Investigating Modern Application's Impact on Multimodal Learning in Higher Education: A Case Study of Bilibili's Online Course

Yaoyang Zhou
Foreign Languages College, Guangdong University of Petrochemical Technology, Maoming, Guangdong, 525000, China

Keywords: Multimodal Interaction, Social Media, College Students.

Abstract: This study explores the impact of multimodal interaction on the learning efficiency and effectiveness of college students when they participate courses on the Bilibili platform in the context of the digital era through qualitative analysis methods and case studies. It was found that multimodal interaction functions to improve students' motivation, engagement and knowledge acquisition. Multimodal learning environments are effective, especially the pop-up commenting feature that enhances interaction and engagement, but may also make slower learners feel anxious. These findings suggest the potential of Bilibili as an educational resource. Educators and platform designers need to balance engagement and distraction to optimise multimodal learning environments to support all learners.

1 INTRODUCTION

With the advent of the digital era, social media platforms such as beeping have become an important channel for interaction and communication between learners, teachers and classmates. According to Bowen Liu's study, modality refers to the way in which information is exchanged between the user and the computer, mainly through the senses of touch, vision, hearing, smell and taste (Liu, 2024). These modalities can be classified as tactile modalities, visual modalities, auditory modalities, etc. Multimodal interaction combines multiple senses. With the development of technologies such as computer vision, artificial intelligence and gesture recognition, multimodal interaction has become increasingly important in the fields of computer science and interaction design.

Social media platforms are able to break through the constraints of time and space, facilitate cross-cultural and cross-geographical learning exchanges, and provide students with opportunities for immediate feedback and knowledge sharing. Such diverse modes of interaction, such as through video comments and real-time discussions, further enhance the effectiveness of cross-cultural communication and promote cooperative learning and the formation of learning communities (Li, 2024). Related literature research shows that social media, with its multimodal characteristics, can provide learners with a rich cross-modal interactive experience through the integrated application of text, image, audio and video, effectively stimulating students' learning motivation. However, there are fewer theories and researches related to online class pedagogy operating today, and it is difficult to teach online classes systematically (Liu & Wang, 2024).

This study adopts a qualitative analytical approach to investigate the impact of multimodal interactions on the learning efficiency and effectiveness of university students when they are engaged in online courses on a popular video-sharing platform, Bilibili. The main objective of the study was to investigate and analyse how different forms of multimodal interactions affect students' motivation, engagement, and learning incentives in the specific online environment of Bilibili (Sun et al., 2015). Through the means of interviews, and case studies, this study attempts to reveal what forms of multimodal interactions are effective in enhancing students' learning outcomes, thus providing educators with references and suggestions when designing and implementing online courses, with a view to optimising the online learning experience and improving the quality of education.

This paper will endeavour to answer the following questions through qualitative analysis (Lim, Toh, & Nguyen, 2022).

1. How can social media platforms (e.g., Bleep) enhance learners' knowledge exchange through

Zhou, Y.
Investigating Modern Application's Impact on Multimodal Learning in Higher Education: A Case Study of Bilibili's Online Course.
DOI: 10.5220/0013962500004912

In *Proceedings of the 1st International Conference on Innovative Education and Social Development (IESD 2025)*, pages 41-46
ISBN: 978-989-758-779-5

multimodal interactions in the context of the digital age?

2. in what specific ways is multimodal interaction important in the field of computer science and interaction design?

3. How do multimodal interactions on the Beep platform affect the efficiency and effectiveness of university students' learning?

2 RESEARCH METHODOLOGY

This study is dedicated to exploring the impact of multimodal interactions on college students' learning experience in a bleeping online course. In order to achieve this research objective, this study used a combination of qualitative research methods.

2.1 Semi-Structured Interviews

Firstly, semi-structured interviews were conducted with learners on the Bilibili platform. The aim was to explore their preferences for different forms of multimodal interactions and understand their feelings about these experiences. The interviews also gathered learners' subjective evaluations of their learning effectiveness. Additionally, data was collected from a broader group of students on the frequency, motivation, engagement, and learning effectiveness of using Bilibili's multimodal interactive features.

In the pre-interview period before the interviews began, the author specially selected ten teacher training students from Guangdong University of Petrochemical Technology (GDUPT) in their junior year and above as the interview subjects (see Table 1). These students not only possessed a solid foundation in educational theory, but also actively used the Bilibili platform in their daily lives to take online courses in preparation for various upcoming exams. The interviews explored in depth their learning experiences of online courses on the Bilibili platform, especially the impact of different teaching formats on learning interest, efficiency and effectiveness (Wang et al., 2023).

