

Progress in a multimodal learning evaluation based on artificial intelligence

Chuyi Chen, Ming Tang*, Aobo Yang

Department of Educational Technology, College of Humanities and Education, Foshan University, Foshan, 523800, China

*Corresponding author: hunutangming@fosu.edu.cn

Abstract: *In the context of artificial intelligence, the rapid development of information technology and the importance of innovative talent training, in order to promote the scientific, precision and effectiveness of learning evaluation, this paper will explore the research progress of multi-modal learning evaluation in the background of artificial intelligence. First, we will define the core concepts in the research, namely artificial intelligence, multimodal, and learning evaluation, and analyze their functions and characteristics. Secondly, we will further summarize the research results of multimodal learning evaluation. These research results mainly focus on the aspects of multimodal accurate evaluation, diversified space, evaluation framework and evaluation model, etc., which promote the research of multimodal learning evaluation and make future teaching evaluation more scientific, more accurate and more effective. This paper will provide a useful reference for the future development of multimodal learning evaluation in the context of artificial intelligence.*

Keywords: *Artificial Intelligence, Multimodal evaluation, Multimodal data*

1. Foreword

The 21st century is an era with artificial intelligence as its core. Countries all over the world have realized the importance of mastering advanced technology and the training of innovative talents, and have begun to pay attention to the development of artificial intelligence. In the past few years, various countries have released relevant policies and documents on the development of artificial intelligence. For example, the UK Department of Science, Innovation and Technology (DSIT) published a white paper on artificial intelligence. The U.S. government's report on Artificial Intelligence, Automation, and the Economy; The Korean government released the National Strategy for Artificial Intelligence, and so on. In February 2019, China released the journal China Education Modernization 2035, which mentioned encouraging the development of diversified learning evaluation, emphasizing the process and development of evaluation. Learning evaluation is not only the evaluation of students' learning results, but also the feedback of teachers' teaching effect. The diversity of learning assessment focuses not only on student learning outcomes, but also on student learning processes and methods. At the same time, China's Education Modernization 2035 also emphasizes the importance of educational evaluation with modern technology. Artificial intelligence technology can provide more efficient, accurate and objective methods for learning evaluation, and help to better play the role of learning evaluation. How to use these technologies to carry out learning evaluation has become a new educational problem. Under the background of artificial intelligence, learning evaluation based on multimodal technology has become a new research hotspot in the field of education, which provides a new way for the study of learning evaluation. Multimodal technology can comprehensively make use of the characteristics of various information transmission, such as text, image, sound and video, to provide more comprehensive and accurate data support for learning evaluation

This paper will study the definition, role and research progress of multimodal learning evaluation in the context of artificial intelligence. First, we will define the core concepts of artificial intelligence, multimodality, and learning evaluation, and analyze their roles and characteristics. Then, we will deeply summarize the research results of multimodal learning evaluation, which mainly focus on multimodal accurate evaluation, diversified space, evaluation framework and evaluation model. These research results have promoted the research of multimodal learning evaluation, which makes the teaching evaluation more scientific, more accurate and more effective in the future. In addition, based on the current research results, we explore the development trend of multimodal learning evaluation in

the context of artificial intelligence, and expect that this paper can provide a useful reference for the future development of multimodal learning evaluation in the context of artificial intelligence.

2. Definition of relevant concepts

2.1 Artificial intelligence

Researchers in the world have many definitions of artificial intelligence with different specific contents. At present, the definition of artificial intelligence mainly includes several representatives as follows: 1. Professor Nelson has given such a definition of artificial intelligence: " Artificial intelligence is a subject of knowledge- -how to represent knowledge and how to acquire knowledge and use it^[1]. " 2. " AI is a study of research on how to make computers do intelligent work that only people could do in the past." 3. Simon, a representative of symbolism, believes that artificial intelligence is a reflection of the real world. As long as it is true in the machine, it is also true in the real world.

Artificial intelligence, a technology development and application, has become a household name in today's society. It is not only a single science, but also an interdisciplinary discipline across multiple fields, whose research goal is to enable computers to think, learn, reason, make decisions, communicate and so on, like human beings, to complete some complex tasks. Through related technologies such as machine learning and deep learning, their laws and trends are found from large amounts of data, so as to make more accurate predictions and decisions. By analyzing students' learning trajectories and relevant data, artificial intelligence has tailored their learning plans and resources for each student, so as to achieve individualized teaching. In addition, artificial intelligence technology can also assist teachers to correct their homework and exams, improve teachers' work efficiency, reduce their work burden, so that teachers have more time and energy for the development of students, and improve teaching efficiency.

