

**Title: An Exploratory Study on a Web-based Formative Assessment Tool for  
Developing EFL Learners' Oral Proficiency at a Chinese Higher Vocational College**

**Name: Yixian Lyu (2000459)**

**Supervised by Dr Md Golam Jamil, Dr Ros Coard**

**Submitted in partial fulfilment for the award of the degree of professional doctor on  
English Language Education and English Literature**

**University of Wales Trinity Saint David**

**2025**

## DECLARATION

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed *Yixian Lyu* (candidate)

Date: January 31, 2025

## STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated. Where correction services have been used the extent and nature of the correction is clearly marked in a footnote(s). Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed *Yixian Lyu* (candidate)

Date: January 31, 2025

## STATEMENT 2

I hereby give consent for my thesis, if accepted, to be available for deposit in the University's digital repository.

Signed *Yixian Lyu* (candidate)

Date: January 31, 2025

## ACKNOWLEDGEMENTS

As I come to the end of this more than five-year-long journey of doctoral research, I would like to express my sincere gratitude to those who have helped and supported me in my learning and study. It has been a path paved with intellectual challenges, moments of self-doubt, and countless sacrifices, yet also one filled with invaluable learning, unwavering support, and the resilience to push forward. Without the encouragement and generosity of many individuals and institutions, this achievement would not have been possible.

First and foremost, my deepest gratitude goes to UWTSD and my esteemed supervisors, Dr Md Golam Jamil, Dr Ros Coard. Their profound knowledge, insightful guidance, and rigorous training in theories and research methodologies have shaped me as a scholar. Their dedication went beyond mere academic instruction—they sacrificed their weekends and personal time to hold online meetings with me, ensuring that I remained on track. Their patience, encouragement, and belief in my potential provided the foundation upon which this research was built. Without their unwavering commitment, I could not have successfully navigated the complexities of this study.

I am profoundly indebted to all the participants who contributed to this research. My heartfelt thanks go to the 130 students from JTC who took part in the survey and the 44 students who endured the rigour of my 20-week experimental study. Their willingness to engage, their perseverance, and their dedication allowed me to gather the crucial data that shaped my findings. A special mention must be made to JTC's foreign teacher, Martin, who stood by my side throughout the lengthy observation period. His meticulous tracking of students' oral English proficiency development in both the experimental and control groups, along with his participation in four rounds of interviews, added significant depth to this study. Additionally, I extend my sincere gratitude to Peter from iFLYTEK, whose technical expertise and support proved invaluable. He ensured that both the students and I became proficient in using the FiF platform and swiftly addressed any technical issues that arose, making the research process considerably smoother.

This journey would have been insurmountable without the love and unwavering support of my family. As a son, husband, and father, my pursuit of this doctoral degree came at a cost—not just my own sacrifices, but those of my loved ones. My mother, my

wife, and my son provided me with the strength and motivation to continue despite the challenges. Their patience in my prolonged absences, their encouragement during moments of despair, and their faith in my aspirations were the silent yet powerful forces that carried me through. To them, this achievement is as much theirs as it is mine.

I am also deeply grateful to JTC. The college granted me the opportunity to pursue my doctoral studies abroad. Their generosity in providing the necessary support and their flexibility in work arrangements allowed me to fully dedicate myself to my research. Without their understanding and commitment, balancing professional and academic responsibilities would have been nearly impossible.

A special note of appreciation is owed to iFLYTEK for their generous contribution to this research. Their provision of 60 extended, two-year licenses for the FiF platform at no cost proved instrumental in establishing the technological foundation for this study.

Furthermore, my heartfelt thanks go to the dedicated staff at UWTSD's Postgraduate Office, Library, Student Centre, and Reception. Their warmth, professionalism, and willingness to assist ensured that I could focus on my academic work with fewer distractions. Their generosity in offering guidance, resources, and support was an integral part of my academic journey.

Lastly, I extend my gratitude to the kind and compassionate individuals I encountered during my time in the UK. Whether it was a simple smile from a stranger, a word of encouragement from a fellow student, or the thoughtful gestures of the local community, these seemingly small acts of kindness made a world of difference during my most challenging days. In a foreign land, far from home, these moments of connection and humanity became my anchor.

This journey has not been easy—it has been fraught with challenges, sleepless nights, and moments of doubt. Yet, it has also been a period of immense personal and academic growth, resilience, and ultimate fulfilment. To everyone who walked this path with me, offering their support, guidance, and belief in my journey, I extend my deepest and most heartfelt appreciation. This achievement is not mine alone—it is a testament to the collective efforts of many.

## **ABSTRACT**

The purpose of this study was to investigate the impact of WBFAT on the oral proficiency development of EFL learners. The research was conducted at a higher vocational college in China, 119 students participated in a survey exploring their attitudes towards WBFAT. Subsequently, 44 of these students were selected for a 20-week experimental study. The study yielded findings distinct from prior research. Firstly, participants exhibited a positive attitude towards WBFAT, with a strong positive correlation identified between their willingness to use the tool, their motivation, and their expectations. Higher willingness was aligned with stronger motivation and higher expectations. However, the demand for personalisation and interactive functions highlighted areas requiring improvement in WBFAT. Secondly, significant individual differences were observed in students' self-regulated learning motivation, habits, and strategies. Nonetheless, the experimental intervention had a marked positive impact on students' autonomous learning behaviours. Notable improvements were seen in time management, task completion, feedback utilisation, strategy adjustment, and peer communication. The experimental group significantly outperformed the control group in these areas, with the intervention fostering more active engagement in self-directed learning tasks. These enhancements translated into substantial gains in critical language skills such as fluency and coherence. Thirdly, the gap in oral proficiency scores between the experimental and control groups widened over time, particularly during the latter stages of the study. This demonstrated the growing advantages of the learning approach adopted by the experimental group. Fourthly, in terms of linguistic features, the experimental group excelled in vocabulary breadth, sentence length, and lexical diversity. However, certain metrics, such as word concreteness and temporal cohesion, showed slight declines. No significant differences were observed between the two groups in coherence, readability, and comprehension difficulty measures. Moreover, the experimental group displayed consistent improvements in classroom participation, speaking accuracy, fluency, and coherence, with progress being particularly notable during the latter stages of observation. These trends were closely linked to the use of WBFAT, underscoring its significant impact on enhancing oral English proficiency. Finally, in the student survey on the perceived importance and satisfaction with WBFAT, features such as feedback, scoring accuracy, and the immediacy and effectiveness of face-to-face interaction were identified as the most valued aspects of WBFAT.

## TABLES OF CONTENTS

|                                                                                                             |    |
|-------------------------------------------------------------------------------------------------------------|----|
| Chapter 1 Introduction .....                                                                                | 1  |
| 1.1 Research Background .....                                                                               | 1  |
| 1.1.1 The Popularisation of Educational Informatisation in China .....                                      | 1  |
| 1.1.2 Changes and Impacts on Education Caused by the Informatisation .....                                  | 2  |
| 1.2 Definition of WBFAT and Introduction to FiF .....                                                       | 4  |
| 1.2.1 Definition of WBFAT .....                                                                             | 4  |
| 1.2.2 Introduction to FiF .....                                                                             | 5  |
| 1.3 Purpose of the Study .....                                                                              | 5  |
| 1.4 Statement of Research Problem .....                                                                     | 5  |
| 1.4.1 Theme One: Students' Attitudes Towards Using WBFAT .....                                              | 6  |
| 1.4.2 Theme Two: Students' Performance of OEAL & FA and Progress in oral English<br>After Using WBFAT ..... | 7  |
| 1.4.3 Theme Three: Students' Perceptions of WBFAT .....                                                     | 10 |
| 1.5 Significance of the Study .....                                                                         | 11 |
| 1.6 Research Methodology .....                                                                              | 12 |
| 1.6.1 Data Collection .....                                                                                 | 12 |
| 1.6.2 Data Analysis and Integration .....                                                                   | 13 |
| 1.6.3 Data Reporting and Interpretation .....                                                               | 14 |
| 1.7 Organisation of the Dissertation .....                                                                  | 15 |
| Chapter 2 General Introduction to Chinese School Education .....                                            | 16 |
| 2.1 The Origin, Heritage, and Transformation of Chinese School Education .....                              | 16 |
| 2.1.1 Ancient School Education: Development and Influence .....                                             | 16 |
| 2.1.2 The Transition to Modern Education .....                                                              | 17 |
| 2.1.3 Future Directions and Challenges .....                                                                | 18 |
| 2.2 The Rapid Development of Contemporary Education in China .....                                          | 18 |
| 2.2.1 The Improvement and Innovation of the Education System .....                                          | 18 |
| 2.2.2 Continuous Improvement in Education Quality .....                                                     | 19 |
| 2.2.3 Narrowing the Urban-Rural Education Gap .....                                                         | 19 |
| 2.2.4 The Popularisation of Digital and Smart Education .....                                               | 20 |
| 2.2.5 Accelerating Internationalisation of Education .....                                                  | 20 |
| 2.3 Modern Higher Education in China .....                                                                  | 20 |
| 2.3.1 The Flourishing Development of Undergraduate Education .....                                          | 20 |
| 2.3.2 The Rapid Rise of Vocational Education .....                                                          | 21 |
| 2.4 Educational Informatisation in China .....                                                              | 24 |
| 2.5 The Educational Philosophy of Chinese School .....                                                      | 25 |
| 2.5.1 Aims of Education .....                                                                               | 25 |
| 2.5.2 Curriculum .....                                                                                      | 26 |
| 2.5.3 Teaching Methods .....                                                                                | 27 |
| 2.5.4 Teacher and Student .....                                                                             | 28 |
| 2.5.5 Teaching Content and Teaching materials .....                                                         | 28 |
| 2.5.6 Assessment for Teaching and Learning .....                                                            | 29 |
| 2.5.7 School Administration .....                                                                           | 30 |
| 2.6 Current Research on Chinese School Education .....                                                      | 31 |
| 2.7 Chapter Summary .....                                                                                   | 32 |
| Chapter 3 Literature Review .....                                                                           | 34 |
| 3.1 Procedure of Literature Selection .....                                                                 | 34 |

|                                                                                                                                                                  |     |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 3.1.1 Data-Bases and Search Strategy .....                                                                                                                       | 35  |
| 3.1.2 Criteria for Inclusion and Identification of Relevant Publications .....                                                                                   | 36  |
| 3.2 Screening, Eligibility, and Quality Appraisal of Literature .....                                                                                            | 36  |
| 3.2.1 Screening and Eligibility .....                                                                                                                            | 36  |
| 3.2.2 Quality Appraisal .....                                                                                                                                    | 37  |
| 3.3 Explanatory Analysis of the Included Publications .....                                                                                                      | 38  |
| 3.3.1 The Research on WBFAT for Learning Has Rapidly Increased .....                                                                                             | 38  |
| 3.3.2 The Research on WBFAT for Learning Are Mainly Concentrated in the United States and China .....                                                            | 40  |
| 3.3.3 The Distribution of the Included Literature in Terms of Participants, Study Duration, Study Context, Disciplinary Distribution, and Research Methods ..... | 41  |
| 3.3.4 Theoretical Basis in the Included Literature .....                                                                                                         | 46  |
| 3.3.5 Web-based or Mobile-based Formative Assessment Tools Utilised in the Literature and Their Impacts on Learning .....                                        | 51  |
| 3.4 Review of Relevant Theories .....                                                                                                                            | 56  |
| 3.4.1 Learning Theories .....                                                                                                                                    | 56  |
| 3.4.2 Theories of Assessment for Learning .....                                                                                                                  | 66  |
| 3.4.3 Technology Acceptance and Application Theory .....                                                                                                         | 72  |
| 3.5 Chapter Summary .....                                                                                                                                        | 75  |
| Chapter 4 Methodology .....                                                                                                                                      | 77  |
| 4.1 Design of the Mixed Methodology .....                                                                                                                        | 77  |
| 4.1.1 An Explanatory Sequential Design of Mixed Methods .....                                                                                                    | 77  |
| 4.1.2 Why Mixed Methods .....                                                                                                                                    | 80  |
| 4.2 Settings .....                                                                                                                                               | 82  |
| 4.2.1 Settings of the Survey .....                                                                                                                               | 83  |
| 4.2.2 Settings of Interviews .....                                                                                                                               | 83  |
| 4.2.3 Settings of the Focus Group .....                                                                                                                          | 84  |
| 4.2.4 Settings of the OEAL & FA in the Phase of Experiment .....                                                                                                 | 85  |
| 4.2.5 Setting of the Instrument in the phase of Experiment - FiF .....                                                                                           | 85  |
| 4.3 Research Procedures .....                                                                                                                                    | 89  |
| 4.3.1 Pilot Study .....                                                                                                                                          | 90  |
| 4.3.2 Survey .....                                                                                                                                               | 92  |
| 4.3.3 Experiment .....                                                                                                                                           | 94  |
| 4.4 Role of the Researcher, Instructor and Technician .....                                                                                                      | 104 |
| 4.4.1 Role of the Researcher .....                                                                                                                               | 104 |
| 4.4.2 Role of the Instructor .....                                                                                                                               | 104 |
| 4.4.3 Role of the Technician .....                                                                                                                               | 106 |
| 4.5 Ensuring the Quality of Findings .....                                                                                                                       | 106 |
| 4.5.1 Pilot Study .....                                                                                                                                          | 106 |
| 4.5.2 Triangulation .....                                                                                                                                        | 107 |
| 4.5.3 Eliminating the influence of the Researcher's Status as a Teacher .....                                                                                    | 108 |
| 4.6 Data Analysis and Integration .....                                                                                                                          | 109 |
| 4.7 Chapter Summary .....                                                                                                                                        | 112 |
| Chapter 5 Findings .....                                                                                                                                         | 113 |
| 5.1 Students' Attitudes Towards Using WBFAT .....                                                                                                                | 113 |
| 5.1.1 Demographic Description of the Participants .....                                                                                                          | 113 |
| 5.1.2 Reliability and Validity of the Scale Items .....                                                                                                          | 113 |

|                                                                                                                                                                 |     |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 5.1.3 Statistical Analysis of 21 Scale Items in the Questionnaire .....                                                                                         | 114 |
| 5.2 Students' Engagement in OEAL & FA .....                                                                                                                     | 117 |
| 5.2.1 Descriptive Analysis of Students' Engagement in OEAL & FA .....                                                                                           | 117 |
| 5.2.2 Analysis of Differences between the EG and CG .....                                                                                                       | 119 |
| 5.3 Students' Progress in Scores of the Five Tests of Oral English .....                                                                                        | 122 |
| 5.3.1 Descriptive Analysis of the Five Tests .....                                                                                                              | 122 |
| 5.3.2 The trend of scores of two groups of students in five oral tests .....                                                                                    | 124 |
| 5.4 Students' Progress in Linguistic Features From Their Audio Transcripts .....                                                                                | 124 |
| 5.4.1 Descriptive Analysis of Students' Five Audio Transcripts .....                                                                                            | 124 |
| 5.4.2 <i>T</i> -test of 20 Indicators' Mean of Language Features in the Five Tests .....                                                                        | 128 |
| 5.5 Student's Performance of Oral English in Class Observation .....                                                                                            | 130 |
| 5.5.1 EG Showed Higher Classroom Engagement, Oral Accuracy, Fluency, and<br>Coherence than the CG, with Notable Improvements after Using FiF .....              | 130 |
| 5.5.2 EG Showed Continuous and Significant Improvements in Oral English Proficiency<br>Over Time With the Use of FiF .....                                      | 131 |
| 5.6 Themes of the Instructor Interviews .....                                                                                                                   | 132 |
| 5.6.1 WBFAT Improved Students' Participation and Initiative in Oral English in Class .....                                                                      | 133 |
| 5.6.2 WBFAT Enhanced Confidence in Oral English Expression Among Higher<br>Vocational College Students .....                                                    | 133 |
| 5.6.3 WBFAT Promoted Students' Pronunciation Accuracy in Oral English .....                                                                                     | 134 |
| 5.6.4 WBFAT Helped Higher Vocational College Students Improve Vocabulary<br>Accuracy, Specificity, and Grammatical Accuracy .....                               | 134 |
| 5.6.5 WBFAT Had Positive Impacts on Fluency and Coherence in Oral English .....                                                                                 | 135 |
| 5.6.6 Higher Vocational College Students Still Face Challenges in the Development of<br>Oral English Proficiency .....                                          | 135 |
| 5.7 Students' Perceptions of WBFAT .....                                                                                                                        | 135 |
| 5.7.1 Reliability and Validity Test of the Scales Used in the Survey on Students'<br>Perceptions of the Satisfaction and Importance of FiF .....                | 136 |
| 5.7.2 <i>T</i> -test Highlights the Need for Improvement in Users' Satisfaction .....                                                                           | 136 |
| 5.7.3 Importance-Performance Analysis Highlights Areas to Prioritise for Improvement .....                                                                      | 138 |
| 5.8 Findings From the Student Focus Group Discussion .....                                                                                                      | 139 |
| 5.8.1 Student Focus on WBFAT Satisfaction and Importance .....                                                                                                  | 139 |
| 5.8.2 Students were most concerned about feedback and grading functions of WBFAT .....                                                                          | 142 |
| 5.8.3 Students Were Dissatisfied with the WBFAT's Functions of Feedback and Scoring .....                                                                       | 143 |
| 5.9 Chapter Summary .....                                                                                                                                       | 144 |
| Chapter 6 Discussion .....                                                                                                                                      | 146 |
| 6.1 Revisiting the Research Aims and Questions .....                                                                                                            | 146 |
| 6.2 Discussion on Theme 1 .....                                                                                                                                 | 147 |
| 6.2.1 High Value of Expectation of Using WBFAT Indicates a Strong Demand for and<br>Trust in WBFAT in Their Learning .....                                      | 148 |
| 6.2.2 The Relatively Low Mean of Willingness of Using WBFAT Reflects a Significant<br>Gap Between Students' Cognition of WBFAT and Their Actual Behaviour ..... | 149 |
| 6.2.3 The Variability in Motivation Necessitates that WBFAT Design Addresses both<br>Intrinsic and Extrinsic Motivational Needs of Students .....               | 151 |
| 6.3 Discussion on Theme 2 .....                                                                                                                                 | 152 |
| 6.3.1 WBFAT Enhanced Students' Motivation and Engagement in OEAL & FA .....                                                                                     | 152 |
| 6.3.2 Long-term Use of WBFAT Contributes to Improving the Test Scores of Students'                                                                              |     |

|                                                                                                                                                     |     |
|-----------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Oral English .....                                                                                                                                  | 154 |
| 6.3.3 WBFAT Has a Positive Effect on Improving Students' Oral English Performance<br>in Class .....                                                 | 155 |
| 6.3.4 WBFAT Provided Positive Support for the Development of Oral English Proficiency<br>in Higher Vocational College Students .....                | 157 |
| 6.3.5 WBFAT Had Limited Impact on the Development of Students' Complex<br>Language Skills .....                                                     | 158 |
| 6.4 Discussion on Theme 3 .....                                                                                                                     | 158 |
| 6.4.1 Learners' Growing Demand for Personalised and Dynamically Updated<br>Learning Resources .....                                                 | 159 |
| 6.4.2 The Importance of Integrating Tools Like WBFAT Into Teaching .....                                                                            | 160 |
| 6.4.3 Students' Satisfaction May Be Influenced by the Gap Between Expectations<br>and the Time Needed for Learning Outcomes .....                   | 160 |
| 6.4.4 Inherent Flaws in the Tool Can Undermine Users' Expectations and Satisfaction .....                                                           | 161 |
| 6.5 Chapter Summary .....                                                                                                                           | 161 |
| Chapter 7 Conclusion .....                                                                                                                          | 163 |
| 7.1 Summary of the Study .....                                                                                                                      | 163 |
| 7.2 Academic Contributions .....                                                                                                                    | 165 |
| 7.3 Practical Implications .....                                                                                                                    | 166 |
| 7.4 Limitations of the Study .....                                                                                                                  | 167 |
| 7.5 Future Research Directions and Prospects .....                                                                                                  | 168 |
| 7.6 Chapter Summary .....                                                                                                                           | 169 |
| Bibliography .....                                                                                                                                  | 171 |
| Appendixes .....                                                                                                                                    | 197 |
| Appendix 1 Scales of Questionnaire and Data Analysis in the Pilot Study .....                                                                       | 197 |
| Appendix 2 Item-Total Statistics of the Scales of Questionnaire and Data Analysis in the<br>Survey of Students' Attitudes Towards Using WBFAT ..... | 198 |
| Appendix 3 Raw Data of Students' Engagement of the 14 Variables in OEAL & FA .....                                                                  | 199 |
| Appendix 4 Raw Data of Students' Scores in the Five Oral English Tests .....                                                                        | 201 |
| Appendix 5 Records of Class Observation on Students Oral English Performance .....                                                                  | 203 |
| Appendix 6 Summary of Four Interviews Between the Researcher and the Instructor .....                                                               | 225 |
| Appendix 7 Summary of the Student Focus Group Discussion .....                                                                                      | 235 |
| Appendix 8 Sample of Participants' Consent Form .....                                                                                               | 250 |

## **LIST OF ABBREVIATIONS**

RQ: Research Questions  
EFL: English as Foreign Language  
WB FAT: Web-Based Formative Assessment Tools  
WBLT: Web-Based Learning Tools  
IT: Information Technology  
AIT: Artificial Intelligence Technology  
MOOCs: Massive Open Online Courses  
VR: Virtual Reality  
AR: Augmented Reality  
OEAL & FA : Oral English Autonomous Learning and Formative Assessment  
SD: Standard Deviation  
ZPD: Zone of Proximal Development  
TAM: Technology Acceptance Model  
SDT: Self-Determination Theory  
TEL: Technology-enhanced Learning  
SLT: Situated Learning Theory  
CL: Collaborative Learning  
CSCL: Computer-supported Collaborative Learning  
CE2: College English, Phase 2  
JTC: Jiangxi Teachers College  
EG: Experimental Group  
CG: Control Group  
MOE of the PRC: Ministry of Education of the People's Republic of China  
DESWC: Word count, number of words  
DESSL: Sentence length, number of words, mean  
PCNARp: Text Easability PC Narrativity, percentile  
PCSYNp: Text Easability PC Syntactic simplicity, percentile  
PCREFp: Text Easability PC reference, percentile  
PCDCp: Text Easability PC Deep cohesion, percentile  
CRFAOa: Argument overlap, all sentences, binary, mean  
CRFCWO1: Content word overlap, adjacent sentences, proportional, mean  
LSASS1: Latent Semantic Analysis, sentences, mean  
SMTEMP: Temporal cohesion, tense and aspect repetition, mean  
WRDFAMc: Familiarity for content words, mean  
PCCNCp: Text Easability PC Word concreteness, percentile  
WRDCNCc: Concreteness for content words, mean  
LDTTRc: Lexical diversity, type-token ratio, content word  
LDTTRa: Lexical diversity, type-token ratio, all words  
WRDHYPn: Hypernymy for nouns, mean  
WRDHYPv: Hypernymy for verbs, mean  
RDFRE: Flesch Reading Ease  
RDFKGL: Flesch-Kincaid Grade Level  
RDL2: Coh-Metrix L2 Readability  
IPA: Importance-Performance Analysis

## LIST OF TABLES

|                                                                                                                                                                |     |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Table 1 The Chart of China's Vocational Education Structure .....                                                                                              | 22  |
| Table 2 Distribution of the Publications on the Frontiers and Hot Issues of China's ..... 32                                                                   | 32  |
| Education Research (2019-2022) .....                                                                                                                           | 32  |
| Table 3 Criteria for the Quality Appraisal of the Literature .....                                                                                             | 38  |
| Table 4 Year Distribution of the Included Publications (N=130) .....                                                                                           | 38  |
| Table 5 Statistical Analysis of Participants, Study Duration, Contexts, Disciplinary Distribution, and Research Methods Within the Included Publications ..... | 41  |
| Table 6 The Distribution of the Included Literature in Terms of Participants, Study ..... 41                                                                   | 41  |
| Duration, Context, Disciplinary Distribution, and Research Methods .....                                                                                       | 41  |
| Table 7 Theoretical Basis in the Included Literature .....                                                                                                     | 47  |
| Table 8 WBFAT, MFAT Utilised and Related Findings in the Included Literature .....                                                                             | 52  |
| Table 9 The Scheme of CE2 Learning in JTC .....                                                                                                                | 85  |
| Table 10 Likert-scale Questions in the Survey on Students' Attitudes Towards Using WBFAT .....                                                                 | 93  |
| Table 11 Demographic Statistics of the Answers to the Questions in the Survey of Students' Attitudes ..... 94                                                  | 94  |
| Table 12 Intervention Measures for OEAL & FA in the Experimental Study .....                                                                                   | 95  |
| Table 13 Selected Variables of Students' Performances in OEAL & FA .....                                                                                       | 96  |
| Table 14 Indicators of Coh-Metrix 3.0 Linguistic Profiling Selected in This Study .....                                                                        | 98  |
| Table 15 The Schedule of Class Observation .....                                                                                                               | 101 |
| Table 16 The Schedule of Instructor Interviews .....                                                                                                           | 102 |
| Table 17 Scales in the Survey of Students' Perceptions of WBFAT .....                                                                                          | 103 |
| Table 18 The Instructor Participant's Contribution to the Present Study .....                                                                                  | 105 |
| Table 19 Design of Data Sources and Data Analysis in the Present Study .....                                                                                   | 111 |
| Table 20 Demographic Statistics of the Scales in the Survey of Students' Attitudes Towards Using WBFAT .....                                                   | 113 |
| Table 21 Reliability Test of the Scales in the Survey of Students' Attitudes Towards Using WBFAT .. 114                                                        | 114 |
| Table 22 KMO and Bartlett's Test of the Scales in the Survey of Students' Attitudes Towards Using WBFAT .....                                                  | 114 |
| Table 23 Descriptive Analysis of the Scales in the Survey of Students' Attitudes Towards Using WBFAT .....                                                     | 114 |
| Table 24 Descriptive Statistics of the Variables in Students' Engagement in OEAL & FA .....                                                                    | 119 |
| Table 25 Independent Sample T-test of the 14 Variables in Students' Engagement in OEAL & FA .... 121                                                           | 121 |
| Table 26 Descriptive Analysis of Students' Average Scores on the Five Oral English Tests .....                                                                 | 122 |
| Table 27 Descriptive Analysis of the Values of the 20 Indicators of Linguistic Characteristics in Students' Five Audio Transcripts .....                       | 125 |
| Table 28 T-tests of the 20 Linguistic Characteristic Indicators' Mean in Students' Five Audio Transcripts .....                                                | 128 |
| Table 29 The Node Coding Frequency of the Six Class Observations .....                                                                                         | 131 |
| Table 30 Query Results of The Matrix Coding of the Six Class Observations .....                                                                                | 132 |
| Table 31 Reliability Test of the Scales in the Survey of Students' Perceptions of WBFAT .....                                                                  | 136 |
| Table 32 KMO and Bartlett's Test of the Scales in the Survey of Students' Perceptions of WBFAT .... 136                                                        | 136 |
| Table 33 T-tests of the Four Dimensions of the Survey of Students' Perceptions of WBFAT .....                                                                  | 137 |
| Table 34 The Node Coding Frequency of the Student Focus Group Discussion .....                                                                                 | 142 |
| Table 35 The Matrix Coding Query Results of the Student Focus Group Discussion .....                                                                           | 143 |

## LIST OF FIGURES

|                                                                                                                             |     |
|-----------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 1 The Flow Chart of the Procedure in Literature Selection and Screening .....                                        | 35  |
| Figure 2 Distribution of the Included Publications by Country or Region .....                                               | 40  |
| Figure 3 An Explanatory Sequential Design of Mixed Methods in this Study .....                                              | 77  |
| Figure 4 A Screenshot of the Digital Resource Base on the Web-Platform of FiF .....                                         | 86  |
| Figure 5 A Screenshot of the FiF's Functions in the Comprehensive Assessment for Oral English Test .....                    | 87  |
| Figure 6 A Screenshot of the FiF's Functions in Rating and Feedback for Oral English Test .....                             | 87  |
| Figure 7 A Screenshot of FiF's Data Panels .....                                                                            | 88  |
| Figure 8 The Research Procedure of This Study .....                                                                         | 89  |
| Figure 9 A Snapshot of Analysis of Audio Transcripts of Students' Oral Tests via Coh-Metrix 3.0 .....                       | 98  |
| Figure 10 Quality Triangulation Designed in the Present Study .....                                                         | 108 |
| Figure 11 The Trend of Student's Average Scores of Oral English Tests by Group .....                                        | 124 |
| Figure 12 Trends of the Values of the 20 Indicators of Linguistic Characteristics in Students' Five Audio Transcripts ..... | 129 |
| Figure 13 The Matrix chart: Number of Codes by Group in Class Observation .....                                             | 132 |
| Figure 14 Quadrants of the Scales in the Survey of Students' Perceptions of WBFAT .....                                     | 139 |
| Figure 15 The Word Cloud of the Student Focus Group Discussion .....                                                        | 139 |

## **Chapter 1 Introduction**

This chapter starts with a quick overview of the educational informatisation in China, as well as the changes and impacts it has caused on teaching, learning, and assessment, providing the context for the present study. Subsequently, the researcher defines the term ‘Web-Based Formative Assessment Tools (WBFAT)’ and makes a brief introduction to FiF, the WBFAT used in this study. The chapter then elaborates on the purpose, research questions, and significance of the study. Following this, the methodology in this study is discussed, including the methods of data collection, analysis, integration, and reporting. Lastly, the chapter outlines the structure and content of each chapter, offering a clear presentation of the framework for this study.

### **1.1 Research Background**

#### **1.1.1 The Popularisation of Educational Informatisation in China**

In 1999, the Central Committee of the Communist Party and the State Council of the PRC (1999) put forth the directive to accelerate the modernisation of educational technology. This was considered to be the inception of educational informatisation in China, as it was the first time it was mentioned in official policy documents. Over the past two decades, significant progress has been made in the area of primary, higher, and vocational education in terms of the informatisation of education.

Firstly, it is evidenced by remarkable advancements in the development of informatisation infrastructure. The official report (MOE of the PRC, 2023) indicates that all Chinese primary and secondary schools possess internet connectivity, with approximately 75% of schools equipped with wireless internet and 99.9% achieving a minimum internet speed of 100 Mbps. Moreover, 99.5% of educational institutions now feature multimedia classrooms, integrating information technology (IT) and artificial intelligence technology (AIT) to enhance teaching and learning. In higher education institutions, including vocational colleges, the integration of technology into teaching and learning has become increasingly common. They have established smart management systems based on technology for instruction, learning, and evaluation. Technology enables comprehensive data integration and resource sharing within educational processes, thereby establishing a robust base for the progression of educational informatisation .

Secondly, it is demonstrated by the vast online digital resources and a large number of users. According to Yang (2024), as of December 2024, more than 30 national online platforms have been developed, offering more than 97,000 Massive Open Online Courses (MOOCs), with a total of 1.39 billion learners, through which on-campus students had collectively earned 440 million academic credits. Additionally, China's Smart Higher Education Platform had developed 31,000 online courses, with 78,000 teachers and 93 million users registered from 183 countries and regions. These figures illustrate the rapid pace and astonishing breadth of the integration of IT and AIT into China's education.

### **1.1.2 Changes and Impacts on Education Caused by the Informatisation**

Educational informatisation is reshaping both the structure and substance of education while also affecting multiple facets of the educational process. With the support of IT and AIT, the process of accessing learning resources has now become more convenient and efficient, and the mode of learning has also become more flexible and autonomous (Panadero, 2017). Students can easily access a vast amount of online learning materials anytime and anywhere and organise their study according to their own needs, free from the constraints of time and space. In addition, IT and AIT promote personalised and interactive learning (Elmiri et al., 2024). Students can tailor their learning experiences to better align with their individual strengths, weaknesses, and interests. Adaptive learning platforms, for example, can track students' progress and offer customised learning paths, providing targeted resources and activities based on their performance (Spector, 2014), which not only boosts students' engagement and motivation (Panadero, 2017; Boud & Falchikov, 2007) but also enhances their ability to learn at their own pace, allowing for a deeper understanding of the subject matter. Furthermore, the integration of technology-based collaborative tools encourages communication and teamwork among students, creating a more interactive and supportive learning environment. As a result, the traditional, one-size-fits-all teaching model gives way to a more dynamic and learner-centred approach, fostering students' ability of self-regulation, independent thinking, and problem-solving, which is essential for lifelong learning.

The use of technology-based tools has also significantly improved the visualisation and interactivity of teaching, enhancing the effectiveness and outcomes of classroom teaching. Through the use of online teaching platforms, virtual reality (VR), and

augmented reality (AR) technologies, teachers can craft immersive experiences for students, enabling them to explore microscopic phenomena, marvel at the wonders of the universe, or experience real-world work scenarios, transforming previously abstract and complex concepts into vivid and engaging content. Furthermore, interactive whiteboards and online platforms allow teachers to deliver materials in real time, initiate polls, and organise group discussions, resulting in heightened attention and a deeper understanding and retention of knowledge, which improves students' engagement in class.

