al., 2015; Tsovaltzi et al., 2017). Thus, it's the potential benefits on performance during subsequent collaboration should be considered when instructors design an IP phase (Mende et al., 2021).

This study proposed to conduct collaborative argumentation in a synchronous CSCA context to detect the effect of IP on collaborative learning. The current study aims to address the following two questions:

1. What is the effect of individual preparation before collaboration on students' collaborative argumentation outcomes?

2. What is the effect of individual preparation with different time duration before collaboration on students' collaborative argumentation outcomes?

3. Methods

A time-series design was applied to examine the effect of IP before collaboration on students' collaborative argumentation that took place on an online platform in one class from a secondary school.

3.1. Participants

A total of 39 Singapore secondary school students participated in this study. All the students were female students in Grade 9, aged between 14 to 15 at the time of this study. They also had the experience of collaborative discussions on the Appletree platform (Authors, 2013). Students were assigned into groups of four or five and received tool training before the lessons. The students in each group knew each other and had prior group learning experiences but not in the same group settings as in this study.

3.2. Design and Procedure

Three collaborative argumentation activities were co-designed by the teacher from the secondary school and the researchers in this study (Table 1). Three topics, including "*Income inequality is worsening in Singapore. Do you agree?*", "*Foreigners are not welcomed in Singapore. Do you agree?", "The Singapore government's efforts have been effective in managing racial and religious tensions. Do you agree?"* were discussed in collaborative argumentation activities as they were controversial social issues that encourage critical conversation and multiple perspectives within group work, with similar difficulty levels. The class separately conducted three 20-minute argumentation activities across three weeks during a period of 41 days with around 20 days in between. During the collaborative argumentation ideation activity, each group discussed and co-constructed their group argumentation in the shared online working space called AppleTree (Authors, 2013). Students sat together with their group members and work on an online collaborative argumentation platform (Appletree. sg) with their individual tablets. In lesson 1, students were asked to write individual ideas for 10 minutes first before they continued to write group ideas collaboratively for another 10 minutes. In lesson 2, students were asked to write individual ideas for 5 minutes first before they continued to write group ideas collaboratively for another 15 minutes. At last, students were asked to write group ideas collaboratively for another 20 minutes immediately in lesson 3. For each lesson, the teachers briefed the argumentation task before students were assigned to the argumentation activities, during which they read into their argumentation, discussed within the group, and wrote their arguments.

Table 1. Overview of the procedure of the study.

Lessons	Topics	Activities	Time
Lesson 1	Income inequality is worsening in Singapore." Do you agree?	Individual Ideation	10 mins
		Collaboration	10 mins
Lesson 2	Foreigners are not welcomed in Singapore." Do you agree?	Individual Ideation	5 mins
		Collaboration	15 mins
Lesson 3	The Singapore government's efforts have been effective in managing racial and religious tensions." Do you agree?	Collaboration	20 mins

3.3. CSCA System

The AppleTree system is a graph-based technological platform that allows users to externalize their thought processes in a mind map structure (Authors, 2013). Figure 1 shows the user interface of the AppleTree system. It was designed for supporting generalized coordination of collaborative argumentation among students and the teacher in the following aspects: (1) argumentation: developing graph-based argumentation to represent argument elements, and relationships between them, in which a cloud represents an idea, an ellipse represents a claim, and a rectangle represents a piece of evidence; (2) collaboration: scaffolding student's continuous knowledge improvement through a phased collaboration. The AppleTree system is a graph-based technological platform that allows users to externalize their thought processes in a mind map structure. When students worked on the argumentative diagrams on the Appletree system, all the data were stored automatically in the back-end database of the Appletree System.