2.2 Multimode

Multimodal technology refers to the integration of two or more kinds of perceptual channels and information interaction, such as vision, hearing and facial expression and other human physiological data, which allows us to comprehensively obtain and express information. These technologies can simulate the natural human perception process, which can provide a better rich and complete information input and interaction experience for the fields such as machine learning and human-computer interaction. Multimodal technology combines a variety of different types of data to improve the understanding and interaction ability of AI systems, making artificial intelligence technology more accurately and naturally understand user intention and situational information, provide more comprehensive and accurate data input for machine learning. Through advanced speech recognition, image recognition and other technologies, people can interact with machines in a more natural way.

In the field of education, multimodal technology effectively improves the interest and attraction of learning by giving students a variety of sensory stimulation. In addition, multimodal technology is used to develop intelligent assessment systems that not only analyze students' answers, but also analyze students' multiple sensory data to assess students' learning effectiveness, understanding, and emotional state.

2.3 Learning evaluation

Learning evaluation refers to the process of systematically collecting information and evidence of students' cognitive and behavioral changes under the influence of teaching and self-study of various subjects on the basis of certain educational and teaching goals and using appropriate and effective tools and approaches, and making value judgments on students' knowledge and ability.

Learning evaluation can not only give timely feedback and guidance to students, help them clearly understand their own learning status, find the existing shortcomings and problems, but also have a significant effect on stimulating students' enthusiasm and autonomy in learning. It can encourage students to participate in the learning process more actively, and cultivate their independent learning and problem-solving abilities. When students receive positive feedback and affirmation in the evaluation, their motivation and self-confidence will be greatly improved, and they will be more engaged in the pursuit of learning progress. It also provides teachers with accurate assessments of

student learning, based on which they can develop personalized guidance for students to improve learning methods, strengthen weak areas, and provide additional resources and support. Learning evaluation can become a bridge of communication and interaction between teachers and students, make teaching more targeted and effective, further optimize the overall teaching effect, and create a more favorable environment for students' learning and growth.

3. Research progress of learning evaluation

In recent years, with the development of quality-oriented education, the reform of educational evaluation has gradually increased. The research on educational evaluation has become the focus of many scholars at home and abroad, and has achieved a lot of fruitful results. These research results mainly focus on the multi-modal accurate evaluation, diversified space, evaluation framework and evaluation model.

3.1 Multimodal data to promote the accurate evaluation of students' learning

Multimodal data is mainly based on the data obtained from different channels, through integration and analysis, so as to provide more complete and rich information for the computer. Multimodal learning evaluation is to capture a variety of data such as emotion, language and attitude of students in the learning process through devices such as eye tracker, sensor and wearing bracelet. Multimodal technology is used to scientifically screen, sort out and analyze these data according to a scientific evaluation system, and evaluate the learning effect of students according to the obtained results. In 2022, Zhang Jiahua et al. divided multi-modal data into four types, namely explicit data, physiological data, psychological data and basic data, as shown in Table 1^[2]. Through the analysis of students' learning situation and the data generated in the learning process, it provides an important basis for carrying out accurate learning intervention and optimizing the teaching process.

Table 1: The multimodal type data

Data Classification	Multimodal data	Utility unit	Evaluation Characterization
Extrinsic data	Text, language, handwriting, voice, video views...	Calculators and mobile devices, monitoring, questionnaires, and statistical tools	It reflects the learning status of learners, including their learning results, learning motivation, knowledge mastery and the interaction between teachers and students.
Physiological data	Brain waves, heart rate, Pressure, breathing rhythm...	Wearable device, Electrocardiogram machine	To deeply understand the physiological responses of learners, such as heart rate, blood pressure, etc., and to predict their learning emotions
Psychological data	Mood, hand gestures, eye movements, and facial expressions	Eye tracker, sensor	It reflects learners' learning attention, learning interest and enthusiasm, insight into cognitive development, representation and information integration ability, and collaborative learning ability
Data base	Student's basic information, learning information, current cognitive level, family situation	Growth file, report card, home visit, telephone communication	Reflect the students' learning style, learning foundation

How to make better use of these multimodal data to promote the accurate evaluation and personalized evaluation of students' learning. In 2021, Huang Tao et al. mentioned that data-driven accurate learning evaluation can better help us interpret and give feedback on teaching information, timely discover problems in education, and carry out effective intervention^[3]. In 2023, Wu Junqi and others also mentioned that different types of multimodal data can provide practical suggestions on the teaching process and teaching evaluation^[4]. The above studies on multimodal data have greatly promoted the precision and personalized research based on the multimodal data, which is conducive to

our subsequent use of multimodal data for accurate learning evaluation, so that we can be able to maximize the use of data while effectively facing the needs of data processing. However, the above results do not take into account the specific data model and evaluation indicators, and can not be a good overall assessment of learners' learning performance.