Moreover, the integration of technology into education facilitates a shift from teacher-centred summative assessments to student-centred, formative, and diversified assessments for learning. With the support of IT and AIT, teachers are able to utilise online platforms and mobile applications to gather real-time learning data, providing a more precise understanding of students' progress. This data-driven approach enables continuous monitoring, allowing for timely interventions that ensure students receive the necessary support throughout their learning journey. By emphasising formative assessments, such as quizzes, peer reviews, and interactive activities, teachers can assess students' development on an ongoing basis, rather than relying solely on final examinations or tests. This shift promotes a more personalised and holistic approach to assessment, taking into account individual learning styles, strengths, and areas for improvement (Hattie & Timperley, 2007; Kaya-Capocci et al., 2022) and encourages students to engage actively in the assessment process and help them reflect on their learning and set their own goals. The shift towards a technology-enhanced, student-centred approach not only supports continuous academic development but also plays a crucial role in cultivating essential skills for lifelong learning.

However, despite the numerous advantages of educational informatisation, teachers and students face challenges in the application of these technologies, such as the discomfort with new technologies, limitations of technologies, and the digital divide (Francom, 2019; Bešić & Delvić, 2022). One of the significant barriers to the implementation of educational technologies is the discomfort and resistance teachers and students experience when using new technologies. Many educators, especially those who are not digitally proficient, face challenges in integrating these tools into their teaching practices. The apprehension can stem from a lack of training or from being overwhelmed by the adoption of a new system (Bešić & Delvić, 2022). Similarly, students who are not

accustomed to digital learning platforms may find it difficult to engage with them effectively. Research by Tondeur et al. (2017) highlight that the successful integration of technology in education heavily depends on teachers' comfort with digital tools, as their attitudes and confidence significantly influence the use of such technologies.

Another is the inherent limitations of the technologies and tools themselves. Despite the rapid advancements in educational technologies, many tools still face functional and technical constraints. For example, tools for formative assessment might lack the ability to offer adaptive feedback that caters to individual learning needs or may not be sufficiently interactive to keep students engaged (Çekiç & Bakla, 2021).

Furthermore, certain tools may require high-performance devices or specialised software that may not be available in all educational settings, especially in lower-income regions. Such limitations hinder the widespread effectiveness of digital tools in enhancing educational practices.

Perhaps the most pressing issue is the digital divide, which limits students' ability to fully engage with educational technologies and contributes to the widening gap in educational outcomes as well as the unequal access to technology based on socioeconomic factors, geographical location, and other disparities in rural or economically disadvantaged areas (Gorski, 2005; Anthony & Padmanabhan, 2010; Liu, 2021; Goin Kono & Taylor, 2021). The digital divide was considered to exacerbate educational inequalities, as those with access to technology can benefit from the enhanced learning experiences provided by digital tools, while others are left behind.

## **1.2 Definition of WBFAT and Introduction to FiF**

### **1.2.1 Definition of WBFAT**

As IT and AIT continue to integrate with education, many technology-supported platforms and applications for teaching, learning, and assessment have emerged, each with varying names, functions, and emphases. Some of them focus on online teaching, and others highlight assessment for learning. While the names and functions of these platforms and applications may vary, they share common characteristics, namely that they fundamentally rely on IT and AI technologies, and most of them offer a combination of features such as providing learning resources, online interaction, automated assessment and feedback, and backend data management, supporting classroom teaching interaction,

learner autonomy, learning evaluation, as well as data management. To facilitate understanding, this study will collectively refer to all these technology-supported teaching, learning, and formative assessment platforms and applications as WBFAT.

### **1.2.2 Introduction to FiF**

FiF, a WBFAT for oral English teaching and formative assessment, was developed by iFlytek. It has several key features. First, it offers a vast array of digital learning resources, covering a wide range of fields such as academics, life, career, travel, culture, science, and sports, catering to learners with varying proficiency levels and interests, and allows learners to create a personalised resource base according to their needs. Secondly, it provides real-time automated scoring and feedback, identifying errors in students' oral English and offering corrective suggestions, which helps learners improve in terms of accuracy, fluency, and completeness. Thirdly, FiF features a powerful data statistics and analysis panel, which tracks students' learning progress in detail. By collecting data such as learning time, task completion, peer evaluations, and interactions, the platform provides comprehensive insights into students' engagement in autonomous learning and the development of oral English proficiency. Finally, it supports online interaction and collaboration among students via peer evaluation, online chats, and message boards, promoting peer interaction and cooperation.

### **1.3 Purpose of the Study**

This study is an exploratory investigation into the impact of WBFAT on the development of oral English proficiency among students in Chinese higher vocational colleges. The study aims to examine students' attitudes towards using WBFAT, its influence on the development of their oral English skills, and their satisfaction with, as well as perceptions of, the importance of WBFAT after its use. Therefore, this study will provide insights into the implementation of autonomous oral English learning and e-learning for Chinese higher vocational students. Additionally, it encourages the developers of WBFAT to refine the platforms and applications, better catering to the specific learning characteristics and needs of Chinese higher vocational students.

### **1.4 Statement of Research Problem**

With the widespread and deep integration of IT and AIT in education, technology-supported teaching, learning, and assessment have become a new attraction in

the field of education and a focus of academic research. To investigate the impact of WBFAT on students' learning effectiveness, it is first necessary to investigate students' attitudes, as attitudes are considered as the basis of engagement and motivation (Miller et al., 1988; Brophy, 2004). Positive attitudes can drive greater engagement, whereas negative perceptions can lead to resistance or disengagement, thus affecting the effectiveness of the tools. Second, students' academic achievement serves as a concrete indicator of the effectiveness of these tools in enhancing students' learning processes. If students who regularly use the WBFAT show significant progress in academic outcomes, it can provide strong evidence of the effectiveness of the tool in supporting student learning and skill development. Furthermore, assessing student satisfaction with the WBFAT and perceptions of its importance is equally critical, because satisfaction surveys can help to understand whether students find the tools helpful, easy to use, and useful for their learning, and investigation of students' views on the importance of the tools helps educators and researchers refine and enhance the tool to better serve students' needs.

To further explore the above issues, this study focuses on the rapidly growing group of EFL students in Chinese higher vocational colleges, exploring the impact of WBFAT on their oral English proficiency development through three themes and six questions. The themes include students' attitudes towards using WBFAT, their progress in oral English proficiency after using WBFAT, and their perceptions of satisfaction and the importance of WBFAT.

#### **1.4.1 Theme One: Students' Attitudes Towards Using WBFAT**

##### ***Q1. What are vocational college students' attitudes towards using WBFAT?***

This research question aims to investigate students' attitudes towards WBFAT and their acceptance of using WBFAT in their learning process at vocational colleges in China.

The attitudes of users towards WBFAT directly affect their efficacy and motivation to utilise these tools (Deci & Ryan, 2000). Previous studies have noted that students' acceptance of online learning tools is often related to the usability of the tool, the personalisation of services, and the applicability of the technology (Venkatesh et al., 2003; Cheung & Vogel, 2013; Sun et al., 2008). Whether the tools can meet students' learning needs and improve learning efficiency is a key factor in determining their attitudes. If students perceive that the tools provide real value in their learning, such as offering instant

feedback, enhancing autonomous learning capabilities, and improving academic performance, their acceptance and willingness to use the tools typically increase (Wahdain & Ahmad, 2014; Venkatesh et al., 2003). Therefore, the design of WBFAT to meet students' learning needs and provide effective feedback significantly influences students' attitudes.

Previous studies have examined teachers' and students' attitudes toward using the WBFAT and have found that overall teachers and students have positive attitudes, while some studies have highlighted students' negative or neutral attitudes stemming from concerns related to the technology's usability, data privacy disclosure (Reidenberg, 2018; Schlosser et al., 2022), or the accuracy of the tool's evaluations, particularly for complex tasks like assessing creativity or critical thinking (Almousa, 2023; Baker & Yacef, 2009). Moreover, as discussed in the literature review (see Chapter 3), most of the previous studies focus on students at primary, secondary schools and universities, there is a paucity of thorough surveys and extensive analyses regarding attitudes towards the WBFAT, specifically among vocational students.

#### **1.4.2 Theme Two: Students' Performance of OEAL & FA and Progress in oral English After Using WBFAT**

##### ***Q2. What is students' performance of oral English autonomous learning?***

This research question seeks to understand students' time investment, task fulfilment, modifications in learning strategies, and peer interaction throughout the process of OEAL & FA to assess the role of WBFAT in fostering students' interest in learning and enhancing students' motivation and self-regulation in the process of students' autonomous learning through controlled analyses of the differences between the performance of the experimental group and the control group.

Motivation and engagement are crucial factors that significantly influence the effectiveness of autonomous learning. Motivation serves as the driving force behind learners' willingness to initiate and sustain their independent learning activities. It encompasses both intrinsic elements, such as personal interest and the satisfaction derived from acquiring knowledge, and extrinsic factors, such as rewards and external recognition. Engagement, on the other hand, reflects the learners' active involvement in the learning process, including their ability to set meaningful goals, maintain focus, and apply strategies to overcome challenges. Zimmerman (2002) and Deci & Ryan (2000) suggest that high

engagement is closely linked to self-regulatory behaviours, such as actively seeking learning resources, setting goals, and adjusting strategies. Engagement is a key predictor of learning outcomes in mobile and online learning environments, as frequent and quality participation in activities on platforms reflects students' interest, motivation, and learning strategies (Lei et al., 2018; Gunuc, 2014; De Villiers & Werner, 2016; Mizani et al., 2022).

Together, motivation and engagement create a foundation for learners to take responsibility for their autonomous learning, enabling them to plan, execute, and reflect on their learning activities effectively. Without these two key components, autonomous learning is likely to lose its potential for fostering autonomy and achieving meaningful educational outcomes.

***Q3. Is there significant progress in students' oral English proficiency after their use of WBFAT over time?***

This research question aims to comprehensively verify whether students' oral English proficiency has improved by analysing their scores from multiple oral English tests, as well as language proficiency indicators such as descriptive indicators, principal component scores for text readability, referential cohesion, latent semantic analysis, lexical information, and readability.

As will be discussed in Chapter 3, many studies have shown that web-based or mobile-based learning and assessment tools have significant positive impact on improving student performance, promoting self-regulation (Baleni, 2015; Chen, 2016; Li et al., 2008; Liu et al., 2018; Marín & Perez Garcias, 2016), and improving academic achievement (Andrade & Heritage, 2018; Lai et al., 2018; Veng, 2023). Most of these studies have explored the tool's impact on academic performance from a single perspective. Akpınar & Çetin (2023), for example, used pre-tests, post-tests, and follow-up tests, employing stratified random cluster sampling and conducting statistical analysis to examine the effectiveness of Student/Teacher Support Systems (ÖDS) and Education Information Networks (EBA) on improving students' academic performance. Some studies have employed survey methods (Elmahdi et al., 2018) or solely focused on literature reviews and theoretical analysis (Fernandez, 2003). Some studies (Andrade & Heritage, 2018) have combined qualitative methods, such as classroom demonstrations, student project analysis, and behavioural observations, with quantitative methods such as test results and surveys to

validate the effects of tools like FACT, ASSISTments, SimScientists, DIAGNOSER, and online assessment tools (e.g., iTunes U, Notability, Socrative, AirPlay) on students' learning motivation and academic performance. Shirley (2009) uses multiple data sources, such as classroom observations, teacher and student interviews, and surveys, to verify the effectiveness of the TI-Navigator system in enhancing classroom interaction and improving student academic outcomes. However, overall, there are relatively few studies that employ a mixed-methods approach and explore WBFAT's effects from multiple perspectives.

This study employs a mixed-method design, combining qualitative and quantitative approaches to investigate the WBFAT's impact on students' oral English proficiency from new perspectives. Firstly, the FiF scores and human ratings of the five oral English tests from both the experimental and control groups will be analysed via analysis of the mean and standard deviation of the data, as well as t-tests to determine the impact of WBFAT on students' oral English performance. Additionally, the research utilises Coh-Metrix 3.0 to analyse language features of students' audio transcripts extracted from their five oral English tests, including 20 indicators such as word count (DESWC), sentence length (DESSL), text readability, syntactic simplicity (PCSYNp), referential cohesion, and latent semantic analysis (LSA), to further verify the WBFAT's impact on the development of students' oral English proficiency.

***Q4. Is there any improvement in students' oral English performance in class after using WBFAT over time?***

This research question further explores the progress of students' oral English proficiency in real classroom settings after 20 weeks of using WBFAT.

Improvement in oral English proficiency is not only reflected in academic achievement but also in students' ability to use English effectively and fluently in real-life situations and practical communication (Zimmerman, 2002; Brown, 2014), which is, for students, typically reflected in their use of oral English in class (O'Malley, 1990). Nicol and Macfarlane-Dick (2006) argue that interaction, discussion, and the opportunity to use language in real situations are pivotal factors in enhancing language proficiency. Students' classroom performance, particularly their ability to express themselves orally, is an important indicator of their language learning progress. Hattie and Timperley (2007)

emphasise the crucial role of feedback in the learning process, noting that feedback not only improves academic performance but also boosts students' engagement and confidence in the classroom.

Therefore, it is imperative to concentrate on students' performance of oral English in class to evaluate whether their oral English proficiency has been effectively improved, wherein their fluency, accuracy, and flexibility can be continuously and clearly observed.

#### **1.4.3 Theme Three: Students' Perceptions of WBFAT**

***Q5. What are students' perceptions of the satisfaction of WBFAT? &***

***Q6. What are students' perceptions of the importance of WBFAT?***

These two research questions aim to understand the satisfaction of students in the experimental group after using FiF for about 20 weeks, as well as their perceptions of the importance of FiF in its main functions, such as automatic scoring and feedback.

Students' satisfaction with and perception of technological tools significantly affect their use and effectiveness. When students receive valuable feedback from WBFAT, they typically show greater willingness to use it and enhanced motivation to learn. Buragohain et al. (2023) find that in e-learning environments, ESL engineering students' satisfaction with online learning tools is closely related to the tool's usability, functionality, and the quality of feedback provided. Gikandi et al. (2011) also point out that students' acceptance and satisfaction with online assessment tools depend heavily on their usability, functionality, and the timeliness and quality of the feedback provided. Moreover, when students evaluate the effectiveness and importance of these tools, they often base their judgements on the tool's impact on their learning outcomes, particularly in terms of academic performance and motivation. Martin & Kummar (2018) emphasise that students' acceptance and perception of the importance of e-learning tools generally depend on whether the tool can significantly improve their academic performance and provide effective feedback, which in turn influences their learning motivation and self-efficacy. If WBFAT helps students improve their oral English skills and provides tangible support during learning, students will regard it as an important tool for their studies. Additionally, students' perceptions affect not only the frequency of tool use but also the effort they invest in their learning process.

Therefore, understanding students' attitudes towards the WBFAT is critical to identifying potential barriers and opportunities for improvement.

### **1.5 Significance of the Study**

There is substantial theoretical and practical value in this study. Theoretically, this study adopts a mixed research method combining qualitative and quantitative approaches, including a survey of students' attitudes towards using WBFAT, a quasi-experimental design to observe the development of students' oral English proficiency after using WBFAT, and a survey of students' perceptions of the satisfaction and importance of WBFAT, to explore the impact of WBFAT on the development of higher vocational students' oral English proficiency. This study, thus, will provide new insights and rich empirical evidence for the existing theories of technology-supported oral English learning, and technology-supported educational assessment theories.

Practically, this study will provide guidance and insights into the learning, teaching, and assessment for higher vocational students and instructors in China. Firstly, students will be encouraged by the findings in this study that participants had made progress in their learning engagement, self-confidence, and oral English proficiency after using FiF, which will stimulate more students to further accept and explore the WBFAT's supportive role in the learning of other subjects. Secondly, the students' performance and development of oral English proficiency in this study can prompt instructors to make timely interventions and adjustments to their teaching strategies in using the WBFAT to support classroom teaching and students' autonomous learning, so as to better facilitate the positive effects of the WBFAT in teaching and students learning. In addition, this study will also promote the decision-makers in higher vocational colleges to select the WBFATs that have been proven to be effective in improving student learning in studies and that are generally approved by most of the students.

Finally, this study holds social significance. On the one hand, with the rapid development of IT and AIT, cultivating high-quality technical talent capable of meeting the demands of the digital age has become a crucial task for vocational colleges. Through this study, the effectiveness of WBFAT in learning, teaching, and assessment will be better understood, providing impetus for education reform and promoting the further popularisation of educational informatisation, and enhancing the students' comprehensive

competitiveness in higher vocational colleges. On the other hand, this study will also provide guidance and inspiration for social training and lifelong learning for other groups.

## **1.6 Research Methodology**

This study adopts a mixed-methods research approach, combining both quantitative and qualitative methodologies. On the one hand, quantitative data were collected through a survey on students' attitudes towards using WBFAT, records of students' performance of their participation during the process of OEAL & FA, scores from five oral English tests, the analysis of the linguistic indicators of students' audio transcripts from the five oral English tests, and a survey on students' perceptions of satisfaction and importance of the WBFAT. Numerical analyses were conducted on the quantitative data to enable a more accurate understanding of the participants' behavioural characteristics, thus providing a basis for subsequent analysis, interpretation, and decision-making. On the other hand, qualitative data, including records of students' oral English use in the classroom, interviews with the instructor, and a student focus group discussion, were collected to further supplement and elucidate the findings derived from the quantitative data.

### **1.6.1 Data Collection**

In order to answer the research questions formulated in this study and to reach the research objectives, the researcher collected a total of eight sets of data in three phases: the survey research phase, the experimental research phase, and the post-experimental research.

In order to answer RQ1, the researcher administered a questionnaire consisting of 21 five-point Likert scale items to the students of the selected higher vocational colleges prior to the experimental study in order to assess the students' overall attitudes and acceptance tendencies towards the use of the WBFAT.

To answer RQ2, the researcher collected data on students' performance in the experimental and control groups during the process of 20-week OEAL & FA, exploring whether the WBFAT had a positive effect on the students' engagement and task fulfilment in the process of autonomous learning.

To answer RQ3, the scores of five monthly oral English tests from the two groups of students were gathered to explore whether WBFAT has a positive role in facilitating students' oral English performance. Further, the researcher selected the audio transcripts of

the free expression part from the five oral tests and used Coh-Metrix 3.0 to characterise the linguistic indicators of their expressions.

In order to respond to RQ4, the researcher entered the classroom to observe and record students' performance of oral English in class 6 times and conducted four interviews with the instructor, which were used to further corroborate the effect of the WBFAT on the development of the students' oral English proficiency.

To answer RQ5 & 6, the researcher conducted a questionnaire consisting of five Likert scale items to investigate students' satisfaction with the FiF and the perceived importance of its functions, as well as organised a focus group discussion with 22 students from the experimental group.

### **1.6.2 Data Analysis and Integration**

#### ***Integration of Qualitative and Quantitative Data***

Integrating qualitative and quantitative data allows for a more comprehensive and nuanced understanding of the research findings, as it combines the strengths of both data types (Creswell & Clark, 2017; Tashakkori, 2020). In this study, the researcher applied descriptive statistics, correlation analysis, and *t*-tests between-groups to quantitative data using SPSS 26.0, providing a clear picture of the measurement indicators and facilitating the comparison of results across groups. On the other hand, qualitative data were analysed using NVivo 11, where themes were identified and categorised through coding and thematic analysis, which allows for a deeper interpretation and understanding of the quantitative data and provides contextual explanations that enrich the overall findings of the study.

#### ***Integration of Analytical Methods***

The integration of multiple analytical methods ensures a well-rounded and robust examination of the research questions. Quantitative analysis methods, including descriptive statistics, correlation analysis, and *t*-tests, were used to measure and compare the performance of students across different conditions. These statistical methods help to quantify the impact of WBFAT on both the experimental and control groups, both before and after its implementation, thus offering clear evidence of the tool's effects. While the qualitative methods, such as coding, thematic analysis, and word cloud analysis, provide rich, context-driven explanations for the quantitative findings and demonstrate students'

behavioural patterns, learning strategies, and the emotional responses towards the WBFAT during the learning process. Therefore, the integration of distinct yet complementary analytical approaches allows for a more comprehensive understanding of the research topic, enhancing the validity and depth of the study's conclusions, overcoming the limitations inherent in relying on a single approach, and providing more robust interpretations for the findings.

### ***Integration of findings***

The integration of findings seeks to synthesise the results from multiple data sources and analytical methods into a cohesive and meaningful conclusion. In this study, the research employs a range of analytical techniques such as descriptive statistics, correlation analysis, *t*-tests, and thematic analysis, each contributing to different facets of the study. By combining the findings from both quantitative and qualitative analyses, the researcher presents a comprehensive view of how WBFAT affects students' oral English proficiency, autonomous learning, and classroom participation, which not only strengthens the scientific rigour of the study but also ensures that the findings are accurate and well-supported by diverse types of evidence and provides a clear framework for understanding the impact of WBFAT on students' learning outcomes as well.

### **1.6.3 Data Reporting and Interpretation**

Data reporting and interpretation is a crucial step to deliver and promote the findings of a study effectively, wherein the findings are presented in a clear and systematic way and in-depth interpretation and discussion from diverse dimensions to ensure the accuracy and comprehensibility of the research findings (Creswell & Clark, 2017).

#### ***Data reporting***

On the one hand, the researcher clearly reported the findings of groups of quantitative data collected in chapter 5 through tables, charts, and graphs, as well as textual descriptions, which further enhances the presentation and readability of the results and helps the readers to intuitively understand the relationship between different variables and their changing trends. On the other hand, qualitative data were coded and thematically analysed to provide richer contextual information for the quantitative data, offering a deeper understanding of students' learning performance and attitudes in terms of emotions and behaviours.

### ***Interpretation and discussion of findings***

The researcher provides a thorough explanation and multidimensional discussion of the findings of this study in chapter 5, analysing in depth the trends and patterns in the data as well as their underlying causes. Comparisons were also made with existing literature to verify whether this study supports or challenges existing theories and hypotheses. Meanwhile, the researcher honestly discussed the possible biases, confounding variables, and sample size limitations in this study, which helps readers to correctly interpret the findings and also provides directions for improvement in future research. The researcher made an outlook for future research on WBFAT-enhanced learning and assessment to provide some enlightenment for further study.

### **1.7 Organisation of the Dissertation**

This dissertation consists of seven chapters. In chapter 1, the overall context of this study was introduced. The research background, the definition of WBFAT, an introduction to FiF, research objectives, research questions, research significance, methodology, and the structure of the dissertation are outlined in general, providing a guide for the entire study. Chapter 2 traces the historical evolution of China's education, providing a macro-level context for this study. Chapter 3 is the literature review, where the researcher demonstrated the process of literature selection, literature analysis and discussion, and made a comprehensive review of related theories of learning and assessment, laying a solid theoretical foundation for the study. Chapter 4 details the research methodology and process, where the mixed-methods design, participants, research settings, research procedures, and methods adopted were elaborated. Chapter 5 presents the data report and discussion of the findings. Chapter 6 revisits the research objectives and questions and discusses students' attitudes towards the WBFAT and students' behaviours, outcome differences in the use of the WBFAT, and their influencing factors in the light of the relevant literature. Chapter 7 makes a summary and acknowledges the limitations of the study, and offers rational suggestions for future research directions as well.

## **Chapter 2 General Introduction to Chinese School Education**

This study is laid in the context of Chinese school education which possesses its own features in the aspects of the evolution process (Li, 2021), education systems (Zhu, 2020; Gao, 2020), educational philosophy (Zhong, 2022; Feng, 2019), learners' characteristics (Yu et al., 2022; Liu & Wang, 2021), teaching, and assessment (Huang, 2022; Huang, 2020). In order to make the research process and findings of this study better understood, it is of great significance to illustrate the context of Chinese school education. Therefore, in this chapter, the evolution of China's school education from ancient times to the present will first be outlined. Then, a detailed exploration of the multifaceted development of contemporary China's education will be outlined, including the improvement of the school educational system, quality, disparities, and the widespread educational informatisation. Subsequently, the research will elaborate on the philosophy of Chinese school education, including teaching objectives, curricula, teaching methods, teacher-student relationships, teaching content and materials, evaluation methods, and school management. Their historical evolution and modern transformations will also be traced in this section. Finally, a summary of the current research landscape of school education in China will be presented, reviewing existing findings and key emerging issues.

### **2.1 The Origin, Heritage, and Transformation of Chinese School Education**

Chinese school education has undergone a profound and multifaceted evolution over thousands of years, adapting to societal transformations while maintaining a strong cultural foundation. Adopting the perspective that 'the history of education is shaped by the model of social development' (Yang, 1995), scholars generally divide Chinese school education into three distinct stages: ancient education during the slave and feudal periods, modern education in the Republic of China era, and contemporary socialist education under the leadership of the Communist Party of China.

#### **2.1.1 Ancient School Education: Development and Influence**

On the grounds of records in ancient works, *The Book of Rites* and *Mencius*, the earliest Chinese school took shape in the Xia Dynasty, the first slavery dynasty in the 21st century BC (Sun, 2019). During more than 3000 years of several slave dynasties and dozens of feudal dynasties, Chinese ancient school education had been gradually developed in terms of education system, teaching management, educational content, and

educational concepts. Researchers have fully affirmed that Chinese ancient school education played a vital role in building human morality and ethics (Zeng, 2019), transforming social traditions, spreading civilisation, maintaining governance stability, promoting social development (Liu & Sun, 2020; Dong, 2010), and inheriting excellent Chinese cultures (Li & Chen, 2015), but condemned its negative side as well in imprisoning people's freedom of thought and inhibiting the development of human innovation ability for its strong nature of hierarchy division and prolonged conformism (Sun, 2021; Ma, 2002; Shen, 1996).

Confucianism and its works, represented by Confucius and Mencius, were the main content of education in Chinese feudal society, which advocated the educational philosophy of 'the nature of human to be followed, virtue cultivation, and then talents development' (Ding & Sun, 2020), which led the feudal education in ancient China for thousands of years. The value of Confucian educational concepts is the unity of benevolence and wisdom, pursuing the spiritual realm of teleology that is ethical but trans-ethical, quasi-aesthetic but trans-aesthetic (Sun, 2021). The advocacy of respecting teachers, teaching by setting a good self-example, caring for learners, equality between teachers and students, being patient in teaching, being good at inspiring, teaching in accordance with students' aptitude, and combining learning and thinking (Li et al., 2022; Wang, 2019) in the *Analects of Confucius* is considered to be the worthy treasure of China's educational history, as it has universal value, not only having an important influence in ancient time but also of great significance nowadays after thousands of years. To study Confucius' educational philosophy by starting from practical problems helps us be inspired (Liu, 2022; Zuo, 2020).

### **2.1.2 The Transition to Modern Education**

The modern schools in China appeared in the 1860s. The reform of educational policies in the late Qing dynasty promoted the transition of China's education from ancient education to modern (Chen, 2022) on account of the transformation of social forms and the introduction of western culture and new trends of educational thoughts. The modern schools could be traced back to the foreign language schools, military schools, and technical schools in the 1860s. Then, as China moved to bourgeois government from feudalism in 1912, school education was comprehensively reformed, and new educational

policy of ‘development of the four educations in morality, wisdom, physics, and aesthetics’ was proposed (Tian, 2016; Wu & Fan, 2011). Chinese school education has gradually entered the stage of individualisation, popularisation, pragmatism, and rationalism.

After the foundation of the People’s Republic of China, Chinese school education entered a new stage of development and made remarkable achievements, especially after the implementation of the policy of ‘Reform and Opening’ in 1978, which propelled Chinese education towards modernisation, global integration, and future-orientated progress. Since then, the philosophy of school education with ‘all-round’ development of morality, intelligence, physics, aesthetics, and labour spirit has been systematically constructed gradually, promoting the coordinated and unified development of education and human beings (Xue & Gao, 2023; Zhang & Wang, 2022). Great achievements have been made in the level of education popularisation, education fairness, education quality, education service ability, the facilities, teaching staff development, and education opening (MOE of PRC, 2022a).

### **2.1.3 Future Directions and Challenges**

Despite the substantial progress achieved in contemporary Chinese school education, several persistent challenges remain. Such as disparities in educational resources across urban and rural areas, as well as between regions and schools, continue to hinder equitable access to quality education (Yang & Lou, 2020; Chen, 2022). Additionally, the current assessment framework requires further refinement to better align with the principles of holistic education and competency-based learning (Wang, 2022). In the future, China’s education system is anticipated to undergo further reforms, with a focus on strengthening educational equity, elevating quality standards, promoting international cooperation and technology integration, and marching on the path of sustainability, innovation, sharing, and openness (Shi, 2022; Yang & Li, 2023).

## **2.2 The Rapid Development of Contemporary Education in China**

### **2.2.1 The Improvement and Innovation of the Education System**

China’s education system has been continuously optimised in recent years, constructing a complete framework spanning from early childhood education to higher education. In early childhood education, significant progress has been achieved under the impetus of inclusive policies, with the gross enrolment rate reaching nearly 90%.

Compulsory education has largely universalised the nine-year system and is progressing towards a twelve-year system, focusing on cultivating core competencies to nurture students' autonomy and creativity. General senior secondary schools are advancing with the implementation of the new National College Entrance Examination (NCEE) reform, granting students the right to select courses, improving teachers' professional capabilities, and upgrading education quality. Higher education, supported by the 'Double First-Class' initiative, has significantly increased the international competitiveness of multiple institutions, with the gross enrolment rate surpassing 50%, entering a stage of universal access. Vocational education, guided by policies, has deepened the integration of industry and education, with institutions optimising curricula to emphasise learners' practical training, thereby enhancing learners' recognition of careers and delivering skilled talents for industries .

### **2.2.2 Continuous Improvement in Education Quality**

The goal of China's education reform has gradually shifted from scale expansion to quality improvement. The framework of students' key competencies development in China focuses on cultivating 'well-rounded individuals', highlighting cultural basis, independent development, and social engagement (Shi, 2018; Yang & Li, 2023). In terms of cultural basis, curriculum reforms have strengthened students' language proficiency, mathematical literacy, and scientific inquiry abilities while also fostering students' sense of national identity and global perspective. For independent development, holistic education has been promoted to cultivate students' abilities in autonomous learning, innovative thinking, and practical skills. In social engagement, education policies increasingly emphasise the cultivation of students' civic responsibility, teamwork spirit, and social responsibility.

### **2.2.3 Narrowing the Urban-Rural Education Gap**

A series of measures has been adopted to narrow the gaps between urban and rural areas, regions, and different social groups. Through the national implementation of the 'Rural Teacher Support Programme', more and more graduates join rural education (Jiang, 2022; Zhou & Xu, 2022). Furthermore, to ensure students from impoverished families can complete their education, the government has established a multi-level financial assistance system, including grants, scholarships, student loans, and special subsidy policies, ensuring that all students can access education.

#### **2.2.4 The Popularisation of Digital and Smart Education**

The rapid development of IT and AIT has brought new changes to school education in China. The emerging teaching models such as online education, smart education, MOOCs, flipped classrooms, and AI-assisted teaching have become increasingly popular, improving the interactivity and personalisation of education (Liu et al. 2018). The mode of ‘Internet + Education’ has been widely popularised, facilitating the sharing of educational resources and narrowing the regional education gap. For instance, the establishment of the National Smart Education Platform has enabled students in remote areas to access high-quality teaching resources from across the country. Moreover, the education evaluation system has gradually shifted from a sole focus on exam results to a more comprehensive assessment of students’ overall achievements, placing greater emphasis on multiple competencies and holistic development.

#### **2.2.5 Accelerating Internationalisation of Education**

With the acceleration of globalisation, the internationalisation of China’s school education has continuously advanced. An increasing number of universities have established cooperative relationships with international institutions, sharing educational resources and jointly cultivating internationally orientated talents. The rise of bilingual education, international curricula, overseas programs, and cross-cultural exchange projects has broadened Chinese students’ global perspectives. At the same time, Chinese education has actively attracted international students to study in China, which assists the global dissemination of Chinese culture.

### **2.3 Modern Higher Education in China**

China has established the largest higher education system in the world, with a total enrolment of over 44.3 million university students and a gross enrolment rate in higher education reaching 57.8% (Wu, 2022). At the same time, the quality of higher education has significantly improved, with certain universities and disciplines, including engineering, agriculture, the humanities, and medicine, reaching globally advanced levels (Ma, 2022). China’s higher education system is increasingly becoming a source of knowledge, talent, and innovation underpinning social development (Zheng & Yang, 2021).

#### **2.3.1 The Flourishing Development of Undergraduate Education**

*Steady increase in the number of higher education institutions*

According to the *2023 National Education Development Statistical Bulletin*, there were 1,242 undergraduate institutions across China, along with 233 research institutions authorised to offer postgraduate education, forming an extensive and diverse undergraduate education system (MOE of PRC, 2024). In addition to prestigious and long-established comprehensive universities such as Peking University and Tsinghua University, numerous specialised institutions have been established across various regions in response to local development strategies and industrial demands.

### ***Notable achievements in quality enhancement***

In 2015, the Chinese government introduced the *Overall Plan for Coordinated Development of World-Class Universities and First-Class Disciplines* (MOE of PRC, 2015), which dismantles disciplinary silos, encourages interdisciplinary collaboration, and optimises resource allocation to comprehensively enhance their institutional strength. Driven by this policy, Chinese universities have performed remarkably well in the QS World University Rankings in recent years.