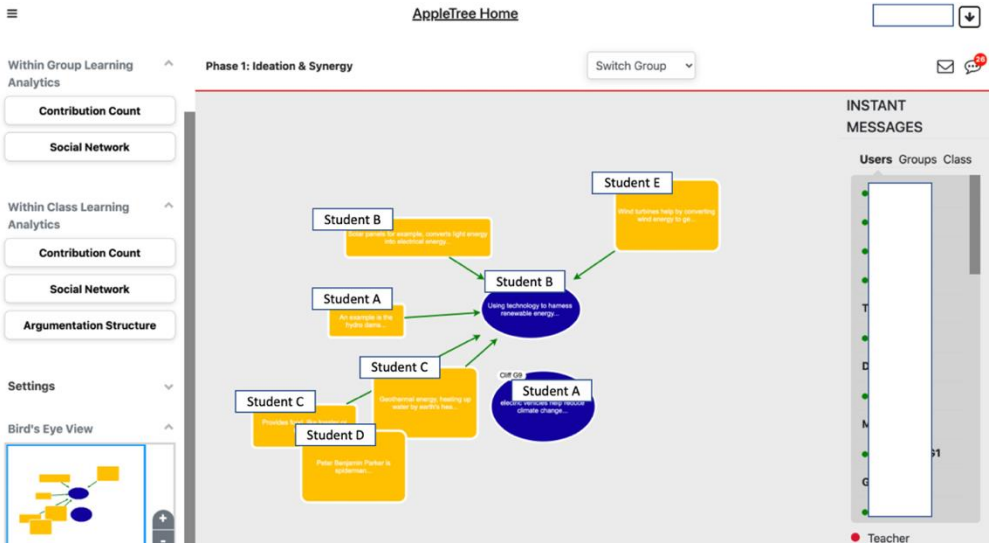


Figure 1. Appletree user interface.

4. Data Collection and Analysis

Data collected for analysis included the argumentation diagrams contributed by students on the AppleTree system in the three lessons. To answer research questions 1 and 2, groups' (N =8) argumentative diagrams were collected through screen captures and system reports to identify the quality of their learning outcomes in different conditions. Class observation videos were also collected to identify the exact time that students spent on each activity.

In this study, the unit of analysis was the collaborative argumentation diagrams in each group's digital space. The researchers coded each group's argumentation diagrams in three conditions from 0 points to 3 points in four dimensions, including Clarity dimension (Andrade et al., 2010; Marzano et al., 2000), Perspectives dimension (Stapleton & Wu, 2015), Evidence dimension (Toulmin, 2003), Elaboration dimension (Marzano et al., 2000). These four dimensions were chosen considering their alignment with the current collaborative argumentation context. Then, all points were added together serving as the final score indicating the overall argumentation quality.

Two independent coders coded 11.1% of all the graphs using the coding scheme mentioned above. The Cronbach alpha for the overall dimensions was .844. The Cronbach alpha for the Clarity dimension was .625. The Cronbach alpha for the Perspectives dimension was .875. The Cronbach alpha

for the Evidence dimension was 1. The Cronbach alpha for the Elaboration dimension was .75. These results indicated satisfied agreement between the coders.

Table 2. The coding scheme of argumentation quality.

Dimensions	0	1	2	3
Clarity	The central idea and clarity of purpose are absent.	The central idea is basically expressed but it is vague. There are many repetitive ideas. Readers have difficulties understanding the argumentation.	The central idea and the clarity of purpose are generally evident. A few repetitive ideas. Readers occasionally have confusions.	The central idea and the clarity of purpose are well developed; no repetitive ideas, very compelling statements. Readers have a clear understanding of the argumentation.
Perspectives	No consideration of multiple perspectives; No stakeholder considered.	Insufficiently explore multiple perspectives Consider limited type of stakeholders.	Display multiple perspectives though not consistently done Consider many type of stakeholders but not systematically.	Synthesizing multiple perspectives while valuing contributions from these perspectives Consider variety of stakeholders systematically.
Evidence	No evidence at all or evidence provided is not credible or relevant.	Limited evidence provided. Evidence provided is somewhat credible and relevant but not supporting the multiple perspectives.	Much evidence has been provided. Evidence provided is credible and relevant and supporting multiple perspectives, though not consistently done.	A rich variety of credible and relevant evidence which consistently supports multiple perspectives.
Elaboration	Irrelevant ideas, or ideas with no elaboration.	Little or superficial development of claims and evidence; lack of explanation and elaboration.	Shows development of claims and evidence, with explanation and elaboration; demonstrate in-depth thinking process such as synthesizing, reflecting, evaluating.	Shows an extensive development of claims and evidence with with thorough explanation and elaboration, Demonstrate in-depth thinking processes such as synthesizing, reflecting, evaluating.

5. Results

Research question 1: What is the effect of individual preparation before collaboration on students' collaborative argumentation outcomes?