3.2 The pluralism of the main body of multimodal learning evaluation promotes teachers' teaching

In the current teaching, the evaluation means of traditional evaluation methods are usually relatively simple, only relying on examination, homework and other methods, only paying attention to the students' knowledge mastery, and the evaluation subject is often dominated by teachers. The biggest difference between multimodal learning evaluation and traditional learning evaluation is that the main body of multimodal learning evaluation is not only teachers, but also students, parents, peers and so on. This often gives people the illusion that in the learning evaluation, the subject position of teachers is weakened and becomes no longer important. In fact, the position of teachers in the new learning evaluation has become more important, especially in the multimodal learning evaluation. In 2019, Lin Ruifang and others also mentioned that the combination of teacher evaluation and student mutual evaluation in the multimodal autonomous learning model can more promote students' learning^[5]. In the multimodal learning evaluation, teachers' evaluation is authoritative and crucial for students' learning, which will guide students' next learning direction. And the students will also improve and move forward in the direction of learning guided by the teacher. Under the guidance of teaching objectives, teachers clearly know the relevant learning situation of students, and interpret and evaluate them, so as to make immediate decisions for the next step of learning. In this process, an immediate and dynamic interactive formative evaluation is formed. In 2022, Zhan Zehui and others also mentioned that students more recognized the combination of teacher evaluation, mutual evaluation and self-evaluation^[6]. The above research plays an important role in promoting the diversification of subjects in multimodal learning evaluation, and has achieved good research results. However, how to play the role of teachers as one of the subjects needs to be deepened, and the precision needs to be improved.

3.3 Diversification of learning space to enrich the multimodal learning evaluation

With the progress of science and technology and the rise of the concept of mobile learning, students' learning space is no longer limited to the traditional schools and classrooms. Nowadays, society, family, community and other fields have become places for learning, and students can learn at any time and anywhere, which also leads to the ubiquitous generation of learning data. Therefore, in addition to the real classroom space, the spatial subject of multimodal learning also has the virtual network space, community and so on. In 2020, Zhang Qi et al. mentioned the importance of data space allocation and annotation in multimodal analysis, which helps to more accurately intervene and guide learning behavior^[7]. In 2023, Liang Aihua and others took the python programming course as an example, and mentioned that the multimodal data collection should cover the physical scenarios and network scenarios to ensure the comprehensiveness and accuracy of the data^[8]. In addition, Zhang Jinghan and others also mentioned in 2023 that multimodal learning data is a mirror representation of learners' learning state, and it is necessary to accurately capture the data of learners in different learning spaces^[9]. Through the above literature, we find that the processing and analysis of multimodal learning data is crucial for learning evaluation. With the diversification of learning space, the future research on multimodal learning should pay more attention to the comprehensive coverage and accurate annotation of data space in order to better support learners' learning and development.

3.4 Improve the multi-modal evaluation framework and better play the role of data evaluation

At present, the multimodal evaluation framework lacks rich experience and needs to be improved. In 2022, Zhang Jiahua et al. mentioned the four main links of establishing goals, obtaining data, establishing models and providing feedback in order to provide learners with a dynamic cycle of evaluation^[3]. In 2023, Fan Fulan et al. proposed to construct a cognitive depth evaluation framework of college students based on multi-modal data in a normal teaching environment from three aspects: data collection, processing and analysis, and the final conclusion, so as to carry out multi-modal learning evaluation^[10]. In the same year, Han Pengjing proposed to design a multidimensional learning evaluation model based on constructivism theory and multiple intelligences theory, which described students' learning status in the wisdom classroom from multiple dimensions and conducted a complete evaluation of students^[11]. At present, there is a preliminary research framework for the process and framework of multimodal learning evaluation, but there is not much discussion about what kind of

framework we should adopt for each scene, what kind of data should be collected and how to collect data.

3.5 The multimodal data model is established to help the future development of multimodal education evaluation

With the development of Internet of Things technology, sensors and wearable devices have been widely used in daily life and education, and multimodal data have also promoted the development of learning evaluation. In 2021, Wang Weifu et al. believed that future multimodal data can effectively combine data in different learning scenarios, providing a more comprehensive, accurate and objective method for constructive higher-order learning assessment ^[12]. At the same time, Luo Fang et al. believed in 2021 that the establishment of a targeted, refined, accurate and interpretable multi-modal data model could provide a new solution for China's educational evaluation reform ^[13]. In 2022, Hu Hang et al. believed that establishing a multi-modal data prediction model can effectively promote students' deep learning ^[14]. By establishing a variety of multimodal data models, we can make multimodal data play a greater role in educational evaluation by sorting and processing multimodal data.