### **2.3.2 The Rapid Rise of Vocational Education**

The modern concept of vocational education in China was initially formed in the ‘Westernisation Movement’ in the 1860s and matured in the 1920s. In 1985, China re-proposed the construction of a modern vocational education system. It developed rapidly from the 1980s to the mid-1990s but experienced a downturn after the mid-1990s (Song, 2022). In 2010, the government put forward the specific goals and plans for modern vocational education, which led to the rapid rise of higher vocational education. The number of vocational students has been greatly expanded and gradually reaches half of higher education students (Zhu & Shi, 2021).

In 2022, China enacted the *New Vocational Education Law*, in which a range of measures, including the positioning of vocational education, development pathways, educational and evaluation models, as well as policy and financial support, were outlined to support the sustainable development of vocational education (MOE, 2022b). This marks a new phase for higher vocational education in China, significantly enhancing both its development quality and future prospects. By 2022, China has built the world’s largest vocational education system, with a total of 11,200 vocational schools and over 29.15 million students (People’s Daily, 2022). Through close combination with industrial

development and cooperation with enterprises, diversified and smooth paths have been established for students' development.

### ***The structure of China's modern vocational education***

As shown in Table 1, China's vocational education system comprises secondary vocational education, higher vocational education, and social vocational training.

*Table 1 The Chart of China's Vocational Education Structure*

| <b>Phase</b>                                    | <b>Targets</b>                                                                                    | <b>Contents</b>                                                                                                        | <b>Source of the students</b>                                                                                                                                                                                                                     |
|-------------------------------------------------|---------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Basic stage: secondary vocational education     | To train entry-level skilled workers                                                              | Secondary professional schools<br>Technician training schools                                                          | Students who fail to enter high school after graduating from junior high school                                                                                                                                                                   |
| Advanced stage: higher vocational education     | To cultivate middle and advanced technical application talents                                    | Three-year higher vocational colleges<br>Four-year vocational colleges (application-oriented undergraduate)            | Students who are willing to become workers and craftsmen after the college entrance examination<br>Students who are willing to become workers and craftsmen after the college entrance examination;<br>Students advancing from vocational college |
| Supplementary stage: social vocational training | To enhance the vocational ability of enterprise employees, laid-off workers and rural labor force | Agricultural technology and domestic service training<br>Re-employment skills training<br>Job skills upgrading program | Rural social personnel<br>Laid-off workers<br>Enterprise employees on duty                                                                                                                                                                        |

The fundamental phase of China's vocational education aims to cultivate entry-level skilled workers. The students in this phase typically come from two main groups: those who have not entered higher secondary education after graduating from junior high school and those seeking to acquire specific skills to enter the workforce immediately. These institutions focus on equipping students with essential practical skills and theoretical knowledge to prepare them for their initial careers in technical and industrial sectors.

The advanced vocational education institutes include three-year higher vocational colleges and four-year vocational universities, aiming at cultivating middle- and advanced-level technical talents to meet the growing demand for a more skilled workforce, particularly as industries evolve and require more specialised knowledge. These institutions cater to students who are eager to advance their technical expertise and become skilled workers or craftsmen after completing the national college entrance examination.

The emphasis in this phase is on deepening practical skills while ensuring students acquire a solid theoretical foundation in their future careers.

The social vocational training, as a supplement to vocational education, focuses on enhancing the vocational abilities of various social groups, including enterprise employees, laid-off workers, and rural labour forces. The social vocational training provides a flexible and practical solution to improving the skill sets of adults who may have already entered the workforce but need further development to keep pace with changing industry demands.

### ***Philosophy of modern vocational education in China***

As early as the 1920s, Huang Yanpei, the pioneer of Chinese vocational education, proposed the establishment of modern vocational education to change the situation of the social poverty and national underdevelopment. Pragmatism was the philosophy of his educational concepts in which school education was suggested to be suitable for the needs of social life and synchronous with the development of industries (You, 2022; Zhang & Cui, 2022), which created the model of the combination of learning and practice, work and study in vocational education (Tan & Chai, 2022; Hao & Jiang, 2020) and provided significant reference and inspiration for the future development of vocational education.

In the new era now, the philosophy of China's vocational education has been further explored and developed into a system of being student-orientated (Hu, 2021), competence-based (Wang, 2022), industry-needs-orientated (Wang, 2022; Ge et al., 2022), and multiple-evaluated (Qin & Mo, 2022), forming a unique discourse system of Chinese vocational education, such as 'high-quality and adaptability of vocational education' 'apprenticeship with Chinese characteristics' 'integration of industry and vocational education' 'college-enterprise cooperation' 'combination of learning and practice' 'institute of industrial' 'college entrance examination of vocational education' 'combination of skill and morality cultivation' and 'craftsmanship spirit' (Li & Teng, 2022).

### ***Characteristics of students at Chinese modern vocational education***

The diversion of students in China occurs in two stages. The first one happens at the end of junior middle school education. Students will enter senior high schools or secondary professional schools and technician training schools, respectively, in a proportion (MOE, 2021) based on students' achievements, scores mostly. Students whose scores rank in roughly the top 60% will go to senior high schools, while others will go to

secondary professional schools and technician training schools. The second one occurs at the end of senior high school education. Generally, students whose scores in the national college entrance examination rank in the top 50 percent will go to research undergraduate universities, while the rest mostly go to the four-year vocational universities or three-year higher vocational colleges. Definitely, there's also a small probability of something happening, some excellent students will be admitted to the top vocational universities and colleges rather than the research universities. Scholars have studied the rational logic of the diversion of students, some of them affirm the positive role in optimising the structure of students, improving the modern educational system, and promoting the coordinated development of general education and vocational education (Hu et al., 2022) while others also concern about the hidden block of education equity and the inhibition of student development (Hua et al., 2022).

Based on the current practice of diversion of students, the researchers have found that, compared to research university students, vocational students are stronger in social communication, hands-on operation, and career development but generally weaker in knowledge base, learning motivation, and self-control and self-efficacy in learning (Wu, 2022; Chen, 2022; Li & Liu, 2022; Zhao & Gao, 2022; Zhan & Zhang, 2021; Yu, 2019). There are also scholars having found that higher vocational students have a higher degree of learning immersion in the operational and application-orientated curriculum studying, and they are more likely to be employed, but their concentration on tasks and self-restraint in learning is poor (Liu & Yang, 2017).

## **2.4 Educational Informatisation in China**

Education informatisation in China originated from electronic education. In 1999, the Chinese government proposed to improve the level of using technology to enhance education and the degree of educational informatisation for the first time (Central Committee of the CPC, State Council, 1999). Over the past two decades, with the support and energisation of informational and digital technologies such as the internet, big data, and artificial intelligence, the development of educational informatisation in China has made extraordinary achievements, driving educational modernisation to march forward (Chen et al., 2022; Chen1 & Chen2, 2020).

Great achievements of educational informatisation are mainly reflected in the following aspects: the overall popularisation of digital campuses, the convenient access to the internet and full coverage in universities, primary and secondary schools across the country, nationwide multimedia classrooms in 98.35% of primary and secondary schools, the establishment of smart campuses, and the exploration and application of massive high-quality digital resources in all levels of education, which are greatly boosting the development of China's education (Zhi-Yan, 2022; CAC, 2022).

In the field of vocational education, a smart educational ecosystem has been gradually built, in which teaching and learning with all-domain knowledge, all-true situations, and whole-process embodied experience have been realised (Song et al., 2021; Dong et al., 2021; Zhang et al., 2020). Technologies, including 'internet plus' 'AI' 'virtual simulation' 'big data' and the 'Internet of Things (IoT)', are widely used in the fields of vocational education. Aided by these technologies, smart teaching and simulation practice are popularised in the whole process of vocational education, and the resource integration of schools, governments, enterprises, and industries has been realised. Besides, educational resource platforms integrating perceiving, teaching, practising, training, and research can be jointly established and shared. Therefore, technology based on IT and AIT has played an important role in the whole process of teaching, learning, and school management (Wu et al., 2022; Ran, 2022; Zhao et al., 2022).

## **2.5 The Educational Philosophy of Chinese School**

Educational philosophy is a kind of philosophical reflection on the nature, purpose, and problems of education (Siegel, 2022). The purpose of studying educational philosophy is to answer questions such as why, who, by whom, where, what, how, and when to educate, in which the purpose of education, curriculum, teaching methods, teachers, students, teaching materials, schools, discipline, evaluation, and supervision and management of education are involved (Ravi, 2022, 114).

### **2.5.1 Aims of Education**

The aims of education are of critical and enduring human concern (Nilholm et al., 2021). General aims of education contain individual and social aims, including the aspects of knowledge, culture, character-building, spirit, career, citizenship, physical development, and self-realisation (Ravi, 2022, 51-72), being determined by not only the factors of

geography, politics, economy, religion, culture, environment, and social factors, but also the philosophy of life, knowledge, and psychology of individuals (Ravi, 2022, 41-42).

School education in ancient China emphasised the aims of individual development in terms of knowledge and skills, ethics and morality, self-adjustment and self-realisation, but more importantly, the social aims, that is, the participation in social management and the continuation of the ruling regime. The main purpose of official school education is to cultivate elites who are conducive to maintaining the governance of the ruling regime via the official selection system, such as the imperial examination system. Influenced by this orientation, for most individuals, the main motivation and purpose of being educated was to become an official (Wang, 2013). In contrast, the aims of school education in the new era of China take into account both individual and social aims, which were mainly reflected in the concept and practice of cultivating builders and successors of socialism with individual ‘all-round’ development of morality, intelligence, physics, aesthetics, and labour spirit, effectively avoiding educational utilitarianism and promoting the aims of school education to return the track of true value (Tan & Tan, 2022; Jiang, 2020).

### **2.5.2 Curriculum**

Curriculum supports to achieve the goals of school education. Educational philosophy is of enormous significance for curriculum based on social needs (Ravi, 2022, 116). Factors affecting the evolution of curriculum include knowledge, social cultures, leadership mechanisms, learners, and teachers (Qin, 2021). The construction of the ancient curriculum system in China services and supports the aims of school education to build the system of human ethics and social governance. Therefore, the curriculum composed of Confucian and its works lasted for more than 2,000 years. Although critical roles in serving politics and individual development did it play, it eventually raised the query and became the focus of modern education reform due to the unitary structure, boring contents, and limitation of thoughts. It was then replaced by the curriculum of a combination of Chinese traditional classics and Western scientific and technological knowledge (Zhou, 2012). When entering modern times, curriculum reform and development in Chinese school education has experienced a long and complex process from experience-based, knowledge-based, labour-based, ability-based, to comprehensive quality-based (Pan, 2021). Up to now, Chinese school education has established a modern curriculum system that

integrates moral education, intellectual education, physical education, aesthetic education, and labour education based on the learners' development of comprehensive quality (Wang<sup>1</sup> & Wang<sup>2</sup>, 2022; Ning & Fan, 2022).

### **2.5.3 Teaching Methods**

Teaching method is an important factor to ensure the effect of education. It is proved that appropriate teaching methods make teaching more attractive and effective. For the research on teaching methods, some Chinese scholars have a strong interest in reviewing the ancient Chinese teaching methods and believe that they have immense value of inspiration and reference for today's teaching practice (Peng, 2022; Hu, 2021; Shan, 2021). More scholars focus their research on the reform and innovation of the teaching methods of specific courses in modern school education (Jin & Sha, 2022; Xu et al., 2022; Cheng, 2022; Li, 2022; Bi, 2022).

Educators in ancient China did effective exploration of teaching methods. For instance, Confucius proposed to teach without discrimination, teach students in accordance with their aptitude, teach for inspiring and inducing, and teach and learn simultaneously (Li et al., 2022). In spite of this, for a long time in the past, under the influence of the examination system, the talent selection system, and the philosophy of official-centred, exam-orientated teaching methods, the cramming method occupied the dominant position in Chinese school education. Chinese students are still considered to be strong in testing but poor in practice and innovation (Zhang et al., 2021).

Inspired by the naturalistic 'learning by doing' and the pragmatic 'experimental method' proposed by Ravi (2022), in recent years, modern educators have reflected and realised that the harms of education for exams and cramming teaching methods are in suppressing the initiative and creativity of students, preventing students from thinking independently, and ignoring the core of learners in education, in the context of the rise of quality education (Wang et al., 2021). Scholars regard exam-orientated and cramming teaching methods as 'a killer' that stifles students' creativity and imagination and trains students to become machines of exams, for they only tell students to know what, but not tell them why (Zheng & Xu, 2021; Chen & Bao, 2021; Ling, 2018). Therefore, heuristic and inquiry-based teaching methods have been popularised in the teaching of various disciplines (Guan, 2022; Wu, 2022; Miao & Chang, 2022), and the project-based teaching

method, behaviour-orientated teaching method, case-based teaching method, task-driven teaching method, cooperative teaching method, inquiry teaching method, and situation-based teaching method are widely adapted in school teaching.

#### **2.5.4 Teacher and Student**

According to Ravi (2022), the importance of teachers and students is determined by the philosophy of education. In idealism the teacher is important and the student is secondary, while on the contrary in naturalism.

The evolution of teacher-student relationships in China has experienced a long journey from unequal, teacher-centred in ancient times to student-centred, equal, and cooperative in modern times (Shao, 2016). In ancient China, teachers had absolute authority deciding how and what the students learnt, which was closely related to the level of human civilisation at that time. With the evolution of human civilisation and the progress of the social system, the old divine authority of teachers has been dissolved. Just as teachers dressed as God in the late Middle Ages and beyond in western countries encountered great challenges in the process of human secularisation, the divine authority of ancient Chinese teachers was also abandoned. When the social system continued to improve, especially knowledge, especially scientific knowledge, started to be respected in the whole society, the basis for the establishment of teachers' authority changed (Liu, 2021). Teachers' dignity and authority should be obtained in the field of education equally and naturally (Zhao & Cai, 2020).

In contemporary China, the teaching profession is still esteemed as a noble and significant vocation by the public, with educators regarded as friends, philosophers, mentors, and social architects, often compared to 'spring cicadas' 'gardeners' and 'candles' (Luo, 2022; Yao, 2021). China has a longstanding tradition of venerating educators since antiquity (Liu & Li Yue, 2022; Guo & Lin, 2020; Liu, 2019). Contemporary teachers are tasked not only with instruction but also with guiding students to form accurate perspectives on global and existential values (Ming & Zhao, 2022; Yu, 2020) and facilitating their physical and mental development of knowledge and competencies (Han & Song, 2021; Ban, 2020). The role of teachers has evolved from that of primary educators to instructors and collaborators (Dai & Chen, 2021).

#### **2.5.5 Teaching Content and Teaching materials**

Teaching materials in the category of educational philosophy are the teaching contents formulated according to the students' ages, interests, attitudes, abilities, and the growing needs of society. The changes of teaching contents can reflect the change of educational philosophy to some extent. In ancient China,

Over thousands of years in ancient China, the teaching content in school education was mainly Confucian and its works, such as the Four Books and the Five Classics. They are criticised fiercely in modern times for their fixed contents and untimely ideology, confining students' eagerness for various knowledge and scope of imagination and hindering the variety of cultures and the introduction of scientific knowledge (Shi, 2021). However, some scholars praise the ancient classics for their incorporation of philosophical ideas, humanistic spirit, and moral concepts. They are regarded as the roots of Chinese culture.

When entering the new era, teaching materials are expected to play the role of enlightening wisdom and ideology in school education (Yan, 2021). The building of teaching contents, represented by high-quality textbooks, has made great achievements (Shi & Zhang, 2021). Massive digital teaching resources are constantly being promoted in teaching now (Tan, 2022). Scholars propose that more polynary, adaptive teaching resources, full of local and Chinese characteristics, should be made in the future (Zhang & Liu, 2022).

#### **2.5.6 Assessment for Teaching and Learning**

Education evaluation is a systematic practice of assessing all aspects of education in terms of educational system, educational philosophy, educational aims, curriculum, teaching, learning, and teaching materials. Assessment of teaching and learning provides proof of the effects of teaching activities by observing both teachers. For a long time in the past, summative assessment was used to assess teaching and learning. Kinds of entrance examinations became the main forms of evaluating students' periodical learning achievement, and scores of the curriculum were regarded as the only proof in the verdict of the learning effect at the end of a semester or an academic year (Niu, 2018; Ye, 2019). It comes to be doubted and distrusted by more and more scholars as learners' formative engagement in learning and comprehensive quality are ignored in this one-off and unilateral assessment (Liu & Wei, 2022; Liu & Xu, 2020). Scholars and educators in China

propose that integrated, multi-dimensional, dynamic, and student quality-centred assessment should be implemented with the aid of technology (Shi, 2022; Liu & Xu, 2022; Xia, 2022). The teaching and learning evaluation methods have been changed from managerialism to constructivism, in which more emphasis has been placed on the consistent construction of teaching, learning, and evaluation (Xia, 2022).

Driven by the rapid development of informational and artificial intelligence technologies, a combination of web-based, technology-based, intelligent formative and summative assessments have become a new trend and a research spot in teaching and learning (Luo et al., 2021; Peng et al., 2021).

### **2.5.7 School Administration**

Ravi (2022) points out that, according to the educational philosophy, the management of schools is democratic. Dewey believes that democracy itself was a kind of educational principle, guideline, and policy in education (Dewey, 2014), and he proposes two standards of democracy: the members of the community participated in and shared consciously and the degree of freedom and adequate communication with other communities (Dewey, 1980). However, Bode criticised Dewey's view as he believes that it focused mainly on the number of problems rather than the quality, lacking the foundation of values (Bode, 1938), which would expand the fragility of democracy (Bode, 1938). Therefore, genuine and high-quality democracy in education should be in the overall process, all stakeholders and all aspects (Zhang, 2022).

The process of educational democracy is mainly reflected in the three educational reforms in China, the first of which was the reform of the new school system and curriculum in 1922, in which the modernisation and democratisation of education were officially confirmed and systematically promoted. It took educational enlightenment and democracy as fundamental purposes and liberalism as a foundation. Since then, social and educational democracy has started to form. The second is the 1988 Shanghai elementary educational reform of curriculum, which emphasises the exploration of students' personalities and encouraged students to choose curriculum freely. It laid a firm foundation for the curriculum reform in the 21st century. The third is the new curriculum reform of elementary schools in 2001, which put forward the idea of 'education for the development

of every student' and advocated the autonomic, cooperative, inquiry-based mode of teaching and learning (Zhang, 2019).

On the whole, educational democracy in Chinese school administration has marched forward a lot. The system of educational democracy has been established, and the practice of educational democracy has continued to go deeper. In primary and secondary schools, parents and students participate in school management in the form of family committees, forming a home-school collaborative parenting mechanism (Dong, 2021; Liu, 2021; Wan, 2020). In universities, stakeholders, including teachers, students, society, and government, are invited into decision-making and routine management, ensuring the autonomy and publicity of university administration (Bao et al., 2022). However, the overall democracy in school education is still insufficient. For instance, the imbalance between schools, the strong hand of government in school management, administrative and official-based ideology at schools, and the limitation of academic power still exist to some extent (Tang, 2021; Wang & Zhu, 2019; Meng & Li, 2018).

## **2.6 Current Research on Chinese School Education**

Regarding the research on the topic of Chinese school education, scholars at home and abroad, especially Chinese scholars, have carried out a lot of research and made great achievements. The domestic works on it mainly include three categories. The first category consists of textbooks (Cheng & Niu, 2021; Sun, 2019; Chen, 2018) contributing to expound the origin, formation, development, downfall, and renaissance of Chinese school education from a diachronic perspective and make detailed descriptions of the educational systems, educational policies, educational philosophy, and educational contents as well in various historical periods, which helps us to understand the past and present of Chinese school education in short time. The second category is academic monographs, some of which focus on the educational system in various historical periods or specific times in China (Guo, 2017), and the other part focuses on systematic reviews of China's educational thought (Sun, 2022; Kong & Huang, 2020; Feng, 2020; Xiao et al., 2020). The third category is academic journal papers and degrees thesis. According to the 2021 annual report on the Frontiers and Hot Issues of Education Research in China (2022), research on educational review and reflection, education development under the new development pattern, optimisation of education supply, the establishment of high-quality education

systems, family-school collaboration education, family-school collaboration education, rural education, higher education popularisation and high-quality development, improvement of educational laws and modernisation of education governance, education opening and internationalisation, education informatisation, education (information) technology, and vocational education are frontier and hot issues of China's education. By searching the literature about the above issues published in the source journals of SCI, EI, CSSCI, CSCD, and Peking University core journals from the years of 2019-2022 through China National Knowledge Infrastructure (CNKI), shown in the following Table 2 we can find that vocational education, educational informatisation and educational technology, and moral education are high-frequency words in the map of education research.

*Table 2 Distribution of the Publications on the Frontiers and Hot Issues of China's*

*Education Research (2019-2022)*

| Subject                                                                   | Yearly publications in the source journals of SCI, EI, CSSCI, CSCD, and Peking University core journals (2019-2022) |      |      |      |       |
|---------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------|------|------|------|-------|
|                                                                           | 2019                                                                                                                | 2020 | 2021 | 2022 | Total |
| Vocational education                                                      | 640                                                                                                                 | 769  | 841  | 683  | 2933  |
| Education informatisation, education (information) technology             | 155                                                                                                                 | 264  | 121  | 81   | 621   |
| Virtue education                                                          | 174                                                                                                                 | 177  | 168  | 97   | 616   |
| Establishment of high-quality education systems                           | 67                                                                                                                  | 114  | 139  | 91   | 411   |
| Rural education                                                           | 54                                                                                                                  | 46   | 52   | 53   | 205   |
| Education opening and internationalisation                                | 62                                                                                                                  | 48   | 48   | 33   | 191   |
| Family-school collaboration education                                     | 41                                                                                                                  | 47   | 53   | 37   | 178   |
| Improvement of educational laws and modernisation of education governance | 22                                                                                                                  | 33   | 24   | 16   | 95    |
| Higher education popularisation and high-quality development              | 15                                                                                                                  | 19   | 20   | 8    | 62    |
| Optimisation of education supply                                          | 0                                                                                                                   | 16   | 16   | 9    | 41    |
| Education development under the new development pattern                   | 0                                                                                                                   | 3    | 22   | 13   | 38    |
| Educational review and reflection                                         | 5                                                                                                                   | 15   | 12   | 5    | 37    |

## 2.7 Chapter Summary

China's school education has undergone a long process of evolution and has achieved remarkable progress in the contemporary era. The education system has been continuously optimised, and higher education has expanded significantly while improving in quality. Vocational education has developed into the largest system in the world with strong industry linkages, and educational informatisation has been widely implemented.

Educational philosophies have also undergone continuous reform in various aspects. The objectives of education have shifted from a traditional emphasis on societal and

personal development to a more comprehensive and balanced approach in modern times. Curricula have evolved from classical texts to a broader competency-based framework. Teaching methods have moved away from examination-orientated instruction towards more diverse and inquiry-based approaches. Teacher-student relationships have become increasingly equal and cooperative, while teaching materials have been continuously updated to align with contemporary needs. Furthermore, assessment methods have become more scientific, and although school governance has progressed towards greater democratic participation, there remains room for further improvement.

For this study, which explores the use of WBFAT in developing oral English proficiency among vocational college students, the rapid advancement of China's vocational education system provides a broad and supportive application context. Vocational education places strong emphasis on practical skills training, and oral English proficiency is a crucial professional competency. The achievements of recent educational reforms, such as the widespread adoption of digital technology and innovations in educational philosophy, create favourable conditions for integrating WBFAT. By leveraging modern educational resources and pedagogical advancements, these tools can effectively enhance oral English proficiency training in vocational colleges, further improving the quality of talent cultivation in vocational education. This aligns with the ongoing trajectory of educational development and innovation in China.

## **Chapter 3 Literature Review**

The literature review is to provide this study a structured theoretical and empirical background. This will help the researcher understand key ideas (Ridley, 2012) and make the research content clearer and more consistent (Creswell, 2014). It also ensures that the research content and methodology build upon the existing body of knowledge while maintaining academic originality (Hart, 1998; Machi & McEvoy, 2016).

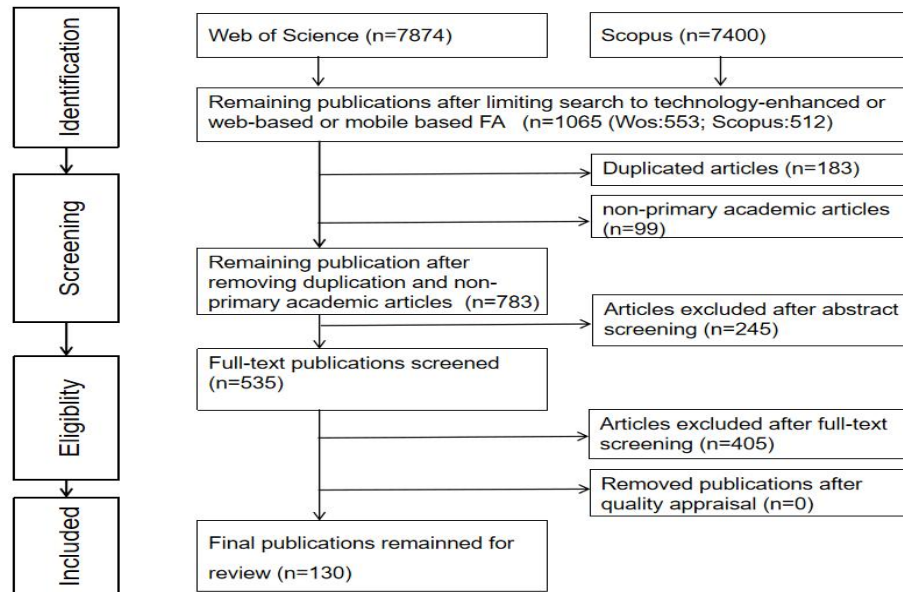
This chapter will first elaborate on the procedures for selecting literature, including the databases used and the search strategies adopted. The research will clarify the criteria for including relevant publications and excluding others to ensure that the final collection of literature for this study undergoes a thorough screening process. Subsequently, the selected literature will undergo eligibility review and quality assessment to ensure its validity and reliability. Following this, an in-depth interpretative analysis of the included literature will be conducted, exploring the development trends of research on WBFAT-enhanced learning, its geographical distribution, participant characteristics, research contexts, disciplinary coverage, research methodologies, theoretical foundations, tools employed, and their effects, including positive aspects, negative feedback, challenges, and user expectations. Finally, the research will provide a comprehensive analysis of the significant theories, including learning theories (like metacognitive theory, self-regulated learning theory, motivation theory, and constructivist learning theory), assessment theories (like summative assessment, formative assessment, diagnostic assessment, and dynamic assessment), and theories about how people accept and use technology. This will provide a solid theoretical base for further research.

### **3.1 Procedure of Literature Selection**

To systematically collect empirical evidence on WBFAT, the researcher followed the general procedures outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines (Moher et al., 2009). This process aims to ensure the transparency, comprehensiveness, and methodological rigour of the review, providing reliable evidence to support the effectiveness and application of WBFAT. Figure 1 illustrates the detailed procedures of this review, clearly outlining the stages from literature search to data extraction. In the remainder of this section, I will describe in detail the specific steps of the literature search process, including the databases used, search

strategies, and keywords; the criteria for inclusion and exclusion, specifying which studies were included and which were excluded; as well as the details of the literature coding procedure, explaining how the selected studies were extracted, categorised, and coded to ensure the accuracy and consistency of the results. The implementation of these procedures will provide comprehensive empirical support for the application of WBFAT in oral English assessments and offer a reference framework for future related research.

*Figure 1 The Flow Chat of the Procedure in Literature Selection and Screening*



### 3.1.1 Data-Bases and Search Strategy

The literature search was conducted in two comprehensive electronic databases widely used by researchers in the humanities and social sciences: Web of Science and Scopus. The reason I chose them is that they are recognised as important, authoritative, international, multidisciplinary academic databases. Web of Science, provided by Clarivate, is a highly authoritative academic database renowned for its rigorous journal selection standards and high-quality scholarly resources. Due to its academic precision and extensive application in citation tracking, it is regarded as an essential tool for academic research (Harzing & Alakangas, 2016). Its core sub-databases, such as the Science Citation Index (SCI) and the Social Sciences Citation Index (SSCI), encompass leading journals across various disciplines, thereby offering a reliable foundation for scholarly inquiry. Web of Science not only allows researchers access to high-quality academic literature but also enables the use of citation analysis tools to identify key publications and influential scholars within a given field. Furthermore, its extensive historical records, which date back

to 1900, are particularly valuable for examining the evolution and trends within academic disciplines. Scopus, developed by Elsevier, is among the largest abstract and citation databases globally, covering over 24,600 journals across a wide range of disciplines, including science, technology, medicine, social sciences, and the humanities. Its breadth of coverage and rapid data updates make it a preferred choice for interdisciplinary research (Mongeon & Paul-Hus, 2016). In comparison to other databases, Scopus offers a broader range of academic resources, particularly excelling in the fields of social sciences and arts and humanities. Additionally, Scopus provides comprehensive bibliometric tools and citation analysis features, which empower researchers to identify emerging research trends, map academic networks, and keep abreast of the latest developments in their fields.

### **3.1.2 Criteria for Inclusion and Identification of Relevant Publications**

As illustrated in Figure 1, an initial search was conducted in the full-text database of Web of Science under All Collections using the topic ‘formative assessment for learning’, with the publication date range set from 1 January 2003 to 30 December 2023. A total of 7,874 documents related to this topic were retrieved. Among these, 4,181 were articles, 1,575 were dissertations and theses, 285 were review articles, and 874 were conference papers. Subsequently, the same search parameters were applied in the full-text database of Scopus under All Collections, resulting in 7,400 documents. This dataset comprised 4,522 articles, 1,707 conference papers, 661 book chapters, and 273 review articles.

To further narrow the scope and precisely focus on the theme of this study, the researcher refined the search by using the keywords ‘web-based formative assessment’, ‘technology-enhanced formative assessment’, and ‘mobile-based formative assessment’. The data range was set from 1 January 2003 to 30 December 2023. This refined search yielded 553 relevant articles in the full-text database of Web of Science and 512 relevant articles in the full-text database of Scopus.

## **3.2 Screening, Eligibility, and Quality Appraisal of Literature**

### **3.2.1 Screening and Eligibility**

As highlighted by Shaheen et al. (2023), the lack of a rigorous eligibility assessment during a systematic review can lead to the inclusion of low-quality studies, thereby compromising the reliability of the review’s conclusions. To ensure that the final

selection of literature was highly relevant and of reliable academic quality, the researcher employed EndNote 21 to merge the 1065 records identified in the previous step. After this, 183 duplicate records were removed, and 99 non-peer-reviewed documents, such as notes, editorials, patents, letters, case reports, and abstracts, were excluded, leaving 783 records for further screening. In the abstract screening phase, the researcher reviewed the abstracts and excluded those that were not aligned with the research topic, resulting in the removal of 245 articles. This left 535 articles for full-text screening. Following the full-text review, 405 articles were discarded for being non-peer-reviewed or irrelevant to the study, including 218 research articles, 72 conference papers, 92 reviews, and 23 book chapters. Ultimately, 130 articles were selected for quality appraisal, comprising 110 peer-reviewed research papers, 12 conference papers, 3 book chapters, and 5 PhD theses.

### **3.2.2 Quality Appraisal**

As shown in Table 3, during the quality appraisal phase, the researchers evaluated the quality of the remaining 130 articles based on the criteria adopted by Theelen et al. (2019), which were developed from the evaluation methods for qualitative research (Savin-Baden & Major, 2007) and quantitative research (NICE, 2012). For qualitative studies, nine criteria were established, focusing on aspects such as methodology, theoretical foundation, ethical processes, the relationship between researchers and participants, consistency between methodologies and data collection methods, interpretation of findings, critical perspective, and acknowledgement of limitations. For quantitative studies, seven criteria were designed, addressing the representativeness of participants, intergroup comparability, accuracy of outcome measurements, appropriateness of statistical analysis, and internal and external validity.

The included studies consisted of qualitative research ( $n = 20$ ), quantitative research ( $n = 27$ ), and mixed-methods research ( $n = 83$ ). The quality of mixed-methods studies was assessed according to both qualitative and quantitative standards. Each article was rated on a scale from 'no mention' (0) to 'extensive mention' (3), and an average score was calculated based on its performance in each category. Following the application of these criteria, all articles received an average score above 2 (Good Mention). As a result, no articles were excluded during the quality appraisal phase. In total, 130 articles were included in the final analysis.