2.2 Case Studies

Then, through the case study method, a specific Bilibili online course was selected, and the role of multimodal elements such as video content and student interactions (including pop-ups and comments) in the actual learning situation was carefully observed and recorded. Through the combined use of these methods, this study aims to provide a comprehensive understanding of how multimodal interactions affect college students' learning experiences on the Bilibili platform.

Table 1: Basic information about the interviewees.

Case number	Distinguishing between the sexes	(a person's) Age	Professions	Learning situation
P1	women	21	English (Teacher Training)	IELTS (International English Language Testing System)
P2	women	22	English (Teacher Training)	Prepare for the Examination
P3	male	21	Chinese Language (Teacher Training)	Preparing for the Examination
P4	male	22	Chinese Language (Teacher Training)	Prepare for the Examination
P5	daughter	22	Chinese Language (Teacher Training)	prepare for public examinations
P6	male	21	Mathematics (teacher training)	Preparing for Teacher Certification
P7	male	22	Mathematics (teacher training)	Grade 6 Preparation
P8	male	22	History (teacher training)	prepare for public examinations
P9	women	22	History (teacher training)	Prepare for the Examination
P10	women	21	Geography (teacher training)	Preparing for Teacher Certification

3 DISCOVERY AND DISCUSSION

3.1 Analysis of Interviews

3.1.1 Content of Online Courses

The purpose of this interview was to gain insight into what types of online classes are popular with students and have good classroom outcomes. Students generally agreed that the combination of animations, images, and human voices in the course content helped them better understand complex concepts and improve their learning efficiency. This form of multimodal learning can stimulate students' interest in learning and deepen their memory through both visual and auditory stimulation (Zhang & Yang, 2024). For example, P6 mentions, "Those courses that combine animation and vivid explanations are much easier to understand than simple textual explanations, especially some abstract theories, and the animation can help me understand them in a more visual way."

The following are some examples from the interviews: Participant 1 (P1) said, "I think Bilibili's multimodal interaction has greatly improved my learning efficiency in preparing for IELTS." P2 mentioned that, "Watching educational videos and participating in community discussions through the Beep platform helped me to better prepare for the exam." Participant 4 (P4) noted, "The interactive courses and hands-on videos on the Beefy platform enabled me to understand the complex content of Chinese language expertise more clearly as I prepared for the exam." Participant 5 (P5) remarked, "The rich visual and auditory material on the Beefy platform helped me to understand complex concepts more deeply as I prepared for the civil service exam." Participant 6 (P6) explained, "I was able to master the problem-solving skills required for the Teacher Certification Exam more effectively through the hands-on instructional videos on the Beehive." Participant 9 (P9) observed, "The interactive learning environment of Beefy promotes my active learning and significantly helps in preparing for the exam." Participant 10 (P10) shared, "Through the first-person perspective instructional videos of Bilibili, I can understand and remember the knowledge points more intuitively when learning geography expertise."

3.1.2 Teaching Methods

An interesting finding from the interviews was that the seven students specifically mentioned an approach to teaching that appears to be different from the traditional multimodal learning model: the use of first-person point-of-view footage, whereby students are taught how to analyse and answer questions by taking a first-person camera perspective. This approach, although on the surface it may appear to be dominated by a single visual stimulus, is in fact fully in line with the core concepts of multimodal learning, especially as it enriches the immersive experience of the learner through the involvement of the whole body's senses.

The teaching method uses a first-person perspective that allows students to participate as if they were there. For example, in maths problem analysis, students can experience the problem solving process, improving concentration and independent thinking. Audio explanations and teacher narration help students understand the logic and enhance the depth and breadth of learning. Students deepen their memory through imitation and manipulation, and P3 said that the first-person teaching videos made him feel involved and more immersed than traditional methods. This type of teaching enhances knowledge acquisition through physical action and thought engagement, and is particularly suitable for subjects that are highly manipulative. The interactivity and immersion of multimodal learning theory are reflected in this kind of teaching, which activates the cognitive process through multi-sensory stimulation, makes the learning content intuitive and concrete, and improves the effectiveness and fun of learning.

3.1.3 Analysis of the Usage and Experience Related to Online Courses

Frequency analysis showed that most of the university students interviewed accessed Bilibili at least three times a week, and four at least once a day, indicating its importance in their studies. As Table 2 shows, they preferred to use it in the evening, which may be related to their work schedule and study plan.