4. Conclusion

In the context of artificial intelligence, in order to promote the scientific, accurate and effective learning evaluation, this paper will discuss the research progress of multimodal learning evaluation in the context of artificial intelligence. We will further summarize the research results of multimodal learning evaluation. These research results mainly focus on the aspects of multimodal accurate evaluation, diversified space, evaluation framework and evaluation model, etc., which promote the research of multimodal learning evaluation and make future teaching evaluation more scientific, more accurate and more effective. At present, the existing research results have played a positive role in the research of teaching evaluation, but there are still some problems to be solved. For example, based on the research on the proportion of multimodal data and its influencing factors in different types of classrooms, the research on the construction of evaluation models based on multi-modal data, and the construction of a comprehensive evaluation framework and index system capable of multi-spatial full data. Further research is needed on the mechanism of multi-spatial data sharing and the influence factors of multi-modal data on students' learning in specific classes. The research on these specific issues will be an important research direction in the field of education based on multimode, which will contribute to the automatic evaluation based on artificial intelligence and its system realization, and will be a breakthrough in education informatization in the era of artificial intelligence.

Acknowledgements

We would like to acknowledge our sincere thanks to the management of our colleges, our colleagues of Foshan University. This research has been supported by Guangdong Provincial Department of Education Characteristic Innovation Project (Humanities and Social Sciences) and research Project of the Steering Committee of Open Online Courses for Undergraduate Universities in Guangdong Province (2022ZXKC482).

References

- [1] Zou Lei, Zhang Xianfeng. *Artificial intelligence and its developmental applications [J]. Information Network Security*, 2012, (02): 11-13.
- [2] Zhang Jiahua, Hu Huizhi, Huang Changqin. *Learning evaluation research supported by multimodal learning analysis technology [J]. Modern Educational Technology*, 2022, 32 (09): 38-45.
- [3] Huang Tao, Zhao Yuan, Geng Jing, etc. *Data-driven evaluation mechanism and method of precision learning [J]. Research in Modern Distance Education*, 2021,33 (01): 3-12
- [4] Wu Junqi, Ren Feifei, Wu Feiyan. *Construction and application of a key index system for data-driven classroom precision teaching [J]. Modern long-distance education*, 2023(02):39-52.
- [5] Lin Ruifang, Dong Liyun. *Discussion on the multimodal autonomous learning mode and evaluation system [J]. Journal of Jimei University (Education Science Edition)*, 2019, 20 (05): 59-63.
- [6] Zhan Zehui, Yao Jiajing, Wu Qianyi, etc. *Design and application of performance evaluation in the*

- AI curriculum [J]. Modern Educational Technology, 2022,32 (05): 32-41.*
- [7] Zhang Qi, Li Fuhua, Sun Ji-nan. *Multimodal Learning Analysis: Learning Analysis in the era of computational education [J]. China Audio-visual Education, 2020, (09): 7-14 + 39.*
- [8] Liang Aihua, Wang Xueqiao. *Multimodal learning data collection and fusion [C] // Network Application Branch of China Computer Users Association. Proceedings of the 27th Annual Conference of Network New Technology and Application of Network Application Branch of China Computer Users Association in 2023. Engineering Comprehensive Experimental Teaching Demonstration Center of Beijing Union University; Beijing Union University Frontier Intelligent Technology Research Institute; 2023:5.*
- [9] Zhang Jinghan, Xu Zhuoxin. *Multimodal data color learning investment evaluation of new development [J]. Journal of Sichuan Vocational and Technical College, 2023,33 (05): 53-59.*
- [10] Fan Fulan, Li Hanting, Qi Tianjiao, etc. *Construction and application of the framework of cognitive depth evaluation of college students based on multimodal data [J]. Modern Long Distance Education, 2023, (03): 57-65*
- [11] Han Pengjing. *Research on the learning evaluation method based on multimodal information fusion from the perspective of intelligent classroom [D]. Northeastern Petroleum University, 2023.*
- [12] Wang Weifeng, MAO Meijuan. *Multimodal learning analysis: a new approach to understanding and evaluating real learning [J]. Audio-visual Education Research, 2021,42 (02): 25-32*
- [13] Luo Fang, Tian Xuetao, Tu Zhuoran, etc. *New trend of educational evaluation: a summary of intelligent evaluation research [J]. Research on Modern Distance Education, 2021,33 (05): 42-52.*
- [14] Hu Hang, Yang Yang. *Evaluation path and strategy of threshold deep learning [J]. Distance Education in China, 2022, (02): 13-19 + 76.*