Table 3 Criteria for the Quality Appraisal of the Literature

| Criteria                                                                                          | No<br>Mention<br>(0) | Some<br>Mention<br>(1) | Good<br>Mention<br>(2) | Extensive<br>Mention<br>(3) |
|---------------------------------------------------------------------------------------------------|----------------------|------------------------|------------------------|-----------------------------|
| Criteria for qualitative studies                                                                  |                      |                        |                        |                             |
| Clear methodology                                                                                 | 0                    | 2                      | 80                     | 21                          |
| Theoretically situated                                                                            | 32                   | 4                      | 52                     | 15                          |
| Transparent ethical process                                                                       | 8                    | 2                      | 72                     | 21                          |
| Researcher(s)' relation to participants is clear                                                  | 10                   | 4                      | 76                     | 13                          |
| Researcher(s)' relation to the data is clear                                                      | 8                    | 6                      | 84                     | 5                           |
| Researcher(s)' takes a critical stance                                                            | 4                    | 6                      | 69                     | 24                          |
| Congruence between methodology and methods used for data collection, analysis, and interpretation | 0                    | 12                     | 50                     | 41                          |
| Participants involvement in data interpretation                                                   | 8                    | 11                     | 55                     | 29                          |
| Limitations voiced                                                                                | 3                    | 8                      | 75                     | 17                          |
| Criteria for quantitative studies                                                                 |                      |                        |                        |                             |
| The source population or area is well-described                                                   | 2                    | 4                      | 83                     | 21                          |
| Interventions and comparisons are well-described and appropriate                                  | 0                    | 5                      | 89                     | 16                          |
| Outcome measures are reliable                                                                     | 0                    | 6                      | 91                     | 13                          |
| Outcomes are relevant                                                                             | 0                    | 9                      | 89                     | 12                          |
| Appropriateness of statistical analysis                                                           | 0                    | 2                      | 94                     | 14                          |
| Internal validity of results (i.e., unbiased)                                                     | 2                    | 9                      | 88                     | 11                          |
| External validity of findings (i.e., generalisability)                                            | 1                    | 15                     | 79                     | 15                          |

### 3.3 Explanatory Analysis of the Included Publications

Based on an intensive reading of the abstracts and full texts of 130 publications, the publications were classified by author, year of publication, source, participants and their number, geographical region, study duration, discipline, research background, WBFAT function, theoretical basis, methodology, data collection methods, and findings. In the following sections, I will make an exploratory analysis and interpretation of these aspects.

#### 3.3.1 The Research on WBFAT for Learning Has Rapidly Increased

As presented in Table 4, over the past two decades, the literature on web-based or mobile-based formative assessment has shown an increasing trend. However, there are notable differences in the pace of development and stability between the various stages, with the overall standard deviation of the number of publications across the four periods being relatively large ( $SD = 7.27$ ). This value reflects the degree of dispersion in the number of publications at different stages in the field.

Table 4 Year Distribution of the Included Publications ( $N=130$ )

| Period    | Frequency | PCT(%) | SD1  | SD2 |
|-----------|-----------|--------|------|-----|
| 2003-2010 | 24        | 18.46  | 2.79 |     |
| 2011-2015 | 32        | 24.62  | 1.73 |     |
| 2016-2020 | 38        | 29.23  | 1.58 |     |
| 2021-2023 | 36        | 27.69  | 3.00 |     |

Between 2003 and 2010, the number of related publications was relatively low, and the growth rate was slow ( $n = 24$ , 18.46%). In particular, during the early 2000s, there was a limited number of relevant publications. Despite significant advancements in technology during this period, its integration into education, especially in the field of assessment, was still in its infancy. Systematic studies on the integration of web and mobile technologies with formative assessment had not yet gained widespread attention.

From 2010 to 2015, with the widespread adoption of smartphones and tablets and the rise of big data and cloud computing, IT and AIT devices rapidly integrated into the education sector, becoming crucial tools for enhancing teaching outcomes. The emergence of social media and online learning platforms also provided new learning opportunities for students, accelerating the increase in research on web-based or mobile-based formative assessment ( $n = 32$ , 24.62%). Dubé et al. (2018) highlight that the rapid development of educational technology, particularly the proliferation of mobile devices, significantly contributed to the transformation of formative assessment practices.

Between 2016 and 2020, the volume of relevant literature increased dramatically ( $n = 38$ , 29.23%), especially after 2017, when web- and mobile-based formative assessment became a central focus in the field of educational technology. During this period, global educational discourse shifted towards leveraging technology to support personalised and self-regulated learning, with formative assessment emerging as a key element of this process. Additionally, the rise of MOOCs has provided significant momentum for this trend, and the ongoing development of educational technologies facilitated deeper exploration and broader application of technology-enhanced formative assessment methods.

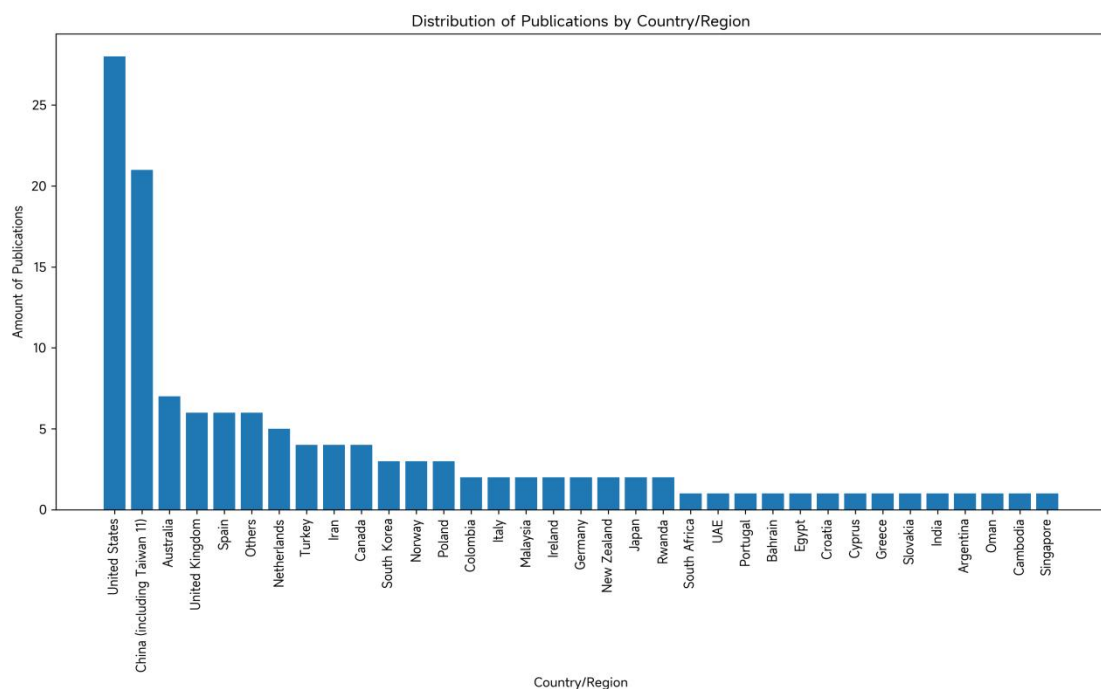
Since 2020, this research trend has continued to accelerate ( $n = 36$ , 27.69%), particularly with the onset of the COVID-19 pandemic, which caused an explosive growth in demand for online learning and remote education. The rapid shift to online teaching created an urgent need to develop effective assessment strategies in virtual environments. Criollo-C et al. (2021) argue that the rapid development of the internet and mobile technologies has spurred innovation in educational assessment, particularly in large-scale online and blended learning settings, where traditional assessment methods proved inadequate, thus driving the advancement of formative assessment research. Furthermore,

studies (Barrot et al., 2021; Gonzalez et al., 2020; Drane et al., 2020) indicate that the COVID-19 pandemic hastens the integration of technology in educational assessment, particularly in meeting the growing demand for self-regulated learning, instant feedback, and personalised learning, making technology-enhanced formative assessment a focal point in educational research.

### 3.3.2 The Research on WBFAT for Learning Are Mainly Concentrated in the United States and China

As can be seen in Figure 2, research in this field is predominantly led by the United States and China, while contributions from other countries are relatively dispersed and limited in number. The highest volume of publications is accounted for by the United States ( $n = 28$ , 21.53%), with its research output surpassing that of other nations by a significant margin. This is likely due to the prominent role played by American research institutions and universities in advancing global educational studies, particularly in the areas of IT and AIT. China (including Taiwan) ranks second in terms of publication volume ( $n = 21$ , 16.15%), a trend that can be attributed to the strong promotion of ‘smart education’ and ongoing educational reforms involving IT and AIT by the Chinese government (Feijóo, 2021; Ma et al., 2015). In contrast, the remaining 32 countries or regions exhibit a noticeable disparity, with publication numbers ranging from 1 to 7.

Figure 2 Distribution of the Included Publications by Country or Region



### 3.3.3 The Distribution of the Included Literature in Terms of Participants, Study Duration, Study Context, Disciplinary Distribution, and Research Methods

From Table 5, which reflects the distribution of the 130 publications, it is evident that the range, central tendency, and dispersion of Participants, Duration of research, Research Contexts, Discipline distribution, and Methodology vary significantly. The subcategories of these items show considerable differences, with a high level of dispersion, particularly in the case of discipline distribution, which exhibits substantial variation ( $SD = 65.05$ ).

*Table 5 Statistical Analysis of Participants, Study Duration, Contexts, Disciplinary Distribution, and Research Methods Within the Included Publications*

|                         | <b>Max.</b> | <b>Min.</b> | <b>SD</b> |
|-------------------------|-------------|-------------|-----------|
| Participants            | 69          | 1           | 27.02     |
| Duration of research    | 45          | 0           | 16.84     |
| Research Contexts       | 90          | 6           | 39.37     |
| Discipline distribution | 116         | 24          | 65.05     |
| Methodology             | 83          | 20          | 34.53     |

From Tables 6, we can gain a more detailed understanding of the distribution of participants, duration of research, research contexts, discipline distribution, and methodologies in these studies.

*Table 6 The Distribution of the Included Literature in Terms of Participants, Study Duration, Context, Disciplinary Distribution, and Research Methods*

| <b>Category</b>      | <b>Sub-category</b>                   | <b>Frequency (%)</b> | <b>Reference</b>                                                                                                                                                                                              |
|----------------------|---------------------------------------|----------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Participants         | Postgraduates & Undergraduates        | 69 (52.31%)          | Antee (2021); Bacca-Acosta & Avila-Garzon (2020); Baleni (2015); Buldu & Buldu (2010); Bjørkli (2014); Broadbent et al. (2020); Çevik (2015); Chang & Lin (2020); Costa et al. (2010); Comiskey (2012) et al. |
|                      | Vocational college students           | 2 (2.31%)            | Alavi et al. (2021); Cigdem, H. & Oncu, S. (2015)                                                                                                                                                             |
|                      | Primary and secondary school students | 39 (30.00%)          | Barana & Marchisio (2016); Chen & Chen (2009); Faber et al. (2017); Fathi et al. (2023); Fjærtøft (2020); Georgiou et al. (2021); Kim et al. (2016) et al.                                                    |
|                      | Nursery school students               | 1 (0.77%)            | Hooker (2017)                                                                                                                                                                                                 |
|                      | Teachers only                         | 9 (6.92%)            | Carey & Stefaniak (2018); Dahalan & Hussain (2010); Feldman & Capobianco (2008); İlhan et al. (2022); Kippers et al. (2018); Nikou & Economides (2019); Tang et al. (2021); Taylor (2011)                     |
| Duration of research | Others                                | 10 (7.69%)           | Wang & Yuan (2023)                                                                                                                                                                                            |
|                      | One year and more                     | 28 (21.54%)          | Ranalli et al. (2017); Rutkowski et al. (2008a); Akpınar & Çetin (2023); Hooker (2017); Koedinger et al. (2010); Lekwa (2012); Sánchez Gómez et al. (2013); Rutkowski et al. (2008b) et al.                   |

Table 6 (continued)

| Category                | Sub-category        | Frequency (%) | Reference                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|-------------------------|---------------------|---------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Duration of research    | 7-11 months         | 0 (0.00%)     | Gil Pla-Campas et al. (2016); Hasan et al. (2021); Huang et al.(2021); Kleitman & Costa (2014); Lai et al. (2018); Macgregor et al.(2011); Martin & Polly (2015); Marín & Perez Garcias (2016); Qomaruddin et al. (2014); Robertson et al. (2019); Zou & Zhang (2013); Yuan & Kim (2018); Shin et al. (2020); Pesare et al. (2015); Nsabayeze et al. (2023); Raman & Nedungadi (2010); Vendlinski et al. (2005); Nadolski et al. (2021); Nsabayeze et al. (2022); Poláková & Klímová (2020); Sheard & Chambers (2014); Joyce, (2018)                                                                                                                                                                                                               |
|                         | 3-6 months          | 34 (26.15%)   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|                         | 1-2months           | 10 (7.69%)    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|                         | less than one month | 13 (10.00%)   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Research Contexts       | NA                  | 45 (34.62%)   | Liu et al. (2018); Georgiou et al. (2021); Hwang & Chang (2011); Broadbent et al. (2020); Chen & Chen (2009); Nikou & Economides (2021); Wang (2007); Kim & Ryu (2013); Lai et al. (2022); Veng (2023)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|                         | Class learning      | 90 (69.23%)   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|                         | Autonomous learning | 26 (20.00%)   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|                         | Working places      | 8 (6.15%)     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|                         | Others              | 6 (4.62%)     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Discipline distribution | Language learning   | 24 (18.46%)   | Alavi et al. (2021); Andrade & Heritage, 2018); Baleni (2015); Buldu & Buldu (2010) et al.; Bacca-Acosta & Avila-Garzon (2020); Bacca-Acosta et al. (2022); Broadbent et al. (2020); Lai et al. (2022) et al.; Duggan et al., 2020); Hwang & Chang, 2011); Rush et al., 2011); Sánchez Gómez et al., 2013) et al.; Vellanki et al., 2023).                                                                                                                                                                                                                                                                                                                                                                                                         |
|                         | Others              | 116 (89.23%)  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Methodology             | Qualitative         | 20 (15.38%)   | Alavi et al. (2021); Chen (2016); Cummins & Davesne (2009); Kapsalis et al. (2020); Bacca-Acosta et al. (2022); Rajendran & Yunus (2021); Bacca-Acosta & Avila-Garzon (2020); Hasan et al. (2021); Chen et al. (2010);Comiskey (2012); De Brún et al. (2022); Georgiou et al. (2021); Duggan et al. (2020) et al. Albinson et al. (2020); Antee (2021); Carey & Stefaniak (2018); Chen (2016); Cummins & Davesne (2009); Darzhinova & Zou (2021); Fernandez (2003); Forkosh - Baruch et al. (2021); Gikandi et al. (2011); Hasan et al. (2021); Hooker (2017); Huang et al. (2021); Kaya - Capocci et al. (2022); Kim et al. (2016); Martin & Polly (2015); Muslu (2017); Romero et al. (2015); Tlili et al. (2022); Veng (2023); Yesilyurt (2023) |
|                         | Quantitative,       | 27 (20.77%)   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|                         | Mixed               | 83 (63.85%)   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|                         |                     |               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |

### ***University, primary, and secondary school students as the primary participants***

In these studies, there is significant variation in participants ( $SD = 27.02$ ). Research focusing on postgraduates & undergraduates ( $n = 69$ ) and primary and secondary school students ( $n = 39$ ) dominates the field, with studies on postgraduates & undergraduates accounting for more than half of the total (52.31%). In contrast, there is only one study involving young children (0.77%) and two studies focusing on vocational college students (2.31%). This indicates that researchers predominantly regard university and school students as the primary subjects of their studies, while largely overlooking the learning assessment of vocational college students and young children.

The tendency to select university and school students as participants may be closely related to the ease of accessing samples. Universities and schools typically possess more structured teaching environments, and their information and communication technology (ICT) infrastructure is more developed. This facilitates the adoption and widespread use of web-based or mobile technology-based formative assessment tools, making it easier for researchers to collect large-scale sample data from these populations. In contrast, research involving young children is more challenging due to their younger age, requiring more complex ethical reviews and parental consent, which makes data collection a more time-consuming and intricate process.

The scarcity of studies on network/mobile technology-based formative assessments for vocational college students may also be attributed to the characteristics, learning goals, and learning environments of this demographic. Vocational college students are generally employment-orientated, with their learning content and assessment methods often focusing on the development of practical skills and workplace adaptability (Tynjälä, 2013). Consequently, vocational education activities frequently alternate between on-campus classrooms and off-campus workplace environments, where internship settings may not fully support the application of network- or mobile-based assessment tools (Mulder, 2017). Compared with traditional academic education, vocational education requires more contextual relevance and practical alignment in its assessments. However, network- or mobile-based formative assessments may face challenges in effectively simulating authentic vocational scenarios and providing in-depth feedback (Vaughan et al., 2013).

Additionally, the varying levels of digital literacy among vocational college students, particularly in regions with underdeveloped technological infrastructure, may hinder the promotion of network- or mobile-based formative assessment tools. Students' motivation and digital skills can also affect the effectiveness of these tools. Research has shown that vocational college students have more diverse motivations for learning, such as employment-driven or skill-enhancement goals. However, their acceptance of technology-based tools is often lower than that of university students (Eyadat, 2020; Zhu et al., 2023).

In summary, the characteristics, learning objectives, and learning environments of vocational college students significantly influence the challenges and feasibility of conducting research on network/mobile-based formative assessments. This suggests that future research should focus on developing tools tailored to the specific needs of vocational college students, improving the digital infrastructure in vocational colleges, and enhancing technical support and training for both teachers and students. Only by addressing these challenges can network/mobile-based formative assessments be more effectively integrated into vocational education to meet the unique needs of this group.

#### ***Long-term research designs as the mainstream approach***

In terms of research duration, a significant number of studies did not specify or elaborate on their duration ( $n = 45$ ). Among those that did, studies lasting 3–6 months ( $n = 34$ ) and over a year ( $n = 28$ ) constituted the majority. In contrast, short-term studies lasting 1–2 months accounted for only 7.69%, while those lasting less than a month comprised 10%. This suggests that most studies in educational research favour longer duration, as they allow researchers to better observe the continuity and stability of interventions or experimental effects. Furthermore, long-term research facilitates more comprehensive data collection, thereby enhancing the reliability and persuasiveness of research findings.

Long-term research designs hold significant value in education for evaluating the effectiveness of interventions and understanding the dynamic processes of student development. By extending the duration of studies, researchers can obtain more robust and reliable data, which enhances the credibility and validity of their findings (Rutter, 1988; Raudenbush & Bryk, 2002).

The distribution of research duration reflects the high demands of educational research for data accumulation and result verification. Future studies should tailor their duration to specific research objectives, striking a balance between the in-depth analysis provided by long-term studies and the exploratory efficiency offered by short-term research.

### ***Research concentrated on technology-enhanced classroom learning assessment***

In terms of research contexts, significant differences exist among various contexts ( $SD = 39.37$ ). Classroom teaching and learning assessments dominate the field ( $n = 90$ , 69.23%), while studies focusing on autonomous learning tools for formative assessment are fewer ( $n = 26$ , 20%). Additionally, formative assessments in workplace learning scenarios have been largely neglected ( $n = 8$ , 6.15%). This suggests that factors such as the appeal, feasibility, and practicality of different contexts greatly influence the number of studies conducted. Moreover, it highlights that educational research has prioritised traditional classroom teaching scenarios over emerging or specialised contexts such as autonomous learning and workplace settings.

Although technology-enhanced formative assessment tools for autonomous learning have begun to attract attention, their research remains limited due to challenges in technological development, resource investment, and application promotion. This indicates that the field is still in its developmental stages, offering significant potential for future exploration and application.

To promote comprehensiveness in educational assessment research, future studies should place greater emphasis on the application of technology in emerging contexts and non-traditional learning environments (e.g., autonomous learning and workplace settings). Exploring the implementation and impact of formative assessments in these scenarios can provide broader theoretical support and practical insights for educational assessment practices and policy making.

### ***Significant differences in discipline distribution***

Studies involving WBFAT in language learning accounted for only 18.46% ( $n = 24$ ), indicating a substantial disparity in research quantity across disciplines. Differences in research resources, attention levels, and application needs among disciplines may contribute to this imbalance. Researchers may be more inclined to explore the application

of WBFAT in other disciplines (e.g., STEM, medical education, social sciences), while underestimating its potential in language learning.

This could be due to the more complex design requirements of formative assessment tools in language learning, which often necessitate advanced technologies such as speech recognition and natural language processing. In some regions or research institutions, these technologies may still be in developmental stages. Additionally, formative assessment in language learning needs to address the diversity and complexity of linguistic features, including grammar, phonetics, and vocabulary. These unique demands make the development and application of WBFAT in language learning more challenging, potentially hindering the rapid advancement of research in this field.

### ***Mixed-methods research is widely adopted***

The distribution of research methodologies reveals significant variation ( $SD = 34.53$ ). Mixed-methods approaches are used most frequently ( $n = 83, 63.85\%$ ), reflecting their recognised ability to comprehensively consider multiple factors and deliver more holistic research results. In contrast, purely qualitative methods are used far less frequently ( $n = 20, 15.38\%$ ). Creswell and Plano Clark (2017), in their seminal work *Designing and Conducting Mixed Methods Research*, highlight that mixed methods, by integrating quantitative and qualitative data, offer multidimensional insights and are particularly suitable for complex research contexts where understanding statistical trends and underlying behavioural motivations is necessary. This explains why mixed methods dominate the distribution of methodologies in the studies reviewed. Of course, researchers' choices of methodologies are influenced by various factors, including research objectives, disciplinary domains, and the researchers' academic backgrounds (Bryman, 2016; Creswell & Plano Clark, 2017).

### **3.3.4 Theoretical Basis in the Included Literature**

As shown in Table 7, the publications encompass a wide range of theories, including learning theories, technology acceptance and application theories, assessment theories, cognitive theories, and sociocultural theories, reflecting the multidimensional perspectives in this field of research.

Table 7 Theoretical Basis in the Included Literature

| Fundamental theory                           | Sub-theory                                                                              | Frequency | References                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|----------------------------------------------|-----------------------------------------------------------------------------------------|-----------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Learning theories                            | Constructivist Learning Theory                                                          | 25        | Baleni (2015); Beatty & Gerace (2009); Bjørkli (2014); Carvalho & Santos (2022); Çevik (2015); Chang & Lin (2020); Chen & Chen (2009); Cummins & Davesne (2009); El Tantawi et al. (2015); Fuad et al. (2018); Gamulin et al. (2010); Georgiou et al. (2021); Ghanizadeh et al. (2015); Hay et al. (2013); Kim & Ryu (2013); Kowalski et al. (2013); Li et al. (2008); Liu et al. (2018); Martin & Polly (2015); Nsabayezu et al. (2022); Nsabayezu et al. (2023); Pesare et al. (2015); Rajendran & Yunus (2021); Rosen (2014); Sheard & Chambers (2014); Shirley & Irving (2015); Shin et al. (2020); Wen (2021) |
|                                              | Social Constructivism Theory                                                            | 4         | Andrade & Heritage (2018); Li et al. (2008); Shirley (2009); Shin et al. (2020)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|                                              | Cognitive Constructivist Theory                                                         | 1         | Chen & Chen (2009)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|                                              | Vygotsky's Sociocultural Theory                                                         | 3         | Chen (2016); Kim et al. (2016); Touw et al. (2019)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|                                              | Situated Learning Theory                                                                | 2         | Rajendran & Yunus (2021); Wen (2021)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|                                              | Distributed Cognition Theory                                                            | 1         | Wen (2021)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|                                              | Interactionist Second Language Acquisition Theory                                       | 1         | Chen (2016)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|                                              | Piaget's Constructivism                                                                 | 1         | Andrade & Heritage (2018)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|                                              | Self - regulated Learning Theory                                                        | 7         | Broadbent et al. (2020); Carey & Stefaniak (2018); Kleitman & Costa (2014); Lai et al. (2018); Mahroeian & Wee (2013); Marín & Perez Garcias (2016); Zou & Zhang (2013)                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|                                              | Self - efficacy Theory                                                                  | 1         | Carey & Stefaniak (2018)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|                                              | Achievement Goal Theory                                                                 | 2         | Carey & Stefaniak (2018); Wang & Lehman (2021)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|                                              | Self - regulation Theory in Bandura's Social Cognitive Theory                           | 1         | Sajadi et al. (2023)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|                                              | Ausubel's Learning Psychology                                                           | 1         | Buldu & Buldu (2010)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|                                              | Process - Oriented Writing Theory                                                       | 1         | Chen (2016)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|                                              | Collaborative Learning Theory                                                           | 1         | Chen (2016)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|                                              | Active Learning Theory                                                                  | 1         | Fuad et al. (2018)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|                                              | Information Theory, Bloom's Taxonomy of Learning Outcomes, Kolb's Learning Cycle Theory | 1         | Moscinska & Rutkowski (2018)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|                                              | Adult Learning Theory (Andragogy)                                                       | 1         | Taylor (2011)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|                                              | Theory of Multiple Intelligences, Humanistic Theory                                     | 1         | Li (2022)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Technology acceptance and application theory | Technology Acceptance Model (TAM)                                                       | 7         | Antee (2021); Bacca - Acosta & Avila - Garzon (2020); Deutsch et al. (2012); Huang et al. (2021); Nikou & Economides (2019); Rajendran & Yunus (2021); Tang et al. (2021); Zhai & Ma (2021)                                                                                                                                                                                                                                                                                                                                                                                                                        |

Table 7 (continued)

| Fundamental theory                           | Sub-theory                                                                      | Frequency | References                                                                                                                                |
|----------------------------------------------|---------------------------------------------------------------------------------|-----------|-------------------------------------------------------------------------------------------------------------------------------------------|
| Technology acceptance and application theory | Technology Acceptance Model 2 (TAM2)                                            | 1         | Cigdem, H. & Oncu, S. (2015)                                                                                                              |
|                                              | Unified Theory of Acceptance and Use of Technology (UTAUT)                      | 1         | Tang et al. (2021)                                                                                                                        |
|                                              | Expectation Confirmation Model                                                  | 2         | Antee (2021); Huang et al. (2021)                                                                                                         |
|                                              | Theory of Planned Behavior (TPB)                                                | 1         | Tang et al. (2021)                                                                                                                        |
| Evaluation theories                          | Formative Assessment Theory                                                     | 6         | Gil Pla - Campas et al. (2016); Hasan et al. (2021); Hooker (2017); Kippers et al. (2018); Rutkowski et al. (2008a); Veugen et al. (2021) |
|                                              | Feedback Theory (Hattie and Timperley's Feedback Model)                         | 1         | Faber et al. (2017)                                                                                                                       |
|                                              | Miller's Technical Competence Assessment Framework                              | 2         | Havnes et al. (2012); Rush et al. (2011)                                                                                                  |
|                                              | Sadler's Formative Assessment Theory                                            | 1         | Havnes et al. (2012)                                                                                                                      |
|                                              | Argument - Based Validation Theory (Kane's Theory)                              | 1         | Ranalli et al. (2017)                                                                                                                     |
|                                              | Personalized Feedback Theory                                                    | 1         | Wang & Lehman (2021)                                                                                                                      |
|                                              | Cognitive Diagnostic Modeling (CDM)                                             | 1         | Akpınar & Çetin (2023)                                                                                                                    |
|                                              | ECD Theory, SAMR Model                                                          | 1         | Webb & Gibson (2015)                                                                                                                      |
| Cognitive Theory                             | Cognitive Theory                                                                | 2         | Hamutoglu et al. (2020); Koedinger et al. (2010)                                                                                          |
|                                              | Cognitive Development Theory                                                    | 1         | Vendlinski et al. (2005)                                                                                                                  |
|                                              | Cognitive Load Theory                                                           | 2         | Hwang & Chang (2011); Lachner et al. (2017)                                                                                               |
|                                              | Cognitive Scaffolding Theory                                                    | 1         | Fernandez (2016)                                                                                                                          |
|                                              | Teacher Cognitive Development Model                                             | 1         | Forkosh - Baruch et al. (2021)                                                                                                            |
|                                              | Cognitive Style Theory                                                          | 1         | Wang (2007)                                                                                                                               |
| Socio-cultural theory                        | Mediation and Zone of Proximal Development (ZPD) Theory in Sociocultural Theory | 1         | Fathi et al. (2023)                                                                                                                       |
|                                              | Sociocultural Learning Theory                                                   | 1         | Muslu (2017)                                                                                                                              |
| Other theories                               | Behaviorism Theory                                                              | 1         | Andrade & Heritage (2018)                                                                                                                 |
|                                              | Conceptual Change Theory                                                        | 1         | Beatty & Gerace (2009)                                                                                                                    |
|                                              | Herzberg's Two - Factor Theory                                                  | 1         | Hamutoglu et al. (2020)                                                                                                                   |
|                                              | Testing Effect Theory                                                           | 1         | Joyce (2018)                                                                                                                              |
|                                              | Integrated Behavioral Prediction Model (IMBP)                                   | 1         | Lai et al. (2022)                                                                                                                         |
|                                              | Ontology                                                                        | 1         | Romero et al. (2015)                                                                                                                      |
|                                              | Student Engagement Theory, Critical Thinking Theory                             | 1         | Veng (2023)                                                                                                                               |
|                                              | Growth Mindset Theory                                                           | 1         | Tan et al. (2021)                                                                                                                         |
|                                              | Activity Theory (AT)                                                            | 1         | Zhai & Ma (2021)                                                                                                                          |
|                                              | Grey System Theory, Fuzzy Set Theory                                            | 1         | Chen & Chen (2006)                                                                                                                        |
|                                              | Data Mining Theory                                                              | 1         | Chen & Chen (2009)                                                                                                                        |
|                                              | Validity Theory, Social Semiotics Theory, Multimodal Theory                     | 1         | Fjortoft (2020)                                                                                                                           |
|                                              | ICAP Framework Theory                                                           | 1         | Wen (2021)                                                                                                                                |

### *Learning theories*

In terms of learning theories, constructivist learning theory and its derivative theories are widely cited, with constructivist learning theory being mentioned 25 times. Researchers frequently reference constructivist theories because, in networked and mobile learning environments, students can engage in learning in a more autonomous, interactive, and contextualised manner, which aligns closely with the constructivist emphasis on learners actively constructing knowledge. For instance, in online learning platforms, students can participate in discussions, collaborative projects, and other activities to construct knowledge systems based on their experiences and understanding. This is exemplified in studies by scholars such as Baleni (2015) and Beatty & Gerace (2009).

Vygotsky's sociocultural theory, particularly the concept of the Zone of Proximal Development (ZPD), is referenced three times, highlighting the importance of social interaction and cultural context in learning. In online learning, students can transcend geographical barriers and interact with people from diverse backgrounds, expanding their cognitive boundaries. This is evident in studies such as those by Chen (2016) and Kim et al. (2016).

Self-regulated learning theory and related theories have also gained attention, being mentioned seven times. This is because, in networked and mobile learning environments, students require greater autonomy and self-management skills. These theories help to explain how students set goals, monitor their learning processes, and adjust their strategies in such environments. Examples include studies by Broadbent et al. (2020) and Carey & Stefaniak (2018).

### *Technology acceptance and application theories*

In the realm of technology acceptance and application theories, the Technology Acceptance Model (TAM) and related models are cited 12 times. With the continuous emergence of networked and mobile-based formative assessment tools, understanding how students and teachers accept these technologies is critical. For instance, scholars such as Antee (2021) and Bacca-Acosta & Avila-Garzon (2020) have used TAM and related models to investigate factors influencing user acceptance and adoption of electronic assessment systems.

### *Assessment theories*

Formative assessment theory and related theories are frequently cited in this field, with formative assessment theory itself being mentioned six times. In networked and mobile learning environments, the modes and methods of formative assessment have evolved. These theories provide guidance on designing effective online assessment activities, delivering instant feedback, and facilitating student learning. For instance, studies by Gil Pla-Campas et al. (2016) and Hasan et al. (2021) explore how network technologies can be used to continuously monitor and support students' learning processes.

#### *Cognitive theories*

Cognitive theories, including general cognitive theory and cognitive development theory, are also referenced. These theories are integral to designing learning resources and assessment tasks in networked and mobile learning environments, as they consider students' cognitive characteristics and developmental stages. This ensures that learning content and assessment methods align with students' cognitive levels and foster effective knowledge construction. Examples include studies by Hamutoglu et al. (2020) and Vendlinski et al. (2005), which apply cognitive theories to optimise the presentation of online learning resources and the design of assessment tasks.

#### *Sociocultural theories*

Sociocultural theories, particularly the concepts of mediation and the ZPD, are referenced once, emphasising the mediating role of social and cultural factors in learning and the potential developmental levels individuals can achieve with assistance. In networked and mobile learning environments, elements such as online communities and digital tutoring can provide various forms of support for students, aligning with the principles of sociocultural theory. This is reflected in research by Fathi et al. (2023).

#### *Other related theories*

Other theories, such as behaviourist theory and conceptual change theory, are also mentioned. These theories provide complementary perspectives on the mechanisms underlying learning and behaviour, contributing to a more comprehensive understanding of networked and mobile learning processes. For example, behaviourist theory highlights the impact of environmental stimuli on behaviour, offering guidance for designing reward mechanisms and feedback systems in online learning environments. Conceptual change theory, on the other hand, aids in understanding how students modify and refine their

knowledge concepts during online learning, as demonstrated in studies by Andrade & Heritage (2018) and Beatty & Gerace (2009).