The quality of students' argumentative diagrams in lesson 3 was scored using the above-mentioned coding scheme and a paired samples t-test was conducted to address the second research question. The sum score of Lesson 3 (M=10, SD= 0.9636) is significantly higher than Lesson 2(M=8.5, SD= 0.9626, t (8) =-2.542, p=0.039). To further understand students' performance in clarity, perspectives, evidence, and elaboration dimensions, paired samples t-test was conducted in each dimension between the two lessons see Table 4. The clarity dimension was significantly higher in Lesson 3(M = 2.88, SD = 0.354) than in Lesson 2(M=2.125, SD= 0.2315), t (8) = -3.969, p=0.005) Significant improvement was also identified in the perspectives dimension in Lesson 3 (M = 2.88, SD = 0.354) than in Lesson 2 (M = 2.188, SD = 0.372, t (8)=-3.274, p=0.014). The evidence dimension was significantly higher in Lesson 3 (M= 2.188, SD=0.2588) than in Lesson 2 (M=2.5, SD= 0.2673), t (8) = 1.93, p=0.095). However, the elaboration dimension was significantly higher in Lesson 3 (M = 2.063, SD = 0.4955) than in Lesson 2 (M = 1.688. SD = 0.372, t (8)= 0.357, p=0.732). This result suggests that students in the immediate collaboration condition performed even better in collaborative argumentation than in the IP before the collaboration condition.

Table 3. Comparison of argumentation diagrams in lesson 2 and lesson 3.

	Lesson 2		Lesson 3			
Dimensions	Mean	SD	Mean	SD	t	p
Clarity	2.125	0.2315	2.88	0.354	-3.969	0.005*
Perspectives	2.188	0.372	2.88	0.354	-3.274	0.014*
Evidence	2.5	0.2673	2.188	0.2588	1.93	0.095
Elaboration	1.688	0.372	2.063	0.4955	0.357	0.732
Sum	8.5	0.9626	10	0.9636	-2.542	0.039*

Research question 2: What is the effect of individual preparation with different time duration before collaboration on students' collaborative argumentation outcomes?

The quality of students' argumentative diagrams in lessons 1 (10-minute IP and 10-minute collaboration) and lesson 2 (5-minute IP and 15-minute collaboration) was scored using the above-mentioned coding scheme and a paired samples t-test between the results of lesson 1 and lesson 2 was conducted to address the first research question. The sum score of Lesson 2 (M=7.500, SD= 1.604), is not significantly higher than Lesson 1 (M=8.500, SD= 0.963 t (8)=- -39.000, p=0.081). To further understand students' performance in clarity, perspectives, evidence, and elaboration dimensions, paired samples t-test was conducted in each dimension between the two lessons (see Table 3). The clarity dimension was not significantly higher in Lesson 2 (M=7.500, SD= 1.604) than in Lesson 1 (M=1.380, SD= 1.061), t (8) = -2.291, p=0.056). An insignificant difference was also identified in the perspectives dimension in Lesson 2 (M = 2.188, SD = 0.372) than in Lesson 1 (M = 2.000, SD = 0.535, t (8)=-4.064, p=0.351). The evidence dimension was significantly higher in Lesson 2 (M= 2.500, SD=0.267) than in Lesson 1 (M=2.130, SD= 0.354), t (8) = -4.583, p=0.003). However, the elaboration dimension was significantly higher in Lesson 1 (M = 1.688, SD = 0.372) than in Lesson 2 (M = 2.000. SD = 0.000, t

(8)=- 2.376, p=0.049). This result suggests that the time duration of IP did not play an important role in promoting collaborative argumentation.

Table 4. Comparison of argumentation diagrams in lesson 1 and lesson 2.

	Lesson 1		Lesson 2			
Dimensions	Mean	SD	Mean	SD	t	p
Clarity	1.380	1.061	2.125	0.232	-2.291	0.056
Perspectives	2.000	0.535	2.188	0.372	-1.000	0.351
Evidence	2.130	0.354	2.500	0.267	-4.583	0.003*
Elaboration	2.000	0.000	1.688	0.372	2.376	0.049*
Sum	7.500	1.604	8.500	0.963	-39.000	0.081

6. Discussion and Conclusion

In this study, students who started collaboration immediately performed better than students who prepared arguments individually before collaboration in terms of collaborative argumentation outcomes; students who spent less time on IP performed better than students who spent more time on IP in terms of collaborative argumentation outcomes. Different from some studies in the literature (Jermann & Dillenbourg, 2003; Olsen, 2017; Rummel & Spada, 2005), the findings in this study resemble the findings of the previous research studies that identified the negative impacts of IP on the subsequent collaborative argumentation activities (Tsovaltzi, Judele, et al., 2015; Tsovaltzi et al., 2017) due to the fact that students were found to focus on self-presentation instead of building on each other's ideas.

The results shed light on previous inconclusive findings and contradictory theoretical views regarding IP before collaboration. The study revisits the issue of whether additional collaborative processes can be considered from the perspective of individual process losses (e.g., Marttunen & Laurinen, 2001; Weinberger et al., 2010). Following social constructivism, collaborative processes may actually be an important constituent of collaborative learning and should be given enough time.