Table 2: Analysis of the use of beeping among university students.

Element	Quorum
Proportion of students visited at least once a day	4 people
Time of use preference	9 people
Motivation for use: searching for supplementary course materials	5 people
Motivation for use: watching educational videos for a better understanding	6 people
Motivation for use: participation in community discussions to expand knowledge	4 people
Percentage of students using the pop-up feature	3 people

In terms of motivation, 5 of them use it to find course materials, 6 of them use it to deepen their understanding, 4 of them participate in discussions to expand their knowledge, and 3 of them use pop-ups to increase interaction and fun, which shows that entertainment and social interaction are also the reasons for using it.

Engagement analyses showed that more than half of the students would actively interact while watching educational content. As Table 3 shows, 7 students felt that communication enhanced the learning experience and made them more engaged. 6 students felt that the multimodal interactive features improved learning efficiency and knowledge acquisition. They emphasised that video intuitiveness and pop-up feedback helped them understand complex concepts. However, 3 students mentioned that too many pop-ups could be distracting and interfere with learning.

Table 3: Students' perceptions of learning experiences and learning outcomes.

Viewpoints	Quorum	Relevant factor
Perceived that interacting with other learners enhanced the learning experience	7 people	Quality of learning content and interactive design
The multimodal interactive features of the Beeps are considered to have improved learning efficiency and knowledge acquisition.	6 people	Visualisation of video content and instant feedback from pop-ups
Point out that too many pop-ups can be distracting and interfere with learning	3 people	-

Multimodal learning theory emphasises the importance of multiple senses and communication modes in learning. The multimodal interactive function of Beep provides college students with an integrated learning environment that meets knowledge acquisition needs, enriches the interactive experience, and promotes learning effectiveness (Li, 2023).

Students' motivations for using Bilibili are mainly to find learning materials, watch educational videos and participate in community discussions. The use of the pop-up function also reflects the entertainment and social motivation, which enhances the fun and interactivity of video watching and the learning experience by actively participating through posting pop-ups or comments.

3.2 Case Studies

In terms of visual modality, images, as the central visual medium, provide intuitive presentation of information. The presentation of textual information is closely related to the Constructivist Learning Theory. This theory emphasises that learners assimilate new knowledge by constructing personal understanding. This visual information helps students to quickly capture and understand complex concepts. With the help of colourful diagrams and clear layouts, students can grasp the structure of language more intuitively, thus enhancing their learning.

In terms of textual modality, the advantage of textual modality in online classroom teaching is its ability to provide structured knowledge that helps students to learn and review on their own without relying on other modalities (e.g. visual or audio). For example, the structure of the text presented in the image is: subject + complex transitive verb + object + (object) complement, along with corresponding figurative images and elaboration of example sentences. The textual content in the images not only includes clear explanations of how complex transitive verbs are used in sentences, but also covers examples of practical application of knowledge (Rahmanu & Molnár, 2024). By reading the textual material, students can gain a deeper understanding and mastery of the knowledge.

In the audio modality, the teacher delivers the lesson through a detailed explanation of the human voice. The intonation, emphasis and rhythm of the voice help students to better understand and remember the content. The advantage of audio modality is its ability to enhance the learning experience through Emotional Resonance and Prosody Enhancement. Emotional Resonance allows students to aurally empathise with the presenter, thus improving retention and comprehension of the information. Another important role of audio modality is to help students consolidate knowledge points through auditory memory (Zhang, 2020). Audio modality adds emotion and rhythm to the learning content through the speaker's intonation and pace of speech. The narrator adopts a gentle and friendly tone of voice with a moderate pace of speech, which is neither too fast to make the information difficult to digest nor too slow to make the learners lose interest. By pausing and emphasising at the right time, the presenter adds emotion and pace to the content, making the learning process more lively and interesting. The narrator emphasised the importance of the knowledge points through an intonation that

enabled the learners to concentrate more and be impressed by the learning material.