The theories referenced in studies on web-based or mobile-based formative assessment aim to deepen our understanding of students' learning processes, the feasibility of technological applications, and the effectiveness of assessments from various perspectives. They collectively construct a multidimensional theoretical framework, providing guidance for the practice and research of networked and mobile-based formative assessment. By drawing on these diverse theoretical perspectives, researchers can optimise learning environments, refine assessment tools, and enhance learning outcomes.

### **3.3.5 Web-based or Mobile-based Formative Assessment Tools Utilised in the Literature and Their Impacts on Learning**

An analysis of 130 selected studies on WBFAT, as shown in Table 8, including those based on mobile technology, highlights their varied impact on teaching, learning, and assessment. While these tools offer significant benefits, they also pose challenges. Moreover, the findings reveal users' expectations in different areas, underscoring both their potential and the need for further refinement.

Table 8 WBFAT, MFAT Utilised and Related Findings in the Included Literature

| Catalogue                                 | Sub-catalogue                                      | Description                                                                                                                                                                                                                                                  | Tool                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | Reference                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|-------------------------------------------|----------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Positive comments</b>                  | Improve learning effect                            | Promote students' academic performance, knowledge and skills enhancement, learning strategy application, and knowledge understanding and mastery through various means, covering different disciplines and learning stages                                   | FACT, SimScientists, Nearpod, Snappet, Mobile Response System(MRS), Pitalica, Oculus Rift, WFPAS, InkSurvey, WriteToLearn, Mobile learning operating system based on iOuerv, Analytic Rubrics, PDA - WATA, Questions for Learning(OFL), TI - Navigator™ System, TI - Navigator, mobile tools based on UML, ACME - DB, Web - based College English Test of Band - 4                                                                                                                                                                                             | Andrade & Heritage (2018); Lai et al. (2018); Veng (2023); Faber et al. (2017); Fuad et al. (2018); Gamulin et al. (2010); Georgiou et al. (2021); Kim & Ryu (2013); Kowalski et al. (2013); Landauer et al. (2009); Martin & Polly (2015); Nikou & Economides (2018b); Nsabavezu et al. (2022); Mahrooian & Wee (2013); Sheard & Chambers (2014); Shirley & Irving (2015); Shirley (2009); Soler et al. (2010); Zou & Zhang (2013) Baleni (2015); Chen (2016); Li et al.(2008); Liu et al.(2018); Marín & Perez Garcias (2016); Nikou & Economides (2018a); Elmahdi et al. (2018); Touw et al. (2019); Wen (2021)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|                                           | Enhance interaction and collaboration              | Stimulate students' interactive participation, cooperation, motivation, improve students' performance and attitude, and assist teaching and assessments                                                                                                      | Blackboard, Daedalus InterChange , Web-based Peer Assessment Site, Web PAI.S, Moodle Workshon (Plug-in) , Nearpod, Socrative , Plickers, GM:Studio (Yoyogames), Raven's Progressive Matrices, ARIS, HP Reveal                                                                                                                                                                                                                                                                                                                                                  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|                                           | Provide effective feedback                         | Provide accurate and high-quality feedback, correct mistakes, improve performance and self-awareness, promote objective assessment and enhance assessment-related abilities                                                                                  | Criterion Online Writing Evaluation Service, CATME, GPT - 4, Concept Map, one45 Software, Socio - Scientific Inquiry Network (SSINet)                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|                                           | Promote teaching and resource utilisation          | Supporting teaching, including providing resources, assisting in the implementation of teaching methods, improving teaching efficiency and fairness, skills certification and motivation.                                                                    | Concept Map, CRS, Digital Badges, Camtasia screencasting, Blackboard, elate portal, eClinic Card, Maple T.A., Moodle, Llamas - Nistal et al. self-developed tool, GAM - WATA                                                                                                                                                                                                                                                                                                                                                                                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| <b>Negative comments &amp; challenges</b> | Technical problems and implementation difficulties | Face multiple challenges in technology application, implementation process, and student feedback, including technical problems, negative student attitudes, implementation difficulties, and lack of empirical support in some aspects of teaching methods   | Oxford Placement Test, K - English, m - LMS, Maple T.A., WebLearn, MDCA , Moodle (Optional) , MS Teams, Blackboard, Moodle, Web - based Peer Assessment Site, Criterion Online Writing Evaluation Service, eClinic Card, WhatsApp, Blackboard Collaborate, Analytic Rubrics, Nearpod, Concept Map, CRS, PeLe, ACCELERATED MATH, STAR MATH, Web PALS, Moodle, TI - Navigator, Study Island                                                                                                                                                                      | Beatty & Gerace (2009); Carev & Stefaniak (2018); Comiskey (2012); Deutsch et al. (2012); Duggan et al. (2021); Barana & Marchisio (2016); Hussmann et al. (2004); Moscinska & Rutkowski (2018); Llamas - Nistal et al. (2013); Mahrooian & Wee (2013)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| <b>user expectation</b>                   | Improve tool performance and functionality         | Expect to improve existing tools and teaching assessment methods in multiple aspects such as solving technical problems, optimizing functions, obtaining resource support, and enhancing teaching-related aspects to meet the needs of teaching and learning | Oxford Placement Test, LMS, Camtasia screencasting, Blackboard, eClinic Card, Snappet, Oculus Rift, WFPAS, Blackboard Collaborate, Web PALS, Analytic Rubrics, CATME, GPT - 4, TI - Navigator, GM:Studio(Yoyogames), Raven's Progressive Matrices, Nearpod, PISA 2015, ECD, ARIS, HP Reveal, Web - based College English Test of Band - 4, m - LMS, Maple T.A., Digital Badges, e - ATLMS, WebLearn, WebTutor, WhatsApp, ACCELERATED MATH, STAR MATH, Evaluation Index System, Moodle LMS, Criterion Online Writing Evaluation Service, Study Island, MS Teams | Alavi et al. (2021); Bacca - Acosta & Avila - Garzon (2020); Bacca - Acosta et al. (2022); Antee (2021); Barana & Marchisio (2016); Dahalan & Hussain (2010); Fernandez (2003), Fjortoft (2020); Rutkowski et al. (2008a); Rutkowski et al. (2021); Vellanki et al. (2023); Baleni (2015); El Tantawi et al. (2015); Li et al. (2008); Ranalli et al. (2017); Duggan et al. (2021); Hasan et al. (2021); Li et al. (2020); Nsabavezu et al. (2022); Veng (2023); Buldu & Buldu (2010); Beatty & Gerace (2009); Björkli (2014); Lekwa (2012); Liu et al. (2018); Pesare et al. (2015); Shirley (2009); Taylor (2011); Alavi et al. (2021); Chen et al. (2010); Comiskey (2012); Duggan et al. (2021); Faber et al. (2017); Georgiou et al. (2021); Kim & Ryu (2013); Li et al. (2020); Liu et al. (2018); Nsabavezu et al. (2022); Sajadi et al. (2023); Shirley (2009); Touw et al. (2019); Veng (2023); Webb & Gibson (2015); Wen (2021); Zou & Zhang (2013); Antee (2021); Barana & Marchisio (2016); Carev & Stefaniak(2018); Dahalan & Hussain(2010); Fernandez (2003); Hasan et al. (2021); Lekwa (2012); Li (2022); Pesare et al. (2015); Ranalli et al. (2017); Taylor (2011); Vellanki et al. (2023) |

## ***Positive aspects***

### ***Significant improvement in learning outcomes***

Research has shown that many tools, such as Nearpod and Snappet, are highly effective in enhancing student performance and improving knowledge processing abilities. For instance, Nearpod's self-regulated scientific inquiry learning approach not only contributes to improved academic achievement but also fosters skills in information-seeking, self-efficacy, and time management (Lai et al., 2018). This aligns with constructivist learning theory, which emphasises knowledge construction through autonomous exploration, thereby enhancing learning outcomes. Similarly, Vygotsky's sociocultural theory posits that the interactive and resource-rich environment provided by technology enables students to develop within their ZPD. The combined use of tools such as FACT and SimScientists has also been found to enhance learning effectiveness and promote self-regulated learning (Andrade & Heritage, 2018), which reflects the role of technology-enhanced assessment in facilitating teaching and learning through real-time monitoring and targeted feedback, in line with formative assessment theory's emphasis on continuous improvement.

### ***Enhanced engagement and interaction***

The online discussion feature of Blackboard provides a platform for students to engage in knowledge sharing and collaborative learning (Baleni, 2015), which promotes student interaction and cooperation, consistent with the principles of social constructivism, which highlights the importance of social interactions in knowledge construction. Additionally, peer assessment tools such as the Web-Based Peer Assessment Site have been found to enhance student motivation, subject comprehension, and interactive learning (Li et al., 2008). Through peer feedback, students gain exposure to diverse perspectives, broadening their cognitive outlook and facilitating deeper knowledge construction and understanding.

### ***Effective feedback to support learning***

Some tools provide highly precise feedback on specific types of errors, aiding students in identifying and correcting mistakes while alleviating cognitive burden. For example, Criterion Online Writing Evaluation Service has been shown to offer targeted feedback that supports error correction and reduces anxiety (Ranalli et al., 2017).

Immediate and accurate feedback is a core component of formative assessment, guiding students in refining their learning strategies and improving performance, which is also reflected in behaviourist learning theory, which underscores the role of feedback in reinforcing positive learning behaviours. Tools such as CATME and GPT-4 generate high-quality feedback, promoting student performance and self-awareness (Sajadi et al., 2023). These findings highlight the advantages of modern technology in delivering personalised and adaptive feedback, catering to the diverse needs of learners.

### ***Negative comments and challenges***

#### ***Technical issues and implementation difficulties***

Many tools, such as the Oxford Placement Test, face technical challenges, dependency concerns, and issues related to digital literacy (Alavi et al., 2021), which underscore the need to consider the technological infrastructure and digital competencies of schools and students when adopting WBFA, as inadequate preparation may compromise their effectiveness. Additionally, as highlighted by Forkosh-Baruch et al. (2021), a lack of teacher expertise and issues with adaptive software can undermine educators' decision-making roles, suggesting that teacher training and software optimisation are critical to ensuring the successful integration of technology into teaching and learning.

#### ***Student concerns and negative perceptions***

Students have expressed concerns about certain tools, particularly regarding exam pressure, security issues, and a preference for traditional paper-based assessments, as seen with platforms such as the Oxford Placement Test (Alavi et al., 2021), which highlight the importance of considering students' psychological experiences and learning habits when designing and implementing such tools, to enhance acceptance and engagement. Additionally, some tools, such as the Criterion Online Writing Evaluation Service, have been found to provide inaccurate feedback in specific cases (Ranalli et al., 2017), affecting student trust and the overall effectiveness of these tools. Students' concerns reflect the importance of ensuring the reliability and accuracy of automated assessment tools.

#### ***Other potential limitations***

Some tools, while recognised for their benefits, also present usability challenges. For example, although concept mapping has been acknowledged as valuable by pre-service

teachers, its complexity, requirement for prior knowledge, and potential to cause frustration highlight the need for improvements in usability and accessibility (Buldu & Buldu, 2010). Similarly, aspects of the CRSTEFA instructional approach lack sufficient empirical support (Beatty & Gerace, 2009), indicating the need for further research to validate the efficacy of innovative pedagogical methods and technologies.

### ***User expectations***

#### ***Demand for improvement and enhancement***

Students and teachers express a strong desire for improvements in feedback quality, response speed, and task design. For instance, students in Li et al.'s (2020) study sought solutions to technical issues with video feedback and desired a reduction in redundancy. Similarly, Liu et al. (2018) highlight the need to enhance students' ability to provide evaluative feedback and make informed judgements. These expectations reflect a broader demand for continuous refinement of formative assessment tools to better support teaching and learning. In terms of assessment effectiveness, students in Duggan et al.'s (2020) study called for improvements in formative guidance feedback, underscoring the recognised importance of high-quality feedback in facilitating learning and skill development.

#### ***Stronger support and resource accessibility***

Some studies have highlighted concerns about digital literacy and economic constraints, with students from disadvantaged backgrounds expressing the need for schools and teachers to provide technological resources and internet access (Antee, 2021). This highlights the issue of educational equity in the implementation of digital assessment tools, emphasising the necessity of ensuring that all students can benefit from these innovations. Teachers also seek greater institutional support for technology integration. For example, Barana and Marchisio (2016) argues that the implementation of Maple T.A. faces resistance and technological barriers in underprivileged areas, with teachers requesting additional resources to facilitate effective technology adoption in their teaching.

In summary, WBFAT holds significant potential in education, yet it remains in a continuous state of development and refinement. From a positive perspective, these tools introduce new possibilities for teaching and learning, aligning with multiple learning theories and contributing to pedagogical innovation. However, challenges such as technological barriers and student acceptance must be addressed to ensure their effective

implementation. Users' expectations provide valuable insights for the future enhancement of these tools, guiding improvements in feedback quality, usability, and accessibility. Future research and practice should focus on overcoming these challenges to maximise the benefits of WBFAT, ultimately enhancing teaching quality and fostering students' holistic development. Additionally, interdisciplinary research integrating educational theory, technological advancements, and empirical practice is essential to ensuring the sustainable development of these tools. In an evolving educational landscape, continuous attention to user needs, optimisation of tool performance, and the seamless integration of technology into pedagogy will be key to modernising education and improving learning outcomes.

### **3.4 Review of Relevant Theories**

#### **3.4.1 Learning Theories**

##### ***Metacognitive theory***

Metacognitive theory holds that learners can manage their learning more effectively by monitoring and regulating their cognitive processes. It has garnered significant attention in educational research, particularly regarding its implications for language learning and self-regulated learning. Flavell (1979) proposes that metacognition encompasses both metacognitive knowledge and metacognitive regulation; the former refers to learners' awareness of their learning processes, while the latter pertains to their capacity to self-monitor and reflect during learning. Brown (1987) and O'Malley (1990) highlight that metacognitive regulation is particularly critical for language learning, as it enables students to identify and address weaknesses in their expression. Studies have shown that learners who possess meta-cognitive awareness, understanding their own learning processes, are more likely to succeed academically (Rogti, 2020; Tong, 2012; Wijaya, 2021; Dağal & Bayındır, 2016). For instance, it was found that metacognitive awareness is closely linked to the effective use of language learning strategies, which in turn enhances learners' self-efficacy and overall performance in second language acquisition (Nosratinia et al., 2014; Di et al., 2022).

Practical activities derived from metacognitive theory include self-assessment, reflection, and strategic planning. These activities encourage learners to take ownership of their learning, allowing them to identify their strengths and weaknesses, set realistic goals, and monitor their progress (Masoodi, 2018). For example, the use of digital narratives has

been identified as an effective tool for developing metacognitive skills, as it engages students actively in their learning process and encourages them to reflect on their understanding (Lavrysh et al., 2023). Such approaches not only enhance learners' meta-cognitive abilities but also promote deeper engagement with the material, leading to improved academic outcomes (Bakar & Ismail, 2020).

Furthermore, the role of self-regulation in conjunction with metacognition is critical for fostering learner autonomy. Research indicates that students who engage in self-regulated learning characterised by setting goals, self-monitoring, and self-reflection are more likely to achieve academic success (Aydemir & Karaman, 2017; Naug et al., 2011). This is particularly relevant in the context of foreign language learning, where learners must often navigate complex linguistic and cultural landscapes. By promoting meta-cognitive strategies, educators can empower students to become more autonomous and effective language learners (Kistner et al., 2010; Jones, 2017). Paris and Winograd (1990) emphasise that metacognitive strategies, such as self-monitoring and self-evaluation, are instrumental in helping language learners improve their oral expression. With feedback from online tools, students can effectively reflect on and adjust their language use. Schraw and Moshman (1995) further highlight the necessity of combining metacognitive knowledge and regulation within online learning; immediate feedback allows students to recognise their weaknesses more clearly and improve their spoken performance. Similarly, White and Frederiksen (1998) design a 'reflection log' and find that students significantly enhance their learning outcomes through reflective journaling and feedback, with self-reflection proving particularly valuable for more precise language expression. Teng (2020) demonstrates that metacognitive reflection modules in online tools lead to marked improvements in students' speaking and writing skills, reflecting the effectiveness of strategy adjustment guided by feedback. Veenman et al. (2006) indicate that training in metacognitive strategies enhances learners' ability to self-regulate language use, with immediate feedback from online tools enabling students to reflect and self-adjust more efficiently in oral tasks.

Critically, while the benefits of meta-cognitive strategies are well-documented, there are challenges in their implementation, particularly in web-based learning environments. The integration of technology into language learning can sometimes

overwhelm students, especially those who may lack prior experience with digital tools (Wijayanti et al., 2023; Fauzi & Hussain, 2016). Therefore, it is essential for educators to provide structured guidance and support to help learners navigate these technologies effectively (Rad, 2018). Moreover, the need for tailored interventions that consider individual differences in learners' meta-cognitive skills is paramount, as not all students will benefit equally from generic strategies (Γούδας et al., 2015; Bulut, 2018).

Metacognitive learning theory offers valuable insights into the processes that underpin effective learning. By fostering metacognitive awareness and self-regulation, educators can enhance students' language learning experiences, particularly in web-based environments. However, it is crucial to address the challenges associated with technology integration and to provide tailored support to meet the diverse needs of learners.

### ***Self-regulating learning theory***

Self-Regulated Learning (SRL) theory, proposed by Zimmerman (2002), emphasises learners' active control over their behaviours, cognition, and motivation in the learning process. This involves goal setting, strategy use, learning monitoring, and self-reflection, all of which help learners continuously optimise their learning outcomes. Pintrich (2000) and Overton & Potter (2008) further describe self-regulation as an active process where learners adjust strategies and monitor their progress to address challenges. In the context of language learning, SRL theory is particularly relevant to the development of English-speaking skills. Online learning environments, especially those equipped with formative assessment tools, provide learners with effective support in achieving these goals. The different stages of SRL naturally align with the features of online formative assessment tools. During the goal-setting stage, learners can establish phased objectives based on modules or tasks provided by the tools. In the strategy use stage, these tools offer a variety of resources, such as speech practice, real-time feedback, and other supporting materials. At the monitoring and reflection stage, learners can analyse data generated by the tools (e.g., speech scores and progress charts) to identify weaknesses and adjust their strategies accordingly.

Research on the integration of SRL with online formative assessment tools has primarily focused on the tools' roles in goal setting, strategy use, monitoring and reflection, and behaviour adjustment in the learning process.

### *Supporting goal setting with tools*

Web-based or online formative assessment tools were proposed to enable learners to visualise their objectives and task modules, helping them gain a clearer understanding of their learning direction and expected outcomes. Dabbagh and Kitsantas (2012) propose that digital platforms that incorporate task breakdowns and goal representation improve learners' sense of direction and engagement. For instance, on the FiF platform, learners can set specific speaking objectives, such as enhancing pronunciation accuracy or fluency, and achieve these through phased tasks. Similarly, Awidi and Paynter (2016) emphasise that personalised task objectives and step-by-step learning pathways play a crucial role in building learners' confidence.

### *Optimising strategy use with tools*

The selection of appropriate strategies is essential to achieving learning goals. SRL theory highlights the importance of learners adopting strategies that meet their individual needs to maximise learning efficiency. Online formative assessment tools support learners by offering diverse strategy aids, such as speech imitation, grammar correction, and speech evaluation feedback. González-Lloret (2020) observes that task-based platforms guide learners in selecting suitable expressions and improve pragmatic strategies through collaborative tasks. Additionally, Kusz (2023) argues that interactive activities and peer evaluation features encourage learners to use varied strategies, such as improving their pronunciation by imitating the intonation of others.

### *Personalised feedback in monitoring and reflection with tools*

The real-time feedback functionality of online formative assessment tools is crucial for the monitoring and reflection stages of SRL. Learners can use this feedback to adjust their strategies in real time and identify areas for improvement through reflection. Zimmerman and Schunk (2013) emphasise the importance of feedback and self-monitoring for task completion, highlighting that feedback systems embedded in online tools can significantly enhance self-regulatory skills. Similarly, Winne and Hadwin (1998) explore the application of self-regulated learning systems in online contexts, demonstrating how feedback and progress tracking enable students to manage oral language tasks effectively. Broadbent and Poon (2015) review the application of self-regulation strategies in online learning, concluding that feedback substantially improves students' time management and

task completion abilities. Broadbent (2018) further demonstrates that self-regulation strategies in online learning environments effectively enhance time management skills and language learning outcomes. Panadero and Järvelä (2019) believe in the role of feedback in facilitating self-monitoring and strategy adjustment in language learning tasks within digital environments. Suraworachet et al. (2022) find the integration of human and analytical feedback mechanisms significantly enhances student engagement and improves their performance in reflective writing tasks. Similarly, Kumar et al. (2024) observe in their evaluation of large language models in classroom settings that students who received immediate and personalised feedback reported increased confidence and demonstrated stronger performance in subsequent assessments.

#### *Data-driven behavioural adjustment with tools*

Research has demonstrated the data analytics capabilities of online tools in providing a reliable foundation for reflection and behavioural adjustment in SRL. Learners can evaluate their performance using reports generated by the tools (e.g., speech fluency trend graphs or grammar error statistics) and refine their strategies accordingly. Ramaswami et al. (2023) examine the effectiveness of learning analytics dashboards (LADs) and find that descriptive analytics helps students identify areas for improvement, while predictive analytics encourages them to set specific goals for enhancement. Lounis et al. (2024) argue in their review that the feedback data provided by advanced automatic speech recognition (ASR) technology in pronunciation error detection (APED) systems enables language learners to accurately identify mispronunciations and enhances their speaking skills. Deeva et al. (2021) explore the impact of computer-based tutoring systems, emphasising how personalised feedback helps learners refine their learning strategies. Similarly, Kumar et al. (2024) find AI-powered adaptive learning systems enhance students' motivation and self-regulation compared to those without such support.

Although previous studies have explored specific functionalities of SRL and online tools, such as goal setting and feedback mechanisms, certain challenges remain. First, while studies have addressed individual stages of SRL, such as goal setting or monitoring, there is insufficient research on the dynamic interaction across stages. For instance, how learners transition from goal setting to strategy use and ultimately engage in reflection for optimisation remains underexplored (Pintrich, 2004). Second, most studies focus on

intermediate or advanced learners, while low-proficiency learners often face challenges in understanding goals, interpreting feedback data, and selecting appropriate strategies. Jin et al. (2023) emphasise that AI applications can effectively support self-regulated learning in online environments, particularly when designed with an intuitive and engaging interface. Likewise, Çakiroğlu et al. (2024) put forward that enhancing the accessibility and usability of online learning tools can improve learner engagement and gradually fosters independent learning skills, especially among those with limited prior experience. Moreover, cultural differences and individual differences significantly impact learning behaviours. Different learners may have large individual differences in their acceptance of these strategies, especially those with low self-discipline. For instance, Asian learners tend to rely more on teacher feedback, while Western learners may prefer independent exploration. Current online tools often fail to consider these variations in their design (Hampden-Turner & Trompenaars, 2020).

### ***Motivational theory***

In this study, the personalised feedback mechanisms and autonomous task selection facilitated by WBFAT align with the motivational needs for ‘Autonomy’ and ‘Competence’ identified in motivation theory. Motivation theory underscores the critical role of intrinsic motivation in autonomous learning. Deci and Ryan’s (1985; 2013) Self-Determination Theory (SDT) proposes that intrinsic motivation is stimulated when learners’ needs for autonomy, competence, and relatedness are satisfied, resulting in improved learning outcomes. Similarly, Dörnyei’s (2014) theory of the ‘Motivational Self System’ highlights that language learning motivation is strengthened through learners’ ‘Ideal Self’ and ‘Ought-to Self’, particularly when supported by personalised task selection and achievement feedback. These elements foster greater motivation and engagement.

Extensive research has demonstrated that web-based or technology-supported tools can provide learners with autonomy and immediate feedback, thereby stimulating intrinsic motivation and enhancing their enthusiasm and engagement in oral expression. Noels et al. (2000) identify that autonomy and competence significantly enhance intrinsic motivation, with feedback mechanisms in online tools effectively meeting learners’ autonomy needs. Li and Dewaele (2018) find web-based evaluation tools, through feedback and autonomous task selection, fulfil learners’ need for control, thereby improving their motivation and

engagement. Ushioda (2003) emphasises the importance of personalised feedback in boosting intrinsic motivation, particularly through immediate feedback that enhances learners' sense of achievement and encourages active participation in language tasks. Alamri et al. (2020) propose that applying personalised learning principles in online courses enhances learners' autonomy and competence, thereby increasing intrinsic motivation and engagement. Similarly, Clément et al. (2024) think that integrating machine learning with learner choice mechanisms in intelligent tutoring systems not only improves learning performance but also strengthens motivation, further underscoring the critical role of personalised feedback in fostering learner engagement and motivation.

These studies provide robust evidence for the relationship between intrinsic motivation, feedback, and autonomy, particularly underscoring the pivotal role of personalised feedback in fostering students' engagement and improving language learning outcomes. However, several limitations remain. Most research focuses on general online learning environments, lacking in-depth exploration of their applicability to specific language learning tasks, particularly concerning whether students with varying proficiency levels benefit equally from personalised feedback. Additionally, cultural contexts and individual differences are often neglected, as the majority of studies centre on Western learners. For example, Chinese learners may exhibit greater reliance on teacher guidance and display different preferences for autonomy and feedback formats compared to their Western counterparts.

Another limitation lies in the insufficient analysis of feedback quality and format. While the role of personalised feedback is acknowledged, little attention is given to how factors such as feedback quality (high vs. low) or format (immediate vs. delayed, standardised vs. personalised) impact learning outcomes. Poor-quality or overly complex feedback, for instance, may hinder learning rather than support it. Furthermore, most studies prioritise short-term benefits, leaving the long-term sustainability of intrinsic motivation and autonomy underexplored. It remains uncertain whether students can maintain motivation and autonomy after transitioning away from web-based tools, highlighting the need for further research into these areas.

Therefore, these studies provide substantial theoretical and empirical support for web-based approaches to enhancing motivation, demonstrating the critical role of

personalised feedback and autonomous task selection in increasing students' intrinsic motivation and improving language learning outcomes. Nonetheless, future research should delve deeper into areas such as cultural differences, feedback formats and quality, and the long-term effects of motivation strategies, which is essential for optimising the design of web-based tools and providing more targeted and sustainable support for language learning.

### ***Constructivist learning theory***

Constructivist learning theory posits that learners actively construct knowledge through their interactions with the environment and others, rather than passively receiving external information (Piaget, 1970). This theory emphasises the agency of learners, asserting that knowledge is progressively built through their experiences, reflection, and exploration, as opposed to being directly transmitted by teachers. Vygotsky's (1978) social constructivism extended this notion, highlighting the pivotal role of social interaction in learning. Through collaboration with others, learners could develop skills within their ZPD, a concept particularly relevant to language acquisition (Vygotsky, 1978).

Another significant branch of constructivism is self-constructivism, which focused on how individuals built personal knowledge systems through independent exploration. autonomous learning underlined the ability of learners to acquire knowledge without external intervention. According to Bandura (1997), learners developed autonomy by actively selecting learning resources, setting goals, and managing their learning processes, which was particularly relevant to enhancing spoken English in authentic contexts.

Constructivist principles provided a robust foundation for autonomous learning, especially in the context of technology-enhanced learning (TEL). Digital learning environments offered learners a wealth of resources and tools, enabling them to regulate their learning pace, tailor content to their needs, and adjust strategies based on immediate feedback. Dabbagh and Kitsantas (2012) observe that TEL environments empowers autonomous learners by granting them greater control over their personalised learning pathways. Specifically, WBFAT allows learners to refine their strategies through personalised guidance and real-time feedback, aligning seamlessly with the constructivist emphasis on learner-centred approaches (Nicol & Macfarlane-Dick, 2006). For instance, interactive platforms facilitates peer evaluation and discussion forums, promoting the core

tenets of social constructivism. Simultaneously, these tools enhance self-constructivism by leveraging personalised learning analytics to foster independent knowledge construction (Anderson, 2008).

Moreover, collaborative tools such as Padlet and Kahoot! simulate authentic communicative contexts, which are vital for language learners. Interaction with peers enables learners to share knowledge, refine their language skills, and iteratively adjust their understanding—a dynamic process embodying constructivist principles (Jonassen, 1999).

Constructivist learning theory has been widely used in the study of foreign language learning, especially the development of oral English ability. This theory emphasises that learners actively construct knowledge through interaction with the environment and others, rather than passively receive information. In oral English teaching, interactivity, situational learning, and cooperative learning are the key aspects of constructivism.

#### *Interactivity*

Online tools and digital environments provide rich interactive platforms that enhance the naturalness and fluency of spoken English. Platforms such as Padlet enables learners to share and discuss content collaboratively, thereby fostering meaningful interaction. Similarly, Kahoot! uses gamified quizzes and instant feedback to boost learners' motivation and improve their linguistic proficiency. These tools support authentic practice in spoken English, enhancing linguistic competence through collaboration and feedback.

#### *Situated learning*

Situated Learning Theory (SLT), proposed by Lave (1991), emphasises that knowledge is acquired through authentic environments and social interactions within communities. A key concept of SLT is Legitimate Peripheral Participation (LPP), which describes how novices begin with simple tasks, gradually advancing through community engagement and practice to become core members. The theory asserts that learning is influenced by social contexts and cultural backgrounds.

In recent years, SLT has been widely applied in educational and technology-enhanced learning environments. Yan et al. (2024) find situated learning in VR-based language environments significantly improved oral fluency, engagement, and self-confidence. Li et al. (2022) argue that when Mobile-Assisted Language Learning

(MALL) is designed in alignment with SLT, it could offer a more contextualised and flexible learning experience, enhancing learners' engagement and language proficiency. This suggests that SLT is equally applicable in mobile learning environments, providing language learners with more interactive and immersive learning resources.

### *Collaborative learning*

Collaborative Learning (CL) is an instructional approach that emphasises learner interaction and cooperation, grounded in social constructivist theory, particularly Vygotsky's (1978) sociocultural theory. Vygotsky proposes that social interaction plays a crucial role in cognitive development, allowing learners to operate within their ZPD and achieve higher cognitive levels through collaboration. Johnson and Johnson (1989) further expand on CL theory, identifying five essential elements for successful collaboration: positive interdependence, face-to-face interaction, individual accountability, social skills, and group processing. Their research demonstrated that integrating these elements effectively enhances both academic performance and social competencies.

However, the implementation of collaborative learning also presents challenges. Kirschner et al. (2010) argue that unguided collaborative learning may impose excessive cognitive load, particularly on novice learners, making it essential to design well-structured CL activities that account for learners' individual characteristics and provide appropriate support. In recent years, research has increasingly focused on digital and online collaborative learning environments. For instance, Stahl et al. (2006) explore computer-supported collaborative learning (CSCL), highlighting the role of technology in facilitating learner interaction and knowledge co-construction. Furthermore, Han and Ellis (2021) investigate collaborative learning in blended courses, concluding that students' learning orientations significantly influence their approach to collaborative learning.

### *Overemphasis on technology, neglecting social interaction*

A significant limitation in existing research is the overemphasis on evaluating the technological capabilities of web-based learning tools, such as the precision of automated scoring systems or the convenience of online assessment platforms, while underappreciating the role of social interaction in language acquisition. Social constructivism highlights that language learning not only relies on individual cognitive

processes but also thrives through interaction, imitation, and collaboration with others (Vygotsky, 1978).

However, many tools fail to incorporate meaningful interactive designs that facilitate peer-to-peer or learner-instructor communication. Van der Kleij et al. (2019) critique this approach, arguing that the majority of studies on digital tools focus on their technical validity rather than their capacity to promote collaboration among learners, which is fundamental for effective language acquisition.

#### *Insufficient evidence for advancing higher-order thinking*

While web-based tools have been shown to enhance learners' language fluency and accuracy, their ability to foster higher-order thinking skills, such as critical analysis and creative expression, remains underexplored. Azevedo and Gašević (2019) emphasise that much of the existing research on technology-enhanced learning primarily centres on knowledge acquisition and retention, with limited attention to how these tools develop complex cognitive abilities over time. The lack of empirical studies addressing these aspects restricts the broader educational value of such technologies in fostering holistic skill development.

#### *Limited longitudinal research on sustained impact*

Another notable gap in the literature is the scarcity of longitudinal studies examining the lasting effects of web-based learning tools. For instance, it remains unclear how these tools influence learners' sustained language behaviours or contribute to long-term improvements in autonomy and self-regulation. Dede (2020) notes that research in this area often focuses on short-term experimental outcomes due to constraints in design and funding, leaving critical questions about the enduring impact of these tools unanswered.

### **3.4.2 Theories of Assessment for Learning**

The theory of learning assessment originated in early educational practices. Initially grounded in behaviourism, it focused on psychometric principles and standardised testing (Thorndike, 1922) to quantify students' knowledge and skills, fulfilling accountability requirements and improving teaching. Scriven (1967) first introduces the concepts of formative and summative evaluation, distinguishing between assessments to improve the learning process and those designed to measure learning outcomes. Subsequently, Bloom

et al. (1971) extends the scope of formative evaluation by emphasising its critical role in enhancing learning outcomes and meeting individual needs.