To summarize, the study investigated the question of how much IP before collaboration is needed for argumentative learning. The study showed that IP before the collaboration condition may not always bring better learning outcomes. At the same time, this study highlighted the need to explore the effect of the IP script in CSCA contexts more deeply.

Some limitations existed in this exploratory study. Firstly, the time duration is limited in this study, the findings should be tested with longer time. Secondly, as prior studies found differences between females and males in terms of the quality of their argumentative feedback (Noroozi et al., 2020), this study is contextualized in a secondary girls' school in Singapore and all participants were female schoolers, which makes the finding less generalizable for secondary schoolers. Future research should test the transfer of these results in larger groups and field studies.

References

Andrade, H. L., Du, Y., & Mycek, K. (2010). Rubric-referenced self-assessment and middle school students' writing. *Assessment in Education: Principles, Policy & Practice, 17*(2), 199-214.

Clark, D., Stegmann, K., Weinberger, A., Menekse, M., & Erkens, G. (2008). Argumentation in Science Education. Perspectives from Classroom-Based Research. In.

Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. *Three worlds of CSCL. Can we support CSCL*, *61491*.

Fischer, F., Kollar, I., Mandl, H., & Haake, J. M. (2007). Scripting computer-supported collaborative learning: Cognitive, computational and educational perspectives (Vol. 6). New York: Springer.

Hoadley, C. (2002). Creating context: Design-based research in creating and understanding CSCL. In G. Stahl (Ed.), Computer Support for Collaborative Learning 2002 (pp. 453–462). Erlbaum.

Hsu, P. S., Van Dyke, M., Chen, Y., & Smith, T. (2015). The effect of a graph-oriented computer-assisted project-based learning environment on argumentation skills. *Journal of Computer Assisted Learning*, *31*(1), 32-58.

Kali, Y., & Hoadley, C. (2021). Design-based research methods in CSCL: Calibrating our epistemologies and ontologies. *International Handbook of Computer-Supported Collaborative Learning*, 479-496.

Kirschner, P. A., Buckingham-Shum, S. J., & Carr, C. S. (2002). *Visualizing argumentation: Software tools for collaborative and educational sense-making*. Springer Science & Business Media.

Kobbe, L., Weinberger, A., Dillenbourg, P., Harrer, A., Hämäläinen, R., Häkkinen, P., & Fischer, F. (2007). Specifying computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning*, *2*(2), 211-224.

Lam, R., & Kapur, M. (2018). Preparation for future collaboration: Cognitively preparing for learning from collaboration. *The Journal of Experimental Education*, *86*(4), 546-559.

Marzano, R. J., Gaddy, B. B., & Dean, C. (2000). What Works in Classroom Instruction.

Mende, S., Proske, A., & Narciss, S. (2021). Individually preparing learners to perform interactive activities in CSCL: do generative tasks work? Proceedings of the 14th International Conference on Computer-Supported Collaborative Learning-CSCL 2021,

Okojie, M. C. (2020). *Handbook of Research on Adult Learning in Higher Education*. IGI Global.

Rummel, N., Spada, H., & Hauser, S. (2009). Learning to collaborate while being scripted or by observing a model. *International Journal of Computer-Supported Collaborative Learning*, *4*(1), 69-92.

Scheuer, O., McLaren, B. M., Weinberger, A., & Niebuhr, S. (2014). Promoting critical, elaborative discussions through a collaboration script and argument diagrams. *Instructional science*, *42*(2), 127-157.

Stapleton, P., & Wu, Y. A. (2015). Assessing the quality of arguments in students' persuasive writing: A case study analyzing the relationship between surface structure and substance. *Journal of English for Academic Purposes*, *17*, 12-23.

Stegmann, K., Wecker, C., Weinberger, A., & Fischer, F. (2012). Collaborative argumentation and cognitive elaboration in a computer-supported collaborative learning environment. *Instructional science*, *40*(2), 297-323.

Vogel, F., Weinberger, A., & Fischer, F. (2021). Collaboration scripts: Guiding, internalizing, and adapting. In *International Handbook of Computer-Supported Collaborative Learning* (pp. 335-352). Springer.

Weinberger, A. (2011). Principles of transactive computer-supported collaboration scripts. *Nordic Journal of Digital Literacy*, *6*(3), 189-202.

ISBN: 978-626-974-780-1
ISSN: 3005-3226

GCCCE 2023

主办方：全球华人计算机教育应用学会(GCSCE)

承办方：北京师范大学
北京师范大学未来教育高精尖创新中心

中国 | 北京