In terms of interaction modality, a pop-up is a type of comment that is displayed on the screen in real time in the form of subtitles. Pop-up comments are sent at specific points in time during video viewing and displayed on the screen in real time, and these comments are superimposed on the screen even when the video is played back. The role of pop-ups can be explained in terms of Social Interaction Theory (SIT), which suggests that learning is a socially interactive process in which students deepen their understanding by interacting with others (e.g., classmates or teachers). In this process, pop-up comments provide a platform for students to share their understanding and receive feedback from others, which helps to enhance the motivation and effectiveness of learning (Li, 2023). There are many comments in the pop-ups that are used to explain the content of knowledge, which has a positive effect on students' understanding. For example, "A complex transitive verb is a single transitive verb followed by a compound word (object plus object complement)", "subject-verb-object + object-complement", and so on. However, words such as "have understood" and "have learnt" also appeared in the pop-ups, which may lead to self-doubt and anxiety among students who are slower to assimilate and respond, which is a potential negative effect of pop-ups (Wang et al., 2023).

4 CONCLUSION

The multimodal interaction function introduced by the Bilibili platform enhances the learning experience of university students and promotes their engagement. The platform provides a variety of content delivery methods, including video, audio, text and real-time pop-up comments, to meet the diverse learning needs of students and build a comprehensive learning environment.

While multimodal learning can provide a rich experience, it needs to be carefully designed and implemented to avoid disruption. Case studies have shown that multimodal learning environments are effective, particularly the pop-up commenting feature that enhances interaction and engagement, but can also be anxiety-provoking for slower learners.

These findings demonstrate the potential of Bilibili as an educational resource that not only provides rich multimodal content but also supports social interaction. Educators and platform designers need to balance engagement and distraction and optimise multimodal learning environments to support all learners. In this way, Bilibili can become a powerful learning tool that helps students explore the ocean of knowledge while having fun socially.

However, the study also has some limitations. The sample of this study was mainly drawn from students using Bilibili at Guangdong University of Petrochemical Technology (GDUPT), which may limit the generalisability of the results. Students from different regions and cultural backgrounds may have different acceptance and preference for multimodal interactions, and the findings may not be representative of all college students. Although qualitative analyses were used to explore student experiences in depth, the lack of quantitative data may affect the objectivity and comprehensiveness of the findings. Interviews and case studies rely on subjective reports, which may introduce bias. Technology continues to advance, research data may be outdated, and new technological tools and learning platforms may change student learning styles and preferences, affecting the long-term validity of conclusions. Future research will expand the sample to include students from different geographic and cultural backgrounds and explore the needs and responses to multimodal interactions among students from different academic fields and professional backgrounds. Long-term follow-up studies will observe how new technologies affect student learning behaviours and outcomes and understand the impact of multimodal interactions on student motivation and effectiveness at different points in time. Further research will explore how multimodal interaction technology can be used to provide personalised learning paths and support for students, adjusting the content and interaction format according to students' individual interests, learning styles and progress, in order to optimise the use of multimodal interaction in learning and enhance the effectiveness and quality of online education.

REFERENCES

Li, M. X. 2023. Study on the relationship between college students' motives for using bullet comments and their behaviour in Bilibili online courses. social sciences II Vol., *Information Technology*, (02): 1-15.

Li, X. 2024. Online learning platforms and their impact on the learning outcomes of higher education. *Science Education and Culture*, 21: 16-21.

Lim, F. V., Toh, W., & Nguyen, T. T. H. 2022. Multimodality in the English language classroom: a systematic

review of literature. *Linguistics and Education*, 69, 101048.

Liu, B. Y. 2024. Research on the design of remote teaching products based on multimodal interaction. *Beijing University of Chemical Technology Journal*.

Liu, W., & Wang, S. 2024. The current situation and implications of college English CET-6 vocabulary online course teaching from the perspective of the lexical chunk theory. *Overseas English*, 11: 22-30.

Rahmanu, I. W. E. D., & Molnár, G. 2024. Multimodal immersion in English language learning in higher education: a systematic review. *Heliyon*, 10(19): e38357.

Sun, J., Zhao, H., Li, P., Chi, X., & Mu, X. 2015. A study on the influencing factors of college students' use of online course teaching. *Chizi (Upper Middle)*, 22: 45-53.

Wang, Y., Tian, P., Pu, D., & Sun, H. 2023. Research on factors affecting and coping strategies for English learners' anxiety in online courses. *Journal of Jilin University of Education*, 39(7): 22.

Zhang, X. L. 2020. A study on the impact of multimodal teaching methods on English vocabulary learning for adult non-English majors. *Journal of East China University of Journal of East China University of Science and Technology*, (1).

Zhang, X., & Yang, M. 2024. A study on the mode of precise teaching in ideological and political courses from the perspective of multimodal learning theory. *Heilongjiang Higher Education Research*, 42(10): 56-64.