As educational theories, especially cognitive learning theories and constructivism, grows, evaluation theory goes from being just a way to measure progress to an important way to help learners learn. Black and Wiliam (1998) highlights the crucial role of formative evaluation in fostering self-regulated learning, metacognitive development, and student achievement. Simultaneously, the focus of evaluation shifts from rote memorisation and knowledge quantification to addressing the complexity of learning processes, as well as critical thinking and the application of knowledge (Patel et al., 2008; Illeris, 2018b).

In recent years, technological advancements have driven significant developments in evaluation theory. For example, automated assessments and web-based formative tools (Bennett, 2011) have considerably expanded the functions of evaluation by providing immediate feedback, supporting personalised learning, and promoting learner autonomy (Cheung et al., 2021; Hwang, 2014). Additionally, diverse methods such as peer evaluation and self-assessment have gained prominence in contemporary education. These approaches increase learner engagement with educational content and encourage students to take responsibility for their learning progress (Soares, 2020; Lu, 2016).

At the same time, ongoing discussions on educational outcomes have drawn attention to the socio-cultural dimensions of learning. Transformational learning theory, which emphasises critical reflection and social context, has become increasingly important, particularly in adult education and professional development (Phuong et al., 2019). This perspective encourages educators to account for not only the cognitive aspects of learning but also the emotional and social factors influencing learners' experiences and outcomes (Alam, 2022; Phuong et al., 2019).

Learning assessment theory has transitioned from its early behaviourist roots to a more holistic framework that incorporates cognitive, social, and emotional dimensions. Today's function extends beyond evaluating educational effectiveness—it supports the optimisation of teaching strategies, motivates learners, and enhances their engagement (Sadler, 1989). As educational contexts continue to evolve, evaluation theory is likely to develop further, emphasising innovation, flexibility, and learner-centred approaches to

meet diverse learning needs and achieve broader educational goals (Patel et al., 2008; Hwang, 2014).

### ***Types of assessment***

Assessment constitutes an indispensable element of the educational process, playing a critical role across various stages of teaching and learning. Based on their distinct purposes and application methods, assessments typically fall into four categories: summative, formative, diagnostic, and dynamic. Each category serves unique functions and objectives while also exhibiting certain limitations.

#### ***Summative assessment***

Summative assessment, often called terminal assessment, is conducted at the conclusion of a learning period and primarily aims to evaluate students' knowledge and skills acquired during the instructional process. Typical examples include final examinations, standardised tests, coursework essays, and project reports, all of which are employed to determine whether students have achieved the intended learning outcomes (Scriven, 1967). This assessment focuses on measuring learning outcomes rather than the learning process, providing students with an overall evaluation of their performance while furnishing educators and administrators with critical data regarding instructional effectiveness.

One of the key strengths of summative assessment lies in its provision of clear, standardised benchmarks for evaluating student achievement, which enhances fairness and comparability. Objective data from summative assessments help many people, such as students, teachers, and education administrators, by showing how well students are learning and helping them make choices about the quality of the curriculum and how well the school is doing overall (Sadler, 1989; Harris et al., 2018). For instance, standardised tests and final exams yield comprehensive data that enables schools and educational authorities to adjust curricula and refine teaching strategies accordingly (Wiliam, 2011).

Despite its significant role in education, summative assessment has attracted criticism for its limited ability to foster learning and improve instructional practices. Scriven (1967) argues that summative assessment excessively prioritises 'judging' learning outcomes over 'facilitating' the learning process. Black and Wiliam (1998) also argue that focusing only on final grades doesn't consider how learning changes over time and doesn't

offer immediate feedback and help, which are especially important for meeting the needs of each student during long learning periods. Standardised tests, in particular, tend to disregard individual differences among learners (Carless, 2007) and may disadvantage certain student groups, especially those with diverse cultural backgrounds, cognitive styles, or learning habits whose performance might not accurately reflect their true learning potential (Brookhart, 2017).

### *Formative assessment*

The concept of formative assessment was initially introduced by Scriven (1967), who differentiated between summative and formative assessment, defining the latter as an ongoing process of evaluation aimed at providing feedback to help students refine their learning and adjust strategies for improvement. Scriven's definition underscores the centrality of feedback and its alignment with the learning process, laying the groundwork for subsequent research and practice in educational reform. Black and Wiliam (1998) further advance formative assessment theory, highlighting its pivotal role in enhancing learning outcomes and fostering autonomous learning, thereby promoting its widespread application in education.

The primary advantage of formative assessment lies in its capacity to deliver continuous, constructive feedback that fosters active student engagement and self-regulation (Sadler, 1989). Such feedback not only enables students to identify areas of mastery and improvement but also facilitates a deeper understanding of content while cultivating metacognitive skills (Hattie & Timperley, 2007). By tracking students' learning progress, formative assessment empowers teachers to identify difficulties and implement targeted interventions promptly. Immediate feedback, such as verbal comments or written evaluations, allows students to identify their strengths and areas for growth, leading to measurable improvements in learning outcomes. Furthermore, formative assessment enhances students' confidence and motivation; specific and instant feedback bolsters self-efficacy and stimulates interest in learning, fostering a sense of ownership over their academic progress (Carless, 2007). For educators, formative assessment offers flexibility to adapt teaching strategies in response to student needs, fostering a dynamic and personalised learning environment that promotes holistic development (Wiliam, 2011).

Nevertheless, formative assessment is not without limitations. Its effectiveness relies heavily on providing timely and high-quality feedback, which demands substantial time and effort from teachers—resources often constrained in large class settings or under-resourced environments (Hattie & Timperley, 2007; Black & Wiliam, 1998). Additionally, formative assessment outcomes can vary depending on students' self-regulation capabilities. For learners with weaker autonomous learning skills or limited confidence, particularly younger students, responses to feedback may be less effective, resulting in inconsistent outcomes across individuals (Wiliam, 2011).

With advancements in IT and AIT, technology-supported formative assessment tools offer solutions to some of these challenges. Online platforms and automated feedback systems can deliver timely, personalised guidance while reducing the burden on teachers and enhancing the efficiency and accuracy of feedback. Such tools dynamically adjust content based on student's progress and needs, thereby facilitating autonomous learning and improving learning outcomes (Bennett, 2011; Cheung et al., 2021).

#### *Diagnostic assessment*

Diagnostic assessment refers to evaluations conducted at the beginning of instruction or during the early stages of learning, with the purpose of identifying students' existing knowledge, skill levels, and potential learning challenges (Wiliam, 2011). Unlike summative and formative assessments, diagnostic assessment prioritises providing teachers with detailed information about students' starting points, enabling them to tailor instructional strategies to individual or group needs (Brookhart, 2017). For example, diagnostic tests often uncover knowledge gaps, misconceptions, or specific difficulties, thus informing subsequent teaching interventions (Black & Wiliam, 1998).

The primary strength of diagnostic assessment lies in its specificity and predictive nature, allowing educators to pinpoint students' strengths and weaknesses early on and design targeted instructional plans (Carless, 2007). By addressing diverse learning needs, diagnostic assessment enhances instructional efficiency and reduces the risk of misallocating teaching resources. However, critics argue that reliance on static assessment tools fosters a narrow view of student potential. Vygotsky's (1978) Zone of Proximal Development Theory emphasises that assessing only current abilities without considering learners' developmental potential may fail to capture their capacity for future learning.

### *Dynamic assessment*

Dynamic assessment, rooted in Vygotsky's (1978) 'ZPD' theory, emphasises evaluating learners' potential development by providing external support or intervention during the assessment process. In contrast to static assessments, dynamic assessments focus on both process and interaction, as teachers offer guidance or prompts to assess students' ability to learn and adapt (Lantolf & Poehner, 2004). Two core elements underpin dynamic assessment: diagnosis and intervention. Through carefully calibrated assistance, evaluators observe how students respond and progressively overcome challenges, revealing their learning potential and cognitive development (Poehner, 2008).

The principal advantage of dynamic assessment lies in its developmental and process-orientated nature. By integrating assessment with intervention, dynamic assessment provides a comprehensive understanding of students' abilities and potential for growth (Lidz & Gindis, 2003). For instance, in language learning, dynamic assessment can identify specific difficulties and guide learners to overcome them through gradual support, thus enabling personalised instruction (Shrestha & Coffin, 2012). Furthermore, this approach transforms the assessment process into an instructional activity, facilitating students' problem-solving abilities while revealing their capacity for independent learning (Poehner & Lantolf, 2010).

However, dynamic assessment poses challenges for implementation, as it demands considerable expertise from teachers, who must accurately diagnose students' developmental levels and provide appropriate, effective guidance (Lidz & Gindis, 2003). Additionally, the interactive and individualised nature of dynamic assessment often necessitates significant time investment, making it less feasible in large class settings or resource-limited environments (Haywood & Lidz, 2007). Moreover, the lack of standardised procedures and quantifiable results limits the comparability and scalability of dynamic assessment outcomes (Shrestha & Coffin, 2012).

The debate about assessment has never ceased, as each assessment method possesses its unique advantages while inevitably presenting certain limitations. In educational practice, selecting an appropriate assessment method based on the specific teaching context and objectives is essential. The purpose of assessment extends beyond merely measuring learning outcomes. Assessment should effectively support and promote

learners' growth and development. The selection and implementation of assessment methods should demonstrate clear strategic intent, goal alignment, and learner-centred focus to maximise their positive impact on teaching and learning. So, a successful assessment method should accurately measure learning outcomes and give learners specific feedback to help them figure out where they need to improve and meet their own needs. This feedback can then be a powerful tool to help them keep progressing.

### **3.4.3 Technology Acceptance and Application Theory**

With the rapid development of technology and the ongoing advancement of digital transformation in the field of education, the integration of technology-supported formative assessment into teaching and learning has garnered significant attention. This approach is characterised by the use of technological tools, including online learning platforms, intelligent assessment systems, and learning analytics software, to accurately capture a comprehensive picture of learners' progress. Through real-time data recording, automated scoring, and instant feedback, this method facilitates precise diagnosis of learners' performance and provides personalised recommendations to support adjustments in teaching and learning. Technology-driven evaluation approach offers a fresh look on educational reform and serves as an indispensable and highly convenient option for fostering independent and lifelong learning among students (Scalise and Gifford, 2006; Chen et al., 2023).

Chen et al. (2023), in their analysis of the application of technology-based assessments (TBAs) across the education systems of 34 countries, argue that TBAs enhance measurement precision through adaptive testing (CAT) and artificial intelligence technologies and also improve learning engagement and interaction through instant feedback, interactive design, and personalised recommendations. Moreover, TBAs provide valuable teaching support, enhance managerial efficiency, and prompt the realisation of educational equity.

#### ***Technology acceptance model (TAM): theoretical developments and research applications***

The Technology Acceptance Model (TAM) was introduced by Davis (1989) to explain and predict users' acceptance of technology. Grounded in the Theory of Reasoned Action (TRA), TAM posits that perceived usefulness (PU) and perceived ease of use

(PEOU) are the two primary determinants of an individual's acceptance of technology. Perceived usefulness refers to the degree to which users believe that using a particular technology will enhance their job performance, while perceived ease of use relates to the effort required to use the technology.

Over time, TAM has undergone multiple extensions and revisions. Venkatesh and Davis (2000) come up with TAM2 as they want to make the model better at explaining things by adding external variables like subjective norms and cognitive instrumental processes like job relevance, output quality, and result demonstrability. Subsequently, Venkatesh et al. (2003) develop the Unified Theory of Acceptance and Use of Technology (UTAUT), which integrates multiple theoretical models to provide a more comprehensive framework. UTAUT introduces four core constructs—performance expectancy, effort expectancy, social influence, and facilitating conditions—alongside moderating variables such as gender, age, experience, and voluntariness, to better explain technology adoption behaviour. However, these expanded models have increased complexity, potentially reducing their practicality and applicability. Furthermore, the cross-cultural and cross-context validity of these models remains a subject of debate, necessitating further empirical validation.

In empirical research, TAM has demonstrated broad applicability and has been widely used across various domains, including e-learning, mobile applications, healthcare information systems, e-commerce, social media, and smart technologies. A meta-analysis conducted by King and He (2006) confirm that TAM effectively predicts users' technology adoption across multiple research fields. Nevertheless, their studies also highlight TAM's limitations in explaining user behaviour, particularly when external factors such as organisational culture, policy support, and social dynamics are considered, leading to a decline in its predictive power.

In the domain of e-learning, TAM has been used to assess students' and educators' acceptance of online learning platforms. For example, Cheng (2011) finds perceived usefulness was the most significant factor influencing students' adoption of online learning systems, whereas perceived ease of use indirectly influenced adoption through its effect on perceived usefulness. This finding suggests that, in educational technology applications,

enhancing users' perception of a system's usefulness is more effective in promoting adoption than merely optimising its user interface.

In healthcare information systems, TAM has been widely applied to predict healthcare professionals' acceptance of electronic health records (EHRs) and telemedicine technologies. Holden and Karsh (2010) find perceived ease of use and perceived usefulness significantly influence medical professionals' adoption of EHRs. However, they also note that social influence and organisational support play a more critical role in the healthcare sector, extending beyond TAM's traditional variables. This study indicates that, in highly specialised and institutionally dependent fields, TAM's applicability is somewhat limited and necessitates integration with additional variables to enhance explanatory power.

In mobile technology and social media, TAM has been used to analyse users' adoption intentions towards smartphone applications and social networking platforms. For instance, Zhou (2011) finds in the context of mobile payment systems, trust and perceived risk can moderate the relationship between perceived usefulness and behavioural intention. This suggests that relying solely on PU and PEOU to explain user adoption may be inadequate, particularly in technology applications involving personal privacy and security risks.

Despite its prominence in technology acceptance research, TAM has notable limitations. Firstly, the model is overly simplistic, focusing exclusively on perceived usefulness and perceived ease of use while neglecting social, psychological, and environmental factors (Legris, Ingham & Collette, 2003). Secondly, TAM assumes that users are rational decision-makers, whereas in reality, their behaviour is often influenced by emotions, habits, and social interactions (Bagozzi, 2007). Additionally, TAM's static nature makes it less adaptable to rapidly evolving technological environments, limiting its effectiveness in explaining the adoption of emerging technologies such as artificial intelligence, blockchain, and augmented reality (Dwivedi et al., 2019).

Future research should focus on expanding TAM by incorporating a broader range of variables, integrating theories from social psychology and behavioural economics to enhance its explanatory power. Additionally, a combination of quantitative and qualitative research methods should be employed to gain deeper insights into technology adoption

behaviour. Further interdisciplinary and cross-cultural studies are also needed to test the model's applicability across different industries and socio-cultural contexts. Finally, artificial intelligence and big data analytics should be leveraged to develop dynamic predictive models capable of adapting to rapidly changing technological landscapes.

### **3.5 Chapter Summary**

The researcher has gained a clear understanding of the relevant theories and research practices by conducting a comprehensive review and in-depth analysis of literature on WBFA, which lays a solid theoretical and practical foundation for subsequent research.

While formative assessment has been extensively studied, research on WBFAT has emerged relatively late (in the early 21st century) but has gained significant attention in recent years. Based on the analysis of the selected literature, research on WBFA learning exhibits multidimensional characteristics. Geographically, it is primarily concentrated in the United States and China. In terms of participants, study duration, research contexts, disciplinary distribution, and research methods, a considerable degree of diversity is evident. Theoretical foundations span multiple domains, including learning theories, technology acceptance and application theories, and assessment theories, providing a rich theoretical perspective for research. Additionally, the literature reflects a dual perspective from both teachers and students regarding web- or mobile-based learning and formative assessment tools. While these tools are recognised for their positive impact on learning outcomes, enhanced interaction, and instant feedback, their limitations, technical challenges, and other potential issues are also acknowledged. This underscores the necessity for tool (or technology) providers and educational institutions to take full account of students' needs and experiences as the primary stakeholders in the design, implementation, and integration of such tools. It is crucial to ensure that these technologies align with students' actual learning needs, technological proficiency, and cognitive load. Furthermore, accessibility, usability, and adaptability should be central considerations in the development process to avoid barriers that could hinder effective learning. Educational institutions should also establish systematic mechanisms to continuously evaluate and refine these tools based on user feedback. This would ensure that technology-enhanced formative assessment not only improves learning efficiency but also minimises technical

difficulties and usability concerns. Only by adopting a student-centred approach can the deep integration of technology and education be achieved, thereby maximising its role in facilitating learning and assessment.

Through a review of relevant theories, the researcher has developed a profound understanding of learning theories such as metacognitive theory, self-regulated learning theory, motivation theory, and constructivist learning theory, as well as assessment theories and technology acceptance and application theories. WBFA requires support from both learning and assessment theories to ensure its scientific and pedagogical effectiveness, as well as guidance from technology acceptance and application theories to enhance its feasibility and user acceptance. Learning and assessment theories provide the fundamental principles and methodologies for formative assessment, enabling it to foster cognitive development, self-regulation, and learning motivation. Meanwhile, technology acceptance and application theories highlight user experience, technological usability, and the influence of these factors on learning behaviour, offering critical insights into the design and adoption of assessment tools.

Future research and practice should prioritise the integration of these two theoretical frameworks to enhance and innovate WBFAT. On the one hand, researchers should further explore the application of learning theories to determine how technology can enhance students' meta-cognitive abilities, autonomous learning skills, and learning motivation. This would enable formative assessment to serve not only as an evaluative measure but also as a means to actively support learning processes. On the other hand, we need to pay attention to technology acceptance and application theories by looking at the main things that make students and teachers use these tools, like how easy they are to use, how useful they are seen to be, and how well they can be adapted to different situations. Addressing these aspects would help reduce barriers to adoption and enhance the educational value of technology.

## Chapter 4 Methodology

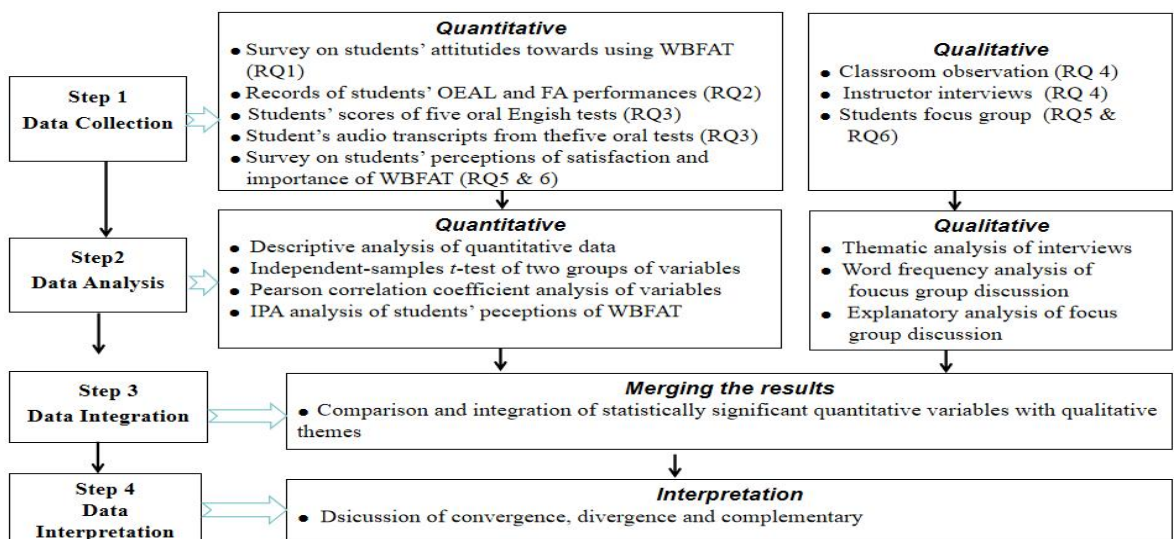
This chapter first provides a detailed introduction to the mixed-methods design, explaining the choice of an explanatory sequential approach and the rationale for using mixed methods in this research. Then the researcher describes the settings for the questionnaires, interviews, focus group discussions, and the autonomous learning and formative assessment (FA) of oral English in the experimental study. Following this, the functions of FiF are depicted, and the research procedure is elaborated in detail to offer a panorama of the research process. Additionally, this chapter makes expressions of the roles of the researcher, instructor, and technician and the measures of ensuring the research findings are valid. At last, the researcher explains the data analysis methods and processes in relation to each research question.

### 4.1 Design of the Mixed Methodology

#### 4.1.1 An Explanatory Sequential Design of Mixed Methods

The selection of an optimal research methodology is of paramount importance for the success of a study, which will determine the types of data, sources of data collection, and the ways of data analysis. There is no single strategy that can solve all kinds of research problems, researchers need to make decisions and use discretion to choose pathways at each distinct stage of a study (Denscombe, 2017) according to the research questions, research objectives, and samples involved. To answer the research questions, the researcher designed a routine for the mixed research methods according to the principles of explanatory sequential design as shown in Figure 3.

Figure 3 An Explanatory Sequential Design of Mixed Methods in this Study



An explanatory sequential design typically consists of three stages. First, quantitative data strands are collected, analysed, and interpreted according to research questions and objectives. Secondly, the related qualitative data strands are designed to further interpret the quantitative results of the previous stage. Finally, quantitative and qualitative data are integrated to draw comprehensive findings (Creswell & Plano Clark, 2018; Schoonenboom & Johnson, 2017; Shorten & Smith, 2017; Dawadi et al., 2021). This approach enables the researcher to provide a comprehensive explanation of the quantitative findings through the insights gained from the qualitative data (Wisdom & Creswell, 2013) as it can explain some unexpected quantitative results in more detail. The research process of all themes and the crucial research questions in this study reflect an explanatory sequential design of mixed methods.

During the data collection, the researcher adopted the means of survey, test, and experiment to obtain quantitative data, and the ways of interview, observation, and focus group as well to gain qualitative data according to the attributes of the research questions in this study. The quantitative data strand in this study includes data from a survey on students' attitudes towards using WBFAT, records of students' performance of their participation during the process of OEAL & FA, scores from five oral English tests, the analysis of the linguistic indicators of students' audio transcripts from the five oral English tests, and a survey on students' perceptions of satisfaction and importance of the WBFAT. While the qualitative data strands include records of students' spoken English use in the classroom, interviews with the instructor, and a student focus group discussion.

In the process of data analysis and integration, this study explored the impact of WBFAT on student learning in a comprehensive and in-depth manner, overcoming the limitations of a single approach and enhancing the validity of the study and the reliability of the conclusions by integrating qualitative and quantitative data types and applying multiple analytical methods to synthesise findings from different sources and types of research.

In the data analysis process, quantitative analysis methods such as descriptive statistics, correlation analysis, and t-tests were used for quantitative data to measure and compare students' academic performance under different conditions. The combination of

various statistical methods helped to quantify the impact of WBFAT on the experimental and control groups before and after its implementation.

Descriptive analysis was the first step in any complex data analysis, and the accuracy and validity of subsequent analysis can be markedly enhanced through preliminary data exploration. The ignorance of descriptive analysis may lead to serious bias and misjudgment in data analysis (Behrens, 1997). Meaningful descriptive analysis could help reveal socially important patterns (Leob et al., 2017). In this study, descriptive statistical methods, including the calculation of variables' frequency, percentage, mean, median, mode, kurtosis, standard deviation to visually show the overall characteristics of the data distribution and tendency and enable the researcher to come to more reasonable conclusions about the stability of the parameters and their prediction (Petruk et al., 2023; Washington et al., 2020; Satyawati et al., 2021). Şirin et al. (2018) and Hatem et al. (2022) point out that the symmetry, concentration, and normality of the data distribution would be further understood and provided more detailed information for in-depth understanding by analysing the skewness and kurtosis of the data.

In addition, other statistical analysis methods for quantitative data were also employed in this study, including Pearson correlation coefficient analysis, variance analysis, and IPA analysis. For instance, to explore differences in the progress of oral English proficiency in holistic and specific language features between students between two groups, the research adopted the Pearson correlation coefficient formula to calculate the coefficient correlation between the two groups' data based on the scores of repeated tests of the overall competence of oral English and the scores of specific language features via the recorded audio scripts of the module of picture speaking questions in the oral test. The significance of the correlation coefficient was tested by calculating the *p* value, and the mean difference between the two groups was compared by the analysis of variance (ANOVA) scores of students' oral tests to evaluate the influence of WBFAT on the development of EFL learners' oral English proficiency in full scales. When exploring students' perceptions on the satisfaction and importance of WBFAT, the researcher calculated the average value of satisfaction and the importance of WBFAT and drew the IPA matrix to identify which attributes should be maintained and which attributes should be improved for the related stakeholders in the future, including WBFAT manufacturers,

learners, instructors, and educational institutes. The IPA analysis has been proved to be an important means to evaluate learning and teaching (Cladera 2021), improve performance and decision-making (Nazari-Shirkouhi et al., 2020), and promote educational service quality (O'Neill & Palmer, 2004).

The researcher compared the findings of quantitative data analysis to the outcomes of qualitative data analysis and integrated them together to provide a more comprehensive explanation for the research questions (Harding et al., 2022), as quantitative data provides numerical evidence for statistical analysis while qualitative data gives the potentiality for detailed and in-depth exploration (Anwar 2015). In the process of data reporting and interpretation, the research effectively conveyed and deepened its findings through clear data presentation and multidimensional interpretation of the quantitative data, such as tables, charts, and textual descriptions, offering an intuitive understanding of the relationships and trends among variables. Simultaneously, qualitative data were coded and thematically analysed, providing contextual and situational support to enrich the quantitative findings. In the interpretation of the findings, convergence, divergence, and complementarity of qualitative and quantitative data were discussed progressively to reveal the existing explanations or possible alternative explanations of the research questions from multiple perspectives, providing a multidimensional, comprehensive, and rich understanding of the research problem in this study.

#### **4.1.2 Why Mixed Methods**

A mixed-methods design was convinced to be beneficial, when properly designed and executed, to exploring complex research issues (Dawadi et al., 2021) and improving the reliability and validity of findings (Creswell & Plano Clark, 2011; Greene et al., 1989) as it integrates philosophical frameworks of both post-positivism and interpretivism (Fetters, 2016) and provides meaningful explanations to research questions via the combination of quantitative and qualitative analysis. Furthermore, it allows researchers to address research questions in a comprehensive manner with sufficient depth and breadth by integrating qualitative and quantitative data (Enosh, Tzafrir, & Stolovy, 2014).

Quantitative methods provide research with extensive and precise evidence as they allow for the data collection from a large range of participants in terms of numbers, genders, occupations, and so on, and quantitative data can be accurately calculated and

analysed statistically by software to derive generalisable characteristics or patterns in the research questions. Thus, the findings from quantitative data can be widely applied and popularised. Qualitative methods, on the other hand, help researchers to obtain data from the participants through first-hand experience, truthful reporting, and conversation quotations, such as their attitudes, experiences, and feelings, including textual, oral, and audio data, which cannot be obtained by quantitative methods. Qualitative data provides comprehensive context and details as well as in-depth understanding and interpretation (Hennink et al., 2020; Alhazm & Kaufmann, 2022; Alam, 2021; Akinyode & Khan, 2018; Dey, 2003) of the phenomena in study, thus answering the research questions in a comprehensive way. In a word, quantitative data provides the breadth of the research while qualitative data provides the depth.

In the practice of research work, no research method could solve all research problems, whether quantitative or qualitative research methods, because each has advantages and limitations (Taherdoost, 2022; Rahman, 2020; Malina et al., 2011; Bryman, 1984). What we can do is to offset the weakness of either approach as much as possible (Heyvaert et al., 2016). Mixing the quantitative and qualitative methods may prove more effective than the use of a single method, as it might yield rich insights and in-depth understandings of the research questions that cannot be fully elucidated by employing either (Creswell, 2009, 2017). The integration of mixed methods is helpful to solve complex problems in research (Poth & Munce, 2020) as it facilitates the integration of complementary multiple data sources and the understanding of the phenomena under study from different perspectives and viewpoints (Shorten & Smith, 2017). Increasingly, methodologists have approved the integration of qualitative and quantitative data as the key to mixed methods (O’Cathain et al., 2007). Integration is viewed as an intentional process by which the researcher brings quantitative and qualitative approaches together in a study (Creswell, 2021), in which the two forms of data then become interdependent in addressing common research questions and hypotheses (Bazeley, 2012). Effective and meaningful integration of mixed methods does yield more persuasive efficiency than the individual qualitative and quantitative methods (Fetters & Freshwater, 2015).

It is obviously not sufficient to adopt only one or unilateral methods to answer the research questions and reach the aims of this study. On the one hand, quantitative methods

were used for exploring learners' general attitudes, the frequency of learners' use of WBFAT, and the significant changes in scores of their oral English tests, because they could be collected through Likert scale questions and repeated tests and be accurately calculated by statistical software. On the other hand, qualitative methods were conducted to obtain the first-hand information and more detailed and in-depth evidence through class observation, teacher's interviews on students' behaviour in class, and a focus group discussion.

Moreover, another important reason for the mixed-methods design in this study is the complex interplay between technology, pedagogy, and learner psychology in the context of online language learning and teaching. On one hand, the wide application of online technology in language learning and teaching provides more opportunities, personalised learning, and instant feedback. On the one hand, technologies such as virtual reality, artificial intelligence, and adaptive learning have been proved to significantly improve learning results (Golonka et al., 2014; Shadiev & Yang, 2020; Budiman, 2020; Su & Zou, 2022; Hazaymeh, 2021; Chapelle, 2006). On the other hand, online technology also brings challenges and greater uncertainty to education (Selwyn, 2016). For example, in the context of online education, due to a lack of technical support, learners are experiencing trauma, psychological pressure, and anxiety to various degrees (Bozkurt, 2020). Relying solely on technology cannot ensure the improvement of learning effect. Only by integrating technology with teaching, learning, and assessment, technology can promote students' active participation and collaboration, thus improving learning results (Chick et al., 2020; Martins & Gresse Von Wangenheim, 2023; Lawless & Pellegrino, 2007; Hew & Brush, 2007). Secondly, learners' psychological factors, such as motivation, self-efficacy, and learning anxiety (Dong, 2022; Barkanyi, 2021) also have an impact on the success of language learning. Therefore, to explore the impact of online or web-based technologies on language learning. It is necessary to comprehensively capture the complex interplay between technology, pedagogy, and learner psychology through a mixed approach, combining quantitative and qualitative data. Quantitative research can reveal the influence of technology and pedagogy on learning outcomes through questionnaires and repeated tests. Qualitative research, on the other hand, can deeply understand learners' psychological experience and feelings through interviews and observations.

## **4.2 Settings**

This study was conducted on the campus of Jiangxi Teachers College (JTC), a local public higher vocational college in mainland China. During the course of the study, the researcher chose different physical environments according to research activities in different stages.

#### **4.2.1 Settings of the Survey**

Two formal surveys by questionnaires in this study were conducted in the present study. Both were carried out in a public classroom of the college, which could accommodate 200 people. Participants came to the classroom within the suggested time range (19:00-20:40 pm, on 28 June and 25 Apr, 2023, respectively) to answer the questionnaires on the spot. The classroom was equipped with a computer and two projectors, a microphone, and monitors in each corner, through which the researcher could clearly show participants the purpose of the study and directions for answering the questionnaires. The scenes were recorded by the monitors. In order to improve the quality of data, each participant was suggested to answer the questionnaire for no less than 30 minutes so that they had enough consideration of each question.

The reasons why the researcher chose to collect questionnaire data in the classroom are, on one hand, active communications and interactions could be easily realised between the researcher and the participants, and on the other hand, participants could answer the questionnaires within the suggested length of time, which helps to improve the quality of data, although more time was consumed.

Researchers now tend to adopt tools like Sojump (wjx.cn), Qualtrics, and Mturk to collect data online, as they are viewed to have the advantages of quickly covering geographically diverse populations, getting faster answers, and encouraging engagement on sensitive topics (Braun et al., 2020; Braun & Clarke, 2013) because of their high anonymity (Terry & Braun, 2017). However, they are not so attractive (Lefever et al., 2007) because online survey tools also have many disadvantages (Ha, 2022), such as participants may not fully understand the items in the absence of face-to-face interaction, or they may suffer from survey fatigue (De, 2021; Fass-Holme, 2022).

#### **4.2.2 Settings of Interviews**

To enhance the interviewer's supportiveness (Lavoie et al., 2021; Blasbalg, 2019) and ensure consistency, interviews in this study were conducted in neutral and trusting

environments. The interviews with the instructor involved in this study were all conducted in the researcher's studio on the campus of JTC. The studio is located in the college's public administration office building which is open to faculty and students at all times. The studio is a separate room of about 12 square meters, with a door, which can reduce the external interference of the interviewees. It is equipped with a computer, a desk, three orange single sofas, and a large window with plenty of sunlight. These provide a quiet, trusting, and comfortable environment for the interview. The time of interviews was agreed upon by both the researcher and the interviewee in advance during the working hours when the instructor did not have classes. Prior to the interview, the interviewee was fully informed of the purpose of the study, the nature of involvement, and measures of confidentiality protection and anonymity. Written consent was obtained from participants before the interview. The interviews were conducted using a semi-structured guide which was valuable to gather rich and in-depth information (Lytje & Dyregrov, 2021; Nachalwe, 2021; Bogi, 2023) in the form of question-and-answer. Questions were open-ended, and participants were encouraged to answer in depth, being allowed to give up answering the questions and express themselves flexibly. Each interview lasted about 15-20 minutes, with the door of the studio ajar throughout. All interviews were recorded with the explicit consent of the interviewee. In addition, the researcher took notes to capture some non-verbal cues, such as participants' immediate action and emotion, enriching the qualitative data.

#### **4.2.3 Settings of the Focus Group**

A focus group discussion based on the significant items of the survey on students' perceptions of satisfaction and importance of FiF was conducted among the experimental group on the Tuesday afternoon of the final week of the experiment study. It was a period when students were not engaged in scheduled classes. Eighteen participants in the experimental group (four were absent due to club activities) attended the discussion. The room was equipped with a computer, a projector, loudspeakers, and an individual microphone, a conference table, and soft chairs, with a full of light and comfortable environment.

Before the discussion, all participants were informed of the purpose of the discussion and the measures of confidentiality and anonymity. They were suggested to sign

a written consent. It was conducted as a semi-structured discussion, in which the researcher presented five topics in turn for participants to discuss. The participants were permitted to present their views in forms of free expression, ask-and-answer, conversation, and debate to be encouraged to offer insights from a variety of perspectives. The focus group discussion lasted approximately 45 minutes. With the consent of all participants, the discussion was audio-recorded throughout to ensure the accuracy of data recording. Additionally, the researcher took some short field notes during the discussion. Through this focus group discussion, the researcher gained a deeper understanding of students' feelings about using the WBFAT.

#### 4.2.4 Settings of the OEAL & FA in the Phase of Experiment

The design of experimental study was based on the second module of College English, Phase 2 (CE2), that is, Oral English autonomous learning and formative assessment (OEAL & FA). The curriculum, CE2, was offered to all students of non-English majors in JTC who have completed the CE1 course (usually sophomores). The learning of CE2 consisted of two parts: classroom learning and autonomous learning. The autonomous learning focused on the practice of oral English. The learning materials, requirements, and evaluation of the two modules are shown in Table 9.

*Table 9 The Scheme of CE2 Learning in JTC*

|                     | Classroom Learning                                                                                    | Autonomous Learning                                                                                                                                                             |
|---------------------|-------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Learning materials  | An assigned textbook                                                                                  | Not specified                                                                                                                                                                   |
| Time allocation     | 2 scheduled lessons per week (80 minutes)                                                             | Recommended no less than 1 class hour (40 minutes) per week                                                                                                                     |
| Learning objectives | Improve learners' English vocabulary, reading, writing and translation skills                         | Improve learner's oral English in accuracy, fluency and completeness<br>Improve learner's autonomy of learning                                                                  |
| Assessment          | Class performance (20%) : attendance, questioning, discussion, innovation<br>Final written test (30%) | Autonomous learning performance (20%) : time spent, number of tasks completed, etc.<br>Final oral examination (30%) : a three-minute monologue and a three-minutes conversation |

#### 4.2.5 Setting of the Instrument in the phase of Experiment - FiF

FiF is a kind of oral English learning and automatic platform, developed by iFlytek Co., Ltd., supporting autonomous learning and formative assessment of oral English for learners (Peng & Xiao, 2023; Zhou, 2018). It has been regarded as one of the mainstream intelligent web-based FA products in oral English teaching and learning and widely used in colleges and universities in mainland China due to its positive role in promoting the

interest of English learners and the improvement of their oral English proficiency (Gong & Xiang, 2022; Wang, 2021; Huang, 2020). The main role of the software is described as ‘thoroughly solving the main problems in oral English teaching and learning, such as insufficient personalised teaching and learning, limited learning resources, and the lack of formative assessment and feedback’ (iFlytek, 2020).

### ***The function to support learners’ oral English autonomous learning online***

There is a huge digital resource base on the FiF platform, covering 9 major themes of various fields, such as academy, life, career, travelling, culture, science, sports, etc. (see Figure 4), English beginners or English masters can easily find learning resources suitable for various needs. In addition, FiF supports learners and instructors to build their own learning resource base where users can systematically construct materials that they are interested in. Therefore, FiF supports learners to conduct personalized autonomous learning of oral English.

Figure 4 A Screenshot of the Digital Resource Base on the Web-Platform of FiF



### ***The function of automatic scoring and feedback in formative assessment***

Through the intelligent assessing engine, FiF can automatically detect the errors of the input sound, score and provide revising reference for oral expression immediately, suitable for the assessment of word and sentence reading, dictation, conversation, man-machine interaction. Learners can receive instant scores, including total and dimensional scores of accuracy, fluency, completeness, and a detailed feedback report ( see

Figure 5). Speaker's errors in expressions are marked with different colors, with green being correct, red being wrong, yellow being non-standard, and black being missed (see Figure 6), which is beneficial for students to conduct personalized and targeted autonomous learning.

Figure 5 A Screenshot of the FiF's Functions in the Comprehensive Assessment for Oral English Test

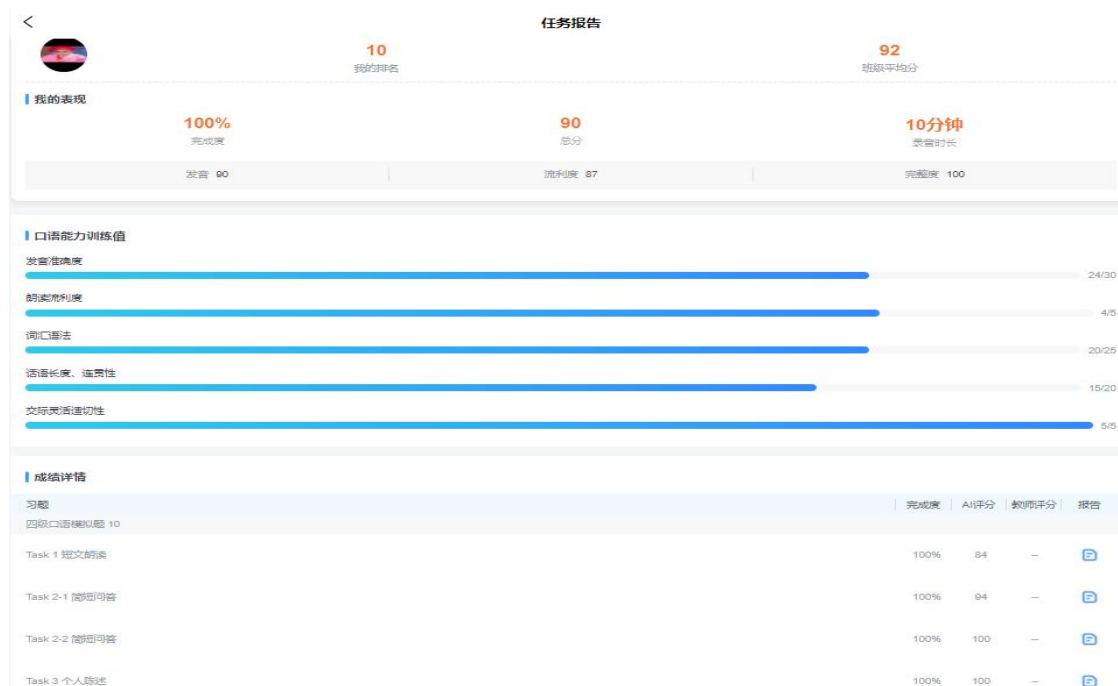
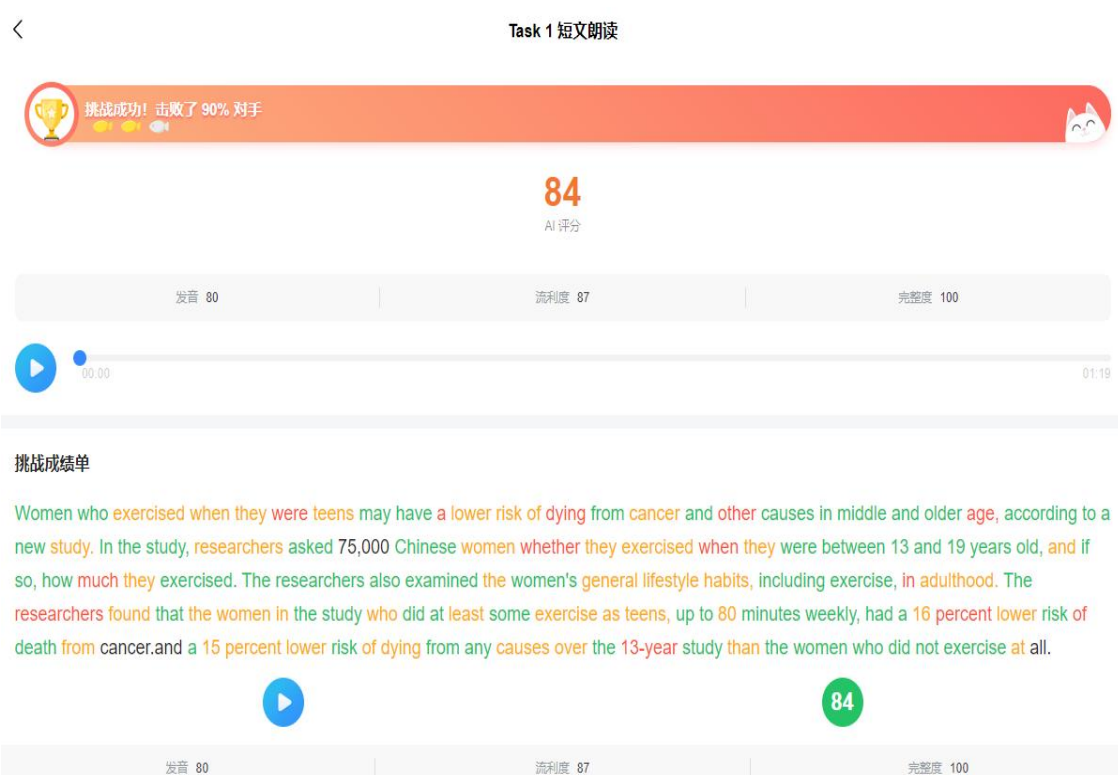


Figure 6 A Screenshot of the FiF's Functions in Rating and Feedback for Oral English Test



### *The function to calculate and present learners' learning performance*

As shown in Figure 7, the statistic and analytic data panel of users' behaviours can be used by the instructors and the researcher to analyse learners' performance and trends of oral English development. The system can collect and calculate the data of learners' overall use on the platform and build a multi-level dynamic assessment index system, offering instructors the panorama of students' autonomous learning after class. It provides a valuable reference for instructors to reform teaching contents and adjust teaching strategies and a scientific basis for school management departments to make teaching plans and decisions as well. The administrator and instructors can view dynamic data on students' autonomous learning, testing and scoring, teacher-student interactions, peer assessment, student participation in contests, and trends of student learning behaviour and oral English proficiency development. For students, they can check their own usage data in each module of the system in detail. The platform supports users to view online, and offers access to downloading the records and assessment documents for local storage. Through the analysis of panel's data, the researcher can grasp the comprehensive situations of users' login, using time length, numbers of task completion, peer evaluation, learners' behavior trends in the pilot study and experimental study, so as to conduct correlation analysis with the development of students' oral English.

*Figure 7 A Screenshot of FiF's Data Panels*



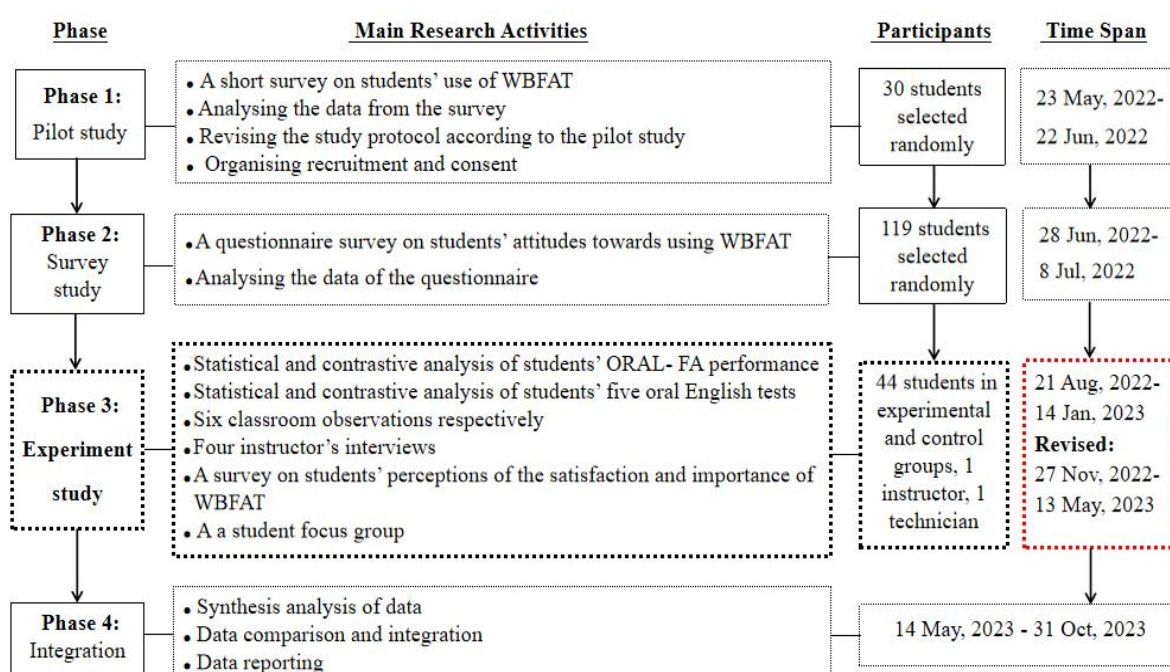
### *The function to support interaction and collaboration*

FiF supports learners to interact with teachers as well as with classmates and peers in different classes or even different colleges, including rating each other, online chatting, and leaving messages on the board.

### **4.3 Research Procedures**

As shown in Figure 8, the procedures of this study include four stages: they are the pilot study, survey study, experimental study, and post-experiment study. The pilot study was conducted from late May 2022, lasting for 1 month, to test the feasibility of the study protocol in a small-scale study, as well as to complete the recruitment of student participants and obtain their written consent. And then, the researcher carried out a survey on students' attitudes towards using WBFAT. Then, according to the original scheme of the study, the experimental research should be followed. However, with the arrival of the summer vacation and several months of lockdown caused by the spread of Covid-19, the experimental research was suspended for nearly four months until the end of November 2022, after students returned back to the college. The experimental study lasted for 20 weeks, in which a quasi-experiment was designed to detect students' performances of OEAL & FA and the impact of the use of WBFAT on their development of oral English proficiency. After the experiment, the main work for the researcher was to carry out data analysis, integration, and interpretation.

*Figure 8 The Research Procedure of This Study*



#### 4.3.1 Pilot Study

A pilot study is considered an important way to promote the quality of research outcomes in a study by identifying limitations in a small-scale trial and modifying a study protocol. It is the first step in an entire study protocol and usually starts with a small-scale trial. The objectives of conducting a pilot study are to examine the feasibility of the study protocol and provide principles and reasonable evidence for revising the study protocol (Avery et al., 2017; Pearson et al., 2020; Thabane et al., 2010). Thus, it can prevent fatal mistakes in the subsequent large-scale experimental studies (In, 2017). Meanwhile, a pilot study can help the researcher to accumulate research experience. In the present study, an internal pilot study was incorporated into the research design of the main study. Through this internal pilot study, the researcher was able to verify the feasibility of the study protocol, recruit participants in accordance with ethical requirements, and obtain their consent.

##### *Examining the feasibility of the study protocol*

After discussing with the supervisor in the preliminary stage, the themes and research questions of this study were determined. The researcher drew a draft of the study protocol according to the research goals and research questions, the main content of which was mainly reflected in PG1. In order to verify the feasibility of the study protocol, the researcher randomly selected 30 students from the classes of CE1 (the foundation module of college English) to conduct a pilot study, who took part in a short questionnaire consisting of six multiple-choice questions (see Appendix 1). From the statistical analysis of these six questions, the researcher preliminarily found that among the 30 students surveyed, most participants (96.67%) had used a web-learning tool, and more than half of the participants (53.33%) had used a WBFAT. Liulishuo English was most commonly used, followed by Italki, Kouyuyi, and FiF, and the frequency of use of web-based learning tool (WBLT) or WBFAT was mainly concentrated in 1-30 minutes per week. As for the effectiveness of WBLT or WBFAT, the majority of participants (60%) thought that its effectiveness was not ideal, and the function deficiencies of the tools in evaluating and giving feedback were considered to be the biggest concern.

The results from the analysis of questions 1 and 2 verify that the present study is realistic and necessary, because most students used or are using WBLT or WBFAT, and

results of questions 5 and 6 show that these users had experienced dissatisfaction or confusion with the WBLT or WBFAT. The researcher conducted a detailed survey and comparison of the four most commonly used WBLT or WBFAT according to answers to questions 3, 5 and 6, and finally decided to use FiF as a tool to assist students' autonomous oral English learning and formative assessment in the later experimental study. According to the slogan on the official website of FiF, it is a systematic professional oral English learning and assessment tool accompanied by a mobile app, which has been adopted by more than 1,000 undergraduate colleges and universities in mainland China as a major teaching aid for oral English teaching, learning, and evaluation. Several students in the pilot study said they did not use FiF because it charged highly. The introduction to the advantages of FiF's functions could be found in the branch (3.2.5) in this chapter. In view of this, the researcher negotiated with iFLYTEK and finally gained the company's support by agreeing to provide free access to the participants in this study free of charge. In addition, the researcher added a variable 'the number of times learners adjust learning strategies according to evaluation and feedback' to the module 'Autonomy', and changed the measurement method of the 14 variables in the control group into the means of manual recording.

In conclusion, the findings in the pilot study provide support for the further understanding of students' use of WBLT or WBFAT and their feelings. Meanwhile, they provide evidence for the researcher to identify the shortcomings of study protocol. From the analysis of the answers to these questions, the researcher obtained the proportion of EFL learners in vocational colleges using Web-learning or WBFAT, the frequency of use, the effect and limitations of their use, and chose FiF as a tool to assist students' oral English learning and formative assessment, and formed a plan of students' autonomous English learning and formative assessment in the following experimental study. Moreover, full consideration had also been given to the variables and key outcome measurements in the experimental study.

### ***Recruitment and consent***

The recruitment of participants in this study consists of two phases. Firstly, student participants were recruited from JTC to take part in a survey on learners' preliminary

attitudes towards using WBFAT. Then, some of those who took part in the first phase of the questionnaire were selected to take part in the second phase of the experimental study.

Ethical rules were complied with when recruiting participants in this study. The researcher firstly submitted a written application and obtained approval from the authority of the JTC academic office before starting the survey and experimental study, then posted an announcement for recruiting participants on a bulletin board on campus, with 300 printed copies of the written consent form on an open shelf next to the bulletin board. The announcement consisted of three sections. The first part was a notification, in which the background to the research, purposes of the research, data collection, data use, participant privacy protection, and the way of withdrawing from the study were expressed. The second part was the consent to participate in the first phase of the study, and the third part was the consent to participate in the second phase of the study. The interested participants were advised to mark in the box that corresponded with the terms, and chose whether to participate in the first phase of the study, the second phase of the study, or both phases of the study. The interested student participants were given one week to sign in by scanning the QR code on the announcement and submitting a written consent form to the researcher. Within one week, the researcher received written consent forms from 131 students who volunteered to participate in the phase 1 study, and 57 of those students agreed to participate in the phase 2 experimental study.

All students who submitted consent forms were by default considered eligible as participants since the first phase of the study only involved a survey and interviews. Participants in the second phase were advised to take part in the CE2 study, so the researcher tended to choose the freshmen, as the sophomores and juniors had already completed this curriculum. Therefore, 49 first-year students were preselected finally for the experimental study.

#### **4.3.2 Survey**

After the pilot study, in order to obtain student participants' attitudes towards using WBFAT and lay a foundation for the experimental research in the later stage, the researcher conducted a questionnaire survey on 28, June 2022 on EFL learners' attitudes towards FA and the use of WBFAT. The questionnaire includes five demographic questions, including participants' gender, major, grade, experience in using WBLT and the use of WBFAT, and

22 five-level Likert-scale questions (strongly disagree-1, disagree-2, neutral-3, strongly agree-4, strongly agree-5, see Table 10) which were divided into three dimensions: willingness to use (Q1-7), motivation to use (Q8-17), and expectation to use (Q18-22). 119 students from the 131 students who had submitted written consent during the pilot study phase participated in the survey.

*Table 10 Likert-scale Questions in the Survey on Students' Attitudes Towards Using WBFAT*

| Questions                                                             | 1 | 2 | 3 | 4 | 5 |
|-----------------------------------------------------------------------|---|---|---|---|---|
| 1.I agree to use WBFAT unconditionally                                |   |   |   |   |   |
| 2.I agree to use the WBFAT if required                                |   |   |   |   |   |
| 3.I agree to use the WBFAT if it was proved useful to learning        |   |   |   |   |   |
| 4.I agree to chose a WBFAT that suits my needs                        |   |   |   |   |   |
| 5.I agree to use of WBFAT in oral English learning                    |   |   |   |   |   |
| 6.I agree to use WBFAT in other courses                               |   |   |   |   |   |
| 7.I agree to start using WBFAT now                                    |   |   |   |   |   |
| 8.I want to use WBFAT to increase engagement in autonomous learning   |   |   |   |   |   |
| 9.I want to use WBFAT to increase interest in learning                |   |   |   |   |   |
| 10.I want to use WBFAT to increase joys in learning                   |   |   |   |   |   |
| 11.I want to use WBFAT to improve confidence in learning              |   |   |   |   |   |
| 12.I want to use WBFAT to enhance capabilities of IT (Deleted)        |   |   |   |   |   |
| 13.I want to use WBFAT to increase interaction with peers remotely    |   |   |   |   |   |
| 14.I want to use WBFAT to increase interaction with teachers remotely |   |   |   |   |   |
| 15.I want to use WBFAT to enhance the personalisation of learning     |   |   |   |   |   |
| 16.I want to use WBFAT to enhance the autonomy of learning            |   |   |   |   |   |
| 17.I want to use WBFAT to enhance the convenience of learning         |   |   |   |   |   |
| 18.I wish I could easily get permission to use WBFAT                  |   |   |   |   |   |
| 19.I wish to use WBFAT in class                                       |   |   |   |   |   |
| 20.I wish to use WBFAT in autonomous learning after class             |   |   |   |   |   |
| 21.I wish to choose a WBFAT that is suitable for my needs             |   |   |   |   |   |
| 22.I wish to use WBFAT for free                                       |   |   |   |   |   |

As shown in Table 11, the gender distribution of participants ( $n = 119$ ) shows an overwhelming majority of females among the participants. The frequency of female ( $n = 104$ , 87.4%) and male ( $n = 15$ , 12.6%) indicates a significant gender imbalance in this sample, with significantly more female than male. From the perspective of grade distribution, first-year students accounted for the largest proportion, with 83 students (69.7%), far more than second-year students ( $n = 23$ , 19.3%) and third-year students ( $n = 13$ , 10.9%). In terms of the distribution of majors, the number of students majoring in education is the largest ( $n = 65$ , 54.6%), overwhelmingly higher than students majoring in literature ( $n = 26$ , 21.8%), students majoring in engineering ( $n = 14$ , 11.8%), students majoring in management ( $n = 9$ , 7.6%), and students majoring in art ( $n = 5$ , 4.2%). Regarding the experience of using WBLT, the data showed that all participants had experience of use WBFA, which indicates that WBLT has a high acceptance among

vocational college students. While the data shows that the number of participants who had never used WBFAT before ( $n = 76$ , 63.9%) was much larger than the number who had ( $n = 43$ , 36.1%), suggesting that the use of WBFAT is not widespread among vocational college students.

*Table 11 Demographic Statistics of the Answers to the Questions in the Survey of Students' Attitudes (N=119)*

|                           | Frequency | Percentage |
|---------------------------|-----------|------------|
| Gender                    |           |            |
| Male                      | 15        | 12.6       |
| Female                    | 104       | 87.4       |
| Major                     |           |            |
| Pedagogy                  | 65        | 54.6       |
| Literature                | 26        | 21.8       |
| Engineering               | 14        | 11.8       |
| Management                | 9         | 7.6        |
| Arts                      | 5         | 4.2        |
| Grade                     |           |            |
| 1st year students         | 83        | 69.7       |
| 2nd year students         | 23        | 19.3       |
| 3rd year students         | 13        | 10.9       |
| Experience of using WBLT  |           |            |
| Yes                       | 119       | 100.0      |
| No                        | 0         | 0          |
| Experience of using WBFAT |           |            |
| Yes                       | 43        | 36.1       |
| No                        | 76        | 63.9       |

### 4.3.3 Experiment

The experimental study was scheduled to take place after the investigation, but due to the following 6-week summer break, the students ended their studies and left school. As a result, the experimental study was expected to be delayed until the end of August, but due to the Covid-19 pandemic, students were unable to return to school at that time, which forced the experimental study to be postponed again, as the experimental study required the researcher to be in the classroom for observation and essential interaction with teachers and students. It wasn't until 25 November 2022 when students returned back to the college, that the experimental study became possible.

The first week of the experimental study began on November 27 and lasted until May 13, 2023 (excluding three weeks of the winter holiday). 44 of the 49 first-year students selected in the pilot study participated in the previous phase of the survey (among 119 people), and with their written consent, these 44 student participants, 1 instructor participant, and 1 FiF technician participated in the experimental study. The 44 participants were randomly divided into two groups, the experimental group and the control group,

which were assigned separately to two different classes taught by Martin\* (pseudonym). The purpose of placing these students in different classes is to achieve the objectives of the quasi-experimental design in this study, which is to contrast the differences in the performance between the two groups in the process of OEAL & FA as well as the differences in the development of oral English proficiency between the two groups through the implementation of an intervention ( see Table 12) to effectively determine the effect of the WBFAT on the development of students' oral English proficiency.

*Table 12 Intervention Measures for OEAL & FA in the Experimental Study*

| <b>Intervention measures</b>                                       | <b>CG</b> | <b>EG</b>                                                                                       |
|--------------------------------------------------------------------|-----------|-------------------------------------------------------------------------------------------------|
| Intervention of assisted tool in OEAL & FA (use of WBFAT)          | /         | Use FiF over time                                                                               |
| Intervention of OEAL & FA                                          |           |                                                                                                 |
| Time spent requirement                                             | /         | No less than 40 minutes per week                                                                |
| The number of tasks completed                                      | /         | Complete one task respectively in accuracy, fluency, and completeness of oral English each week |
| Adjusting learning strategies according to evaluation and feedback | /         | Adjust learning strategies according to each evaluation and feedback                            |
| Peer evaluation and interaction                                    | /         | Take part in the peer assessment and interaction with classmates                                |

### ***Records of students' engagement in OEAL & FA during these 20 weeks***

Based on the learning objectives and assessment of the CE2 autonomous learning module and the intervention measures for OEAL & FA, with the instructor, Martin, the researcher defined the contents of 'Autonomy', 'Competence', and 'Relatedness' (Ryan & Deci, 2020) of participants' OEAL & FA for the experimental study under the guidance of Self-Determination Theory (SDT) of motivation (Deci & Ryan, 1985; Deci & Ryan, 2000).

#### ***Autonomy***

- a1: Set learning goals that suit the level of oral English
- a2: Choose learning materials that meet your needs
- a3: Setting of a reasonable workload
- a4: Choose the appropriate learning strategies
- a5: Adjust learning strategies based on feedback
- a6. Ask for support necessary to solve the problems in autonomous learning and FA

#### ***Competence***

- c1. Micro competence of oral English learning (accuracy, fluency, integrity)
- c2. Situated and authentic learning
- c3. Interactive learning

## ***Relatedness***

r1. Collaborative learning

r2. Community learning

Meanwhile, to obtain data on participants' performance of OEAL & FA, the researcher determined 14 variables to be observed. Among them, 4 variables are about 'Autonomy', they are 'setting learning goals' (a1), 'learning engagement' (a3), 'learning strategy adjustment' (a5), and 'seeking support or help' (a6); 6 about 'Competence', which are 'phonetic accuracy exercise' (c1-1), 'lexical accuracy exercise' (c1-2), 'grammatical accuracy exercise' (c1-3), 'expression coherence exercise' (c1-4), 'expression integrity exercise' (c1-5), and 'situational and simulation integration exercise' (c2); and 4 about 'Relatedness', namely, 'interaction with teachers' (r1-1), 'peer evaluation' (r1-2), 'cooperation with classmates' (r2), and 'community learning' (r3). Table 13 lists the specific variables and data acquisition methods.

*Table 13 Selected Variables of Students' Performances in OEAL & FA*

| First level of Variables | Second level of Variables (Indicators)                                        | Data source     |                                                                |
|--------------------------|-------------------------------------------------------------------------------|-----------------|----------------------------------------------------------------|
|                          |                                                                               | EG              | CG                                                             |
| Autonomy                 | Number of goals made for OEALFA (a1)                                          | Recorded by FiF | Students keep records and submit them weekly to the researcher |
|                          | Time spent on OEALFA (a3)                                                     |                 |                                                                |
|                          | Number of learning strategy adjustments based on evaluation and feedback (a5) |                 |                                                                |
|                          | Times of seeking support or help (a6)                                         |                 |                                                                |
| Competence               | Number of tasks completed in phonological accuracy practice (c1-1)            | Recorded by FiF | Students keep records and submit them weekly to the researcher |
|                          | Number of tasks completed in vocabulary accuracy practice (c1-2)              |                 |                                                                |
|                          | Number of tasks completed in grammatical accuracy practice (c1-3)             |                 |                                                                |
|                          | Number of tasks completed in practising coherence of oral English (c1-4)      |                 |                                                                |
| Competence               | Number of tasks completed in practising completeness of oral English (c1-5)   |                 |                                                                |
|                          | Number of tasks completed in simulation practice (c2)                         |                 |                                                                |
|                          | Times of interactions with teachers (r1-1)                                    |                 |                                                                |
| Relatedness              | Times of peer evaluation (r1-2)                                               |                 | researcher                                                     |
|                          | Times of cooperation with classmates (r2)                                     |                 |                                                                |
|                          | Times of participation in community learning (r3)                             |                 |                                                                |

## ***Scores of tests on students' oral English proficiency***

In order to explore the impact of WBFAT on the development of EFL learners' oral English proficiency in vocational colleges, more specifically, the researcher wanted to see whether WBFAT could significantly promote the development of students' oral English

proficiency, finding direct evidence to answer RQ3. Five monthly tests of oral English proficiency were held, respectively, at the 5th, 8th, 12th, 16th, and 20th weeks in the experimental study stage. The tests were based on level 3 of the Public English Test System (PETS) , consisting of two sessions: 1) Two minutes of personal information introduction to talk about past and present experiences and future plans, and 2) Four minutes of description and discussion of a chosen picture or topic. To facilitate automatic scoring and manual scoring, the 44 students in both the experimental group and the control group were asked to log in to FiF for testing. The student's answering voice had been automatically recorded and scored by FiF, including the overall score and the sub-scores of the four dimensions of pronunciation, semantics, fluency, and completeness. In order to make the evaluation more scientific and make up for the deviation in FiF automatic scoring, the researcher and Martin logged in to FiF, respectively, and scored manually after listening to the recording of the students' answers, which also included the total score and the four sub-scores of pronunciation, semantics, fluency, and completeness. Finally, the average of the FiF's score, the researcher's score, and Martin's score was used as the final score of the test and collected as data.

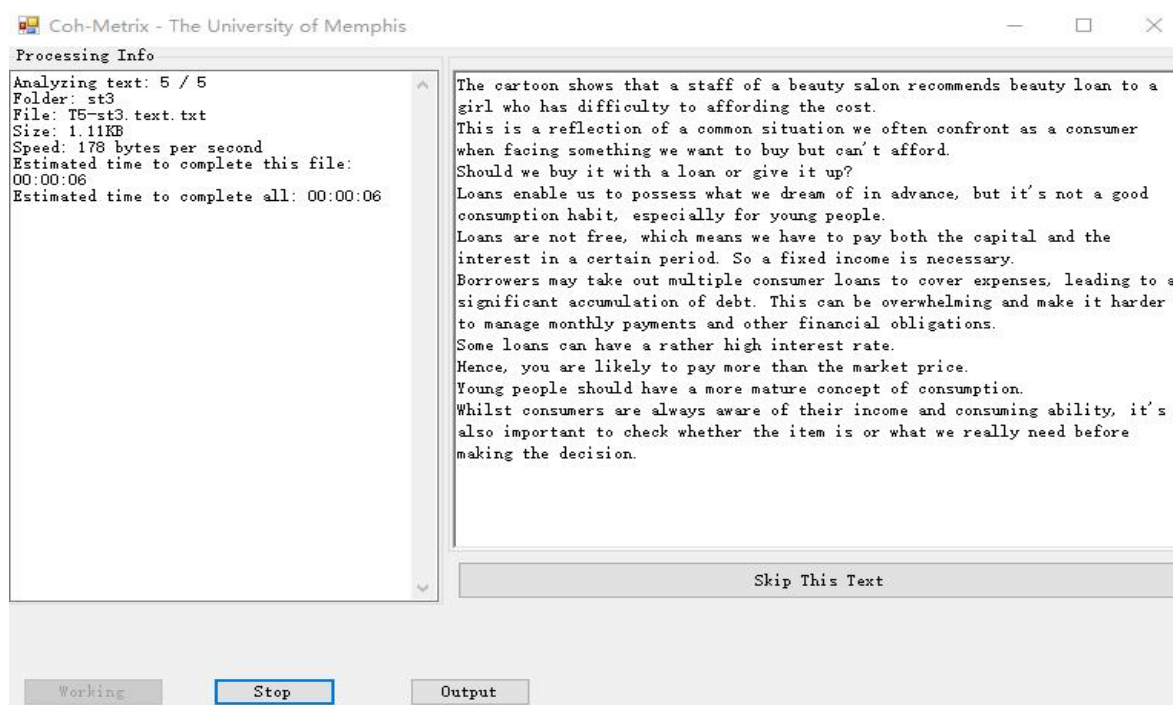
#### ***Data of students' oral English language characteristics***

The reliance on scores alone as the primary measure in the assessment of oral English proficiency may not provide a panorama of the individual's language abilities. The measure of oral English proficiency focuses on the speaker's fluent and accurate spontaneous expression with appropriate vocabulary and correct grammatical sentences in various situations (Guerra Ayala et al., 2024; Sánchez-Verdejo et al., 2019).

In order to further explore the development of students' oral English proficiency, the researcher selected students' recordings of the second part of five monthly tests and converted these audio files into text files using the automatic tool named Text-to-Speech Master (produced by iFlytek). These converted text files were then analysed by using Coh-Metrix 3.0 which is a computational tool of text and widely used to measure cohesion and linguistic sophistication by capturing linguistic features of language, discourse, and conceptual analysis (McNamara et al., 2010). Studies about Coh-Metrix showed that it could be used to generate indicators of linguistic and discourse representation of texts via the analysis of spoken and written English (Louwerse, 2001; Graesser et al., 2004). As

shown in Figure 9, after importing a text set into the database of Coh-Metrix 3.0, texts were processed one by one, and then 108 indices based on the linguistic features of the texts were automatically generated.

*Figure 9 A Snapshot of Analysis of Audio Transcripts of Students' Oral Tests via Coh-Metrix 3.0*



These indices could be categorised into 11 thematic groups: descriptive indicators, principal component scores for text readability, referential cohesion, latent semantic analysis (LSA), lexical diversity, connectives, situational models, syntactic complexity, syntactic pattern density, lexical information, and readability, which were widely used to study the cohesion and linguistic complexity of texts produced by second language writers and speakers (Crossley & McNamara, 2012). Finally, 20 indicators shown in Table 14 were selected to assess students' language features of students' audio transcripts extracted from the five tests.

*Table 14 Indicators of Coh-Metrix 3.0 Linguistic Profiling Selected in This Study*

| Indicators (Abbr.)                                            | Referential meaning                                 |
|---------------------------------------------------------------|-----------------------------------------------------|
| Descriptive indicators                                        |                                                     |
| DESWC                                                         | Word count, number of words                         |
| DESSL                                                         | Sentence length, number of words, mean              |
| Indicators of principal component scores for text readability |                                                     |
| PCNARp                                                        | Text Easability PC Narrativity, percentile          |
| PCSYNp                                                        | Text Easability PC Syntactic simplicity, percentile |
| PCREFp                                                        | Text Easability PC reference, percentile            |
| PCDCp                                                         | Text Easability PC Deep cohesion, percentile        |

Table 14 (continued)

| Indicators (Abbr.)                                                 | Referential meaning                                          |
|--------------------------------------------------------------------|--------------------------------------------------------------|
| Indicators of referential cohesion, latent semantic analysis (LSA) |                                                              |
| CRFAOa                                                             | Argument overlap, all sentences, binary, mean                |
| CRFCWO1                                                            | Content word overlap, adjacent sentences, proportional, mean |
| LSASS1                                                             | Latent Semantic Analysis, sentences, mean                    |
| SMTEMP                                                             | Temporal cohesion, tense and aspect repetition, mean         |
| WRDFAMc                                                            | Familiarity for content words, mean                          |
| Indicators of lexical information                                  |                                                              |
| PCCNCp                                                             | Text Easability PC Word concreteness, percentile             |
| WRDCNCc                                                            | Concreteness for content words, mean                         |
| LDTTRe                                                             | Lexical diversity, type-token ratio, content word            |
| LDTTRa                                                             | Lexical diversity, type-token ratio, all words               |
| WRDHYPn                                                            | Hypernymy for nouns, mean                                    |
| WRDHYPv                                                            | Hypernymy for verbs, mean                                    |
| Indicators of L2 readability                                       |                                                              |
| RDFRE                                                              | Flesch Reading Ease                                          |
| RDFKGL                                                             | Flesch-Kincaid Grade Level                                   |
| RDL2                                                               | Coh-Metrix L2 Readability                                    |

#### *Measures of descriptive indicators*

Two descriptive indicators, DESWC and DESSL, were selected as they are relatively easy to calculate and are more achievable in terms of improvement for beginner-level second language learners. DESWC reflects the length of the text by counting words via the Charniak parser, while DESSL reflects the complexity of sentences by calculating the average word of each sentence in the text (Graesser et al., 2004).

#### *Measures of principal component scores for text readability*

McNamara et al. (2014) and Graesser et al. (2011) suggest that Coh-Metrix plays a significant role in assessing text readability and comprehensibility. Through principal component scores such as narrativity (PCNARp), syntactic simplicity (PCSYNP), and word concreteness (PCCNCp), Coh-Metrix captured structural complexity and cohesion within texts, thus providing a scientifically grounded approach for analysing language materials. Crossley et al. (2012) and Wichien Gamla (2021) highlight that Coh-Metrix, by optimising textual structure and simplifying linguistic complexity, effectively enhances comprehension among non-native speakers, especially within specific reader groups such as second language learners and secondary school students. Furthermore, McNamara and Graesser (2012) highlight Coh-Metrix's unique contribution to theoretical and applied linguistics, where its multi-level analysis of textual features (especially word specificity

and narrative) reveals how textual cohesion affects the reader's experience of comprehension.

#### *Measures of cohesion*

McNamara et al. (2014) argue that Coh-Metrix is highly effective in measuring textual cohesion, as it comprehensively evaluates coherence through indicators such as referential cohesion, deep cohesion, and argument overlap. These measures enable researchers to better understand text complexity and readability. Graesser et al. (2011) and McNamara et al. (2014) believe that the multi-layered feature analysis provided by Coh-Metrix has widespread applications in education. In particular, indicators such as referential cohesion, deep cohesion, and Latent Semantic Analysis (LSA) support educators in evaluating the suitability of reading materials, thereby optimising educational content to improve student comprehension. McCarthy et al. (2007) suggest that cohesion metrics in Coh-Metrix, including LSA and content word overlap, effectively quantifies textual cohesion, showing that highly cohesive texts enhance reading comprehension in educational settings. Ozuru et al. (2009) find that, in scientific texts, cohesion measures from Coh-Metrix significantly impact comprehension across different reader levels. Using referential and deep cohesion indicators, they observe that higher cohesion particularly aids understanding among lower-proficiency readers.

#### *Measures of lexical diversity and lexical information*

McNamara et al. (2014) argue that Coh-Metrix is highly valuable in assessing lexical diversity and lexical information within texts, as it uses measures such as LDTTRc (content word type-token ratio) and WRDFAMc (content word familiarity) to evaluate lexical richness and familiarity. This approach aids in understanding text complexity and the cognitive load on readers. Duran and McCarthy (2010) suggest that lexical diversity is a key indicator of writing quality and proficiency. They find that Coh-Metrix's analysis of lexical diversity reveals that higher quality writing is often associated with more varied lexical use, an insight that also holds significance in automated essay scoring systems. Salsbury et al. (2011) note that lexical information plays a crucial role in the oral expression of second language learners. By examining measures such as WRDFAMc and WRDCNCc (word concreteness), they find that higher familiarity and concreteness in vocabulary enhance fluency in second language learners' spoken language. Crossley et al.

(2012) further observe that in simplified texts, reducing lexical diversity while increasing lexical familiarity can aid lower proficiency readers in comprehending the text more effectively.

### *Measures of Readability*

McNamara et al. (2014) suggest that Coh-Metrix readability indices, RDFRE, RDFKGL, and RDL2, are effective for assessing text difficulty and are widely used in selecting educational and academic materials. Graesser et al. (2011) and McNamara et al. (2014) note that these indices support educators in aligning content with students' grade levels, optimising curriculum design. Crossley et al. (2008) argue that RDFRE and RDL2 are particularly effective for evaluating reading difficulty for non-native learners, helping teachers identify key learning challenges. Snow et al. (2015) further observe that lowering RDFRE and RDFKGL scores enhances readability, making materials more accessible to language learners. Allen et al. (2014) highlight that these readability indices provide a reliable basis for selecting suitable materials in language teaching.

### *Class observation*

The classroom observation in this study started from the 5th week after the experimental study began. According to the CE2 course schedule of EG and CG's class, the researcher entered the class every 2-3 weeks for observation. The specific arrangement is shown in Table 15. It can be seen that the researcher arranged a total of 12 classroom observations during the experimental research stage, of which the class where EG is and the class where CG is are 6 times each.

*Table 15 The Schedule of Class Observation*

| <b>Week</b>  | <b>EG observation</b> | <b>Date</b> | <b>CG observation</b> | <b>Date</b> |
|--------------|-----------------------|-------------|-----------------------|-------------|
| 5th          | Class observation     | 28 Dec,2022 | Class observation     | 29 Dec,2022 |
| 8th          | Class observation     | 8 Feb,2023  | Class observation     | 9Feb,2023   |
| 11th         | Class observation     | 2 Mar,2023  | Class observation     | 2 Mar,2023  |
| 14th         | Class observation     | 22 Mar,2023 | Class observation     | 23 Mar,2023 |
| 17th         | Class observation     | 12 Apr,2023 | Class observation     | 13 Apr,2023 |
| 20th         | Class observation     | 10 May,2023 | Class observation     | 11 May,2023 |
| <b>Total</b> | <b>6</b>              |             | <b>6</b>              |             |

The performance of students' oral English in classes in the EG and CG was recorded by the researcher. The researcher focused on three aspects in the class observation: 1) the initiative of students using oral English in class; 2) the accuracy, cohesion, and coherence of oral English expression in class activities; 3) the performance of peer evaluation and collaboration, The observations were not recorded or videotaped, as the

level of detail in the language used in the classroom was not necessary for the study, but handwritten or typed notes were recorded. All handwritten notes were later printed out on a computer, and all notes were used for analysis. Classroom observation was mainly used to answer RQ4, so as to explore the differences in oral English performance between the experimental group and the observation group in class, so as to infer whether the independent oral English learning supported by WBFAT could promote the development of students' oral English level in class. In addition, findings from classroom observations could be used as complementary evidence for RQ3.

### ***Teacher interview***

A total of 4 teacher interviews were conducted in this study, as shown in detail in Table 16. Interview is one of the most common and effective ways we learn about others (Fontana & Frey, 2005), as it could provide us with subtly different lenses on the world (Warren, 2002). The purpose of the interview was to obtain a deeper and more comprehensive understanding of the relevant research questions from another perspective, such as the voice, attitude, and expression of the participants.

*Table 16 The Schedule of Instructor Interviews*

| <b>Interview</b>     | <b>Interviewee</b>  | <b>Objectives</b>                                                      | <b>Date</b>        | <b>Places</b> |
|----------------------|---------------------|------------------------------------------------------------------------|--------------------|---------------|
| Instructor interview | Instructor (Martin) | To get in-depth understanding of students' performance in class (RQ5 ) | 9, Feb 2023 (W8)   | Meeting room  |
|                      |                     |                                                                        | 2, Mar 2023 (W11)  | of English    |
|                      |                     |                                                                        | 12, Apr 2022 (W17) | Department    |
|                      |                     |                                                                        | 11, May 2023 (W20) |               |

The four interviews with the instructor, Martin, were all conducted after the class observation. It was the purpose that the researcher wondered to learn more about students' performance of oral English in classes in the two groups and to find whether there were significant differences in their performance between the two groups of students in oral English in class. The findings could directly answer RQ4 and provide support for RQ3. All interviews met ethical requirements, and the respondent was consulted and informed of the purpose of the interviews before each interview. Questions in the interviews were all semi-structured. The list of questions used in all interviews can be seen in Appendix 6. As the interviewees are proficient in Chinese, in order to make questions and answers easier to be understood and recorded, Chinese was used in some questions and answers during the interview. On the premise of obtaining the consent of the interviewee, all interviews were recorded throughout.

### *Survey of satisfaction and importance of FiF*

On April 25, 2023, the researcher conducted a third survey on the importance and satisfaction of FiF via a questionnaire. Participants were invited to the appointed classroom to answer the questions on the spot. The questionnaire included 20 five-level Likert scale questions (see Table 17), which included four dimensions, namely: ease of operation of the WBFAT (6 questions), ease of operation of the WBFAT (3 questions), supports for OEAL & FA (5 questions), and functions of evaluation and evaluation feedback (6 questions).

*Table 17 Scales in the Survey of Students' Perceptions of WBFAT*

| Satisfaction |   |   |   |   | Dimension                                                                                 | Importance |   |   |   |   |
|--------------|---|---|---|---|-------------------------------------------------------------------------------------------|------------|---|---|---|---|
| 1            | 2 | 3 | 4 | 5 | <b>Dimension1: Ease of operation of the WBFAT (FiF)</b>                                   | 1          | 2 | 3 | 4 | 5 |
|              |   |   |   |   | 1. It is easy to log in to                                                                |            |   |   |   |   |
|              |   |   |   |   | 2. It has friendly interface                                                              |            |   |   |   |   |
|              |   |   |   |   | 3. It is easy to navigate                                                                 |            |   |   |   |   |
|              |   |   |   |   | 4. It offers instructions for tasks                                                       |            |   |   |   |   |
|              |   |   |   |   | 5. It is smooth to use                                                                    |            |   |   |   |   |
|              |   |   |   |   | 6. It has satisfactory online service                                                     |            |   |   |   |   |
|              |   |   |   |   | <b>Dimension2: Ease of operation of the WBFAT (FiF)</b>                                   |            |   |   |   |   |
|              |   |   |   |   | 7. It provides massive learning resources                                                 |            |   |   |   |   |
|              |   |   |   |   | 8. Learning resources are updated regularly                                               |            |   |   |   |   |
|              |   |   |   |   | 9. It supports learners to build their own learning resources                             |            |   |   |   |   |
|              |   |   |   |   | <b>Dimension3: Supports for oral English autonomous learning and formative assessment</b> |            |   |   |   |   |
|              |   |   |   |   | 10. It supports authentic learning                                                        |            |   |   |   |   |
|              |   |   |   |   | 11. It supports human-machine interaction                                                 |            |   |   |   |   |
|              |   |   |   |   | 12. It supports peer communication                                                        |            |   |   |   |   |
|              |   |   |   |   | 13. It supports teacher-student interaction                                               |            |   |   |   |   |
|              |   |   |   |   | 14. It supports gamified learning                                                         |            |   |   |   |   |
|              |   |   |   |   | <b>Dimension4: Functions of evaluation and feedback</b>                                   |            |   |   |   |   |
|              |   |   |   |   | 15. It provides synchronous assessment for learning                                       |            |   |   |   |   |
|              |   |   |   |   | 16. It provides accurate assessment for learning                                          |            |   |   |   |   |
|              |   |   |   |   | 17. It provides consistent assessment for learning                                        |            |   |   |   |   |
|              |   |   |   |   | 18. It provides immediate feedback for learning                                           |            |   |   |   |   |
|              |   |   |   |   | 19. It provides targeted and instructive feedback for learning                            |            |   |   |   |   |
|              |   |   |   |   | 20. It provides statistical analysis of each task/learning                                |            |   |   |   |   |

### *Focus group discussion*

To obtain more qualitative data about students' perceptions of the use of WBFAT, a student focus group discussion was held on May 9, 15:00-15:45 PM. 18 of the 22 students who participated in the experimental group participated in the focus group in the public conference room of the Academic Affairs Office in the last week of the experiment study. Six main questions were raised based on RQ5 and RQ6 during the focus group discussion that lasted 45 minutes. In order to get better communication and understanding, the discussion was conducted in the participants' native language, Chinese. The whole process of the discussion was recorded and transferred into English with the prior consent of all

participants. During the focus group, the researcher took notes on the key points of what the participants said, which will be used to reconstruct the focus group discussion in case there was any problem with the recording. Since all the audio data was well recorded, only the translated text was used in the data analysis.

#### **4.4 Role of the Researcher, Instructor and Technician**

##### **4.4.1 Role of the Researcher**

The role of the researcher is crucial in the process of defining the research question, setting the research objectives, designing the research framework, selecting the research object, and exploring the research findings, but it is highly debated in research as well. The researcher role is highly debated in qualitative research. This article concerns the researcher-researched relationship. Karagiozis (2018) points out the complexities of the researcher's role. The researcher played four roles in the current study.

First, the collaborators. During the study, I cooperated with the instructor participant, Martin, to design the GE2 oral autonomous learning and evaluation plan before the experimental study, and completed the scoring for participants' tests. Moreover, in the whole process of this study, I cooperated with all student participants, the technician of FiF, and the relative staff of JTC. The second was the coordinator. In order to make the study run smoothly, I continued to coordinate the dynamic relationship between the researcher, the one instructor, the student participants, and the FiF technician throughout the study phase, especially when conflicts arose between them or between more than one of them. The third was the observer. In the experimental research stage, I reserve the right to check the performance of students' autonomous oral learning and evaluation on the FiF system at any time, which was informed to the students when recruiting participants for the experimental research. In addition, in order to further explore whether the oral performance of EG students with the support of WBFAT is significantly better than that of CG students in the GE2 class, I went to the classrooms of their respective classes six times, respectively, to observe and record their performance of oral English in class. Fourthly, as a rater. In the experimental research stage, 6 staged tests were set up. After the test was scored by IF-S for students, the researcher and Martin scored manually for the tests.

##### **4.4.2 Role of the Instructor**

There was one instructor participant in this study, Martin, the instructor of the classes class 1, 2022 - GE2 (in this study randomly set as the experimental group), class 2, 2022 - GE2 (in this study randomly set as the control group). The work and contributions undertaken by Martin in this study are shown in Table 18.

*Table 18 The Instructor Participant's Contribution to the Present Study*

| <b>Class</b>             | <b>Tasks and contribution</b>                                                                                                                                                                                                                                                                                                                                                                                                       |
|--------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Class A, 2022 - GE2( EG) | <p>Worked with the researcher to develop learning plan for students in the experimental group</p> <p>Encouraged and supervised the 22 students in the experimental group to use the FiF in oral English autonomous learning and assessment, and provided necessary supports when students encountered difficulties and problems</p> <p>Accept 4 interviews during the experiment study</p> <p>Manual rating the students' tests</p> |
| Class B, 2022 - GE2(CG)  | <p>Includes the 4 tasks mentioned above in the EG class</p> <p>Recorded students' performance of OEAL &amp; FA in CG, and gave students feedback on their performance</p>                                                                                                                                                                                                                                                           |

The reason why I chose Martin as the instructor participant of this study was that he was the teacher of classes of GE2, which students of EG and CG were in. On the other hand, after I expressed the purpose of the research to him and invited him, he showed strong interest in participating in my research. Besides, he was serious and energetic, and he hoped very much to explore ways and paths to effectively improve students' oral English proficiency.

Martin was an experienced foreign teacher from the Republic of Slovenia with a master's degree in education. He had a certificate of TEFL (i.e. teaching English as a foreign language). He grew up and was educated in the UK, so he was proficient in English. He came to China 10 years ago. After three years working at a normal university in Jiangxi province, he came to JTC, where he has been working for six years. Therefore he could carry out smooth communication with his colleagues and students in Chinese. Since coming to JTC, he had been engaged in teaching oral English and GE. Therefore, he was very clear about the problems of the student's oral English development. He was so outgoing and humorous that he liked to communicate with students after class and was very popular among them. Martin said he was very interested in technology-assisted teaching and assessment because he needed to use technology, such as web-based teaching platforms, oral English practice apps, frequently to aid his teaching of oral English. This experience provided lots of help for my experimental study. It was particularly worth mentioning that Martin, as the instructor of both the experimental group and the control

group, had to manually score 22 students' oral English practice tasks and give them oral or written feedback every week, which was very heavy and patient work. However, after the end of the experimental study, he provided me with the weekly recording hours and grading records for each student, which was an amazing job.

#### **4.4.3 Role of the Technician**

This study was supported by IFlytek Co., LTD., which provided the researcher and participants with free access to the FiF for 24 months. The 22 students in the control group were granted temporary permission to freely use the FiF system to attend the oral English tests. The technician of FiF, Peter\* (pseudonym) was appointed to provide technical support and service for all participants. FiF, unlike other free mobile apps, is a professional web-based oral learning and evaluation system that requires high fees to log in and use without authorisation. Therefore, it was very strange to all participants, although a small number of students in the survey (36.1% of 119 participants in the first phase and 11.4% of 44 participants in the experimental phase) had ever used the WBFAT. During the pilot study, Peter held a 2-hour training and demonstration on the operation of FiF for all participants. During the experimental research, whenever the researcher, the instructor, or student participants encountered technical troubles, warm and professional service was always provided from him.

#### **4.5 Ensuring the Quality of Findings**

To ensure the validity of the data collected and the quality of the findings in this study (Anney, 2014; Treharne & Riggs, 2015), the researcher adopted measures such as pilot study (Thabane, 2010; In, 2017; Van, 2013), triangulation (Hartley, 1997; Thurmond, 2001; Heale, 2013; Oppermann, 2000), and inter-rater reliability checking (Gwet, 2001; Gwet, 2008; McDonald et al., 2019), statistical validation techniques (Belouafa et al., 2017; Kleijnen, 1999; Mayer & Butler, 1993), and eliminating the influence of researcher's own privilege to improve the validity and reliability of data.

##### **4.5.1 Pilot Study**

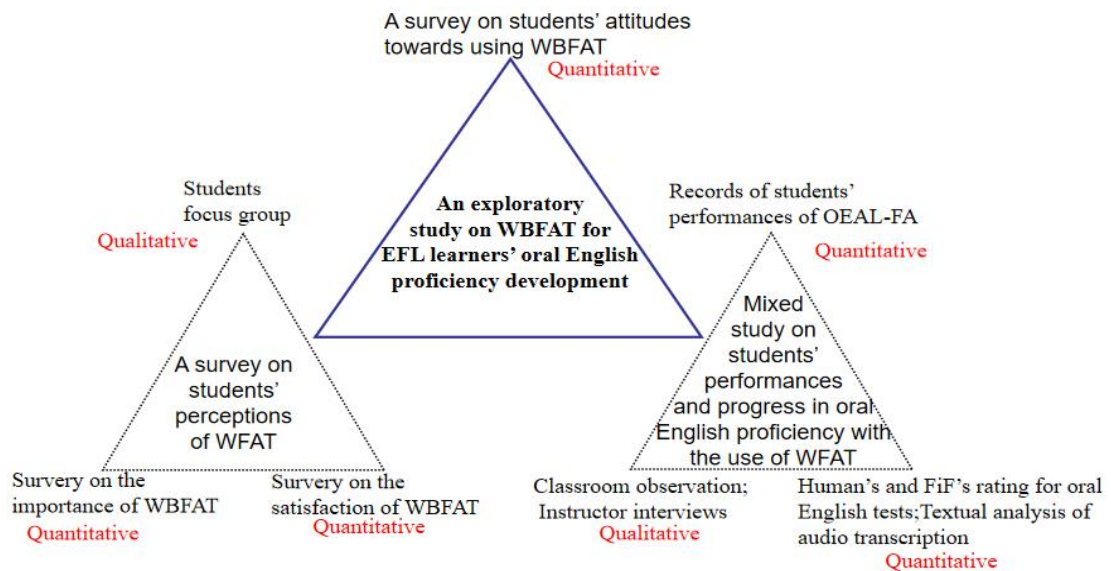
In order to verify the feasibility of the research scheme, a pilot study was conducted before carrying out a larger-scale investigation and the experimental study, in which rectification and preparation for the subsequent research activities were carried out. Firstly, the research gained a general recognition of students' experience of WBFAT use via a short

questionnaire among 30 randomly selected students. Secondly, FiF was decided to be used as a tool to assist students' autonomous oral English learning and formative assessment in the subsequent experimental research through investigation and comparison of the four types of WBLT/WBFAT most commonly used by students. Thirdly, a scheme of students' autonomous English learning and formative assessment had been formed for the experimental research. Fourthly, the selection of variables and the measurement of key results in the experimental study were fully considered. Last but not least, in the pilot study phase, the recruitment of candidates for participants in the later survey and experimental study was completed. All in all, the work in the pilot research stage had played a great role in promoting the smooth progress of the later research and achieving better results.

#### **4.5.2 Triangulation**

In order to enhance the rigour of the findings, the triangulation method (Carter et al., 2014; Korstjens & Moser, 2017) was introduced in this study. Firstly, a mixed approach that incorporated surveys, recording, testing, interview, and observation for data collection ensured the reliability of the data through multiple perspectives in both qualitative and quantitative methods, avoiding the one-sided understanding of the research questions from either quantitative data or qualitative data. As shown in Figure 10, based on the research question, in the process of data collection, one external trigonometry and two internal trigonometry were included in this study. The outer triangle revolved around the overall goal of this study, including a survey on students' attitude towards WBFAT, a mixed study on students' performance and progress in oral English proficiency with the use of WBFAT, and a survey on students' perceptions of WBFAT. One of the inner triangulations was the mixed study on students' performance and progress in oral English in the experimental study, which consisted of records of students' performance, human and FiF's rating for students' oral tests, textual analysis of audio transcription, classroom observation, and instructor interviews, and the other one was the survey on students' perceptions of WBFAT, which consisted of the survey on students satisfaction with WBFAT, the survey on the importance of WBFAT, and a students focus group.

Figure 10 Quality Triangulation Designed in the Present Study



While ‘Investigator Triangulation’ was employed as well by having at least two independent coders analyse qualitative data to guarantee a robust and dependable qualitative research process by integrating diverse perspectives and corroborating the interpretation of results through inter-coder agreement (Delve & Limpaecher, 2023, Keshf & Khanum, 2021). At the beginning of the experimental study, the researcher and the instructor, Martin, held several discussions before finalising the plan for students’ oral English automatic learning. Scores for students’ five tests of oral English during this phase were given by FiF automatically, while also Martin and the researcher manually. Prior to the manual scoring, we agreed on the scoring criteria and ranges, basing on the framework provided by FiF to ensure the consistency of rating.

In addition, the data obtained in the observation and interview were sorted out and returned to the interviewees and two teachers for reading and feedback. After receiving the responses from all the interviewees and two teachers, it was confirmed that everything reported was correct and objective, and then these materials were used as valid data for this study.

#### 4.5.3 Eliminating the influence of the Researcher’s Status as a Teacher

Before starting the pilot study and experimental study, the researcher declared to all participants the outline of research contents and purposes of the study, research methods, data collection methods, use of data results, their rights to privacy and confidentiality, and a clear declaration of the benefits and risks as well (Glesne, 1999; Miles and Huberman, 1994; Silverman, 2006). The study was conducted after obtaining the written consent

signed by all participants, which made the collection of research data and materials meet the ethical requirements.

Given the complex role of the researcher in this study, there was inherent influence of the researcher's position as a teacher in JTC, but this was accompanied by the vulnerabilities that researchers might encounter (Bashir, 2020; Muhammad et al., 2015; Råheim et al., 2016). To address potential conflicts arising from these various roles and to improve the accuracy of the research findings, the researcher applied reflexivity through ongoing dialogue and self-evaluation of their role. This approach aimed to minimise the potential influence of the researcher's positionality on both the research process and outcomes (Berger, 2015; Bradbury-Jones, 2007; Guillemin, 2004).

As the primary investigator, I recognised that my roles as a teacher at JTC and as the head of the Foreign Languages Department could have influenced the authenticity of students' responses during the survey or their behaviour during the experiment. I made every effort to reduce these potential biases. Therefore, for both the web-based attitude survey and the two surveys assessing student satisfaction and the perceived importance of FiF, all participants were invited to complete the questionnaires in a shared, public classroom. Before they began, I consistently reminded them to answer truthfully based on their own experiences and preferences. During the classroom observations, I sat quietly at the back of the room, listening and taking notes without offering any commentary, to minimise any interference with the students' authentic behaviour.

#### **4.6 Data Analysis and Integration**

As shown in Table 19, the researcher collected a total of eight datasets in response to three research themes and six research questions. Upon completing data collection, different analytical tools and research methods were employed depending on the nature of the data, aiming to provide a comprehensive and multi-perspective response to the research questions.

First, SPSS 26 was used to conduct descriptive statistical analysis of the survey data concerning students' attitudes towards web-based formative assessment tools (WBFAT). This analysis included the calculation of maximum values, minimum values, mode, mean, and standard deviation, as descriptive statistics are fundamental in understanding the basic characteristics of data. During the integration phase, it was

essential to systematically synthesise various descriptive statistics results. For instance, descriptive statistical analysis of questionnaire responses facilitated an overall understanding of students' attitudes towards WBFAT, while the recorded performance metrics of students in both the experimental and control groups (including maximum values, minimum values, and mean scores) laid the foundation for subsequent comparative analysis. By presenting statistical results in a structured format, such as tables or reports, a clearer and more organised representation of findings was achieved, allowing for more in-depth interpretation.

Second, SPSS 26 was used to perform descriptive statistical analysis on students' performance records during the oral English autonomous learning and formative assessment (OEAL & FA) process across three dimensions—autonomy, competence, and relatedness, covering fourteen variables (a1, a3, a5, a6, c1-1, c1-2, c1-3, c1-4, c1-5, c2, r1-1, r1-2, r-2, r3). This analysis included maximum values, minimum values, mode, mean, standard deviation, and kurtosis analysis. Subsequently, a *t*-test was conducted to examine differences between the experimental and control groups across these variables.

Third, the mean scores and standard deviations of students' five English-speaking tests were analysed descriptively to compare the performance of the experimental and control groups over time.

Fourth, Coh-Metrix 3.0 was employed to analyse 20 linguistic feature indicators extracted from speech transcripts of students' five English-speaking tests across the two groups, including 2 descriptive indicators (DESWC, DESSL), 4 indicators of principal component scores for text readability (PCNARp, PCSYNp, PCREFp, PCDCp), 5 indicators of referential cohesion, latent semantic analysis (CRFAOa, CRFCWO1, LSASS1, SMTEMP, WRDFAMc), 6 indicators of lexical information (PCCNCp, WRDCNCc, LDTTRc, LDTTRa, WRDHYPn, WRDHYPv), and 3 indicators of L2 readability (RDFRE, RDFKGL, RDL2). Following this, SPSS 26 was utilised to conduct mean score and standard deviation analysis, as well as *t*-tests, on these linguistic feature indicators for the five speech transcripts across both groups.

Fifth, NVivo 11 was used to code, categorise themes, and conduct frequency matrix analysis on six classroom observation records.

Sixth, NVivo 11 was also employed to conduct thematic analysis of four teacher interview transcripts, identifying recurring patterns and key insights.

Seventh, SPSS 26 was applied to perform descriptive statistical analysis on students' satisfaction with and perceived importance of the FiF tool. This analysis included mean score and standard deviation calculations, *t*-tests, and Importance-Performance Analysis (IPA) to assess students' perceptions comprehensively.

Eighth, NVivo 11 was utilised to analyse focus group discussion transcripts. The analysis included word cloud visualisation, coding frequency analysis, and matrix coding query analysis to identify prominent themes and frequently occurring terms.

By employing a combination of quantitative and qualitative analytical tools, the study ensured a rigorous and multifaceted examination of students' experiences and performance in the web-based formative assessment environment.

*Table 19 Design of Data Sources and Data Analysis in the Present Study*

| Themes                                                                                            | Research Questions                                                                                                                                                              | Data sources                                                                                                                                                            | Data Analysis Method                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|---------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Theme1:<br>Students' attitudes towards using WBFAT                                                | RQ1:What are students' attitudes towards the use of WBFAT?                                                                                                                      | Answers to the questionnaire                                                                                                                                            | Use SPSS26 to make descriptive statistical analysis (including calculating max., min., mode, frequency and percentage), correlation analysis between items and between dimensions                                                                                                                                                                                                                                                                                                                       |
| Theme2:<br>Students' performance and progresses in oral English proficiency with the use of WBFAT | RQ2:What is students' performance of oral English autonomous learning?<br><br>RQ3:Is there significant progress in oral English proficiency after students use WBFAT over time? | Records of students' engagement in OEAL & FA during 20 weeks<br>(1)Scores of five tests of oral English<br>(2)Students' audio transcripts extracted from the five tests | SPSS26 was used for descriptive statistical analysis (including max., min., mode, mean, SD and kurtosis), and <i>t</i> - test was conducted to analyse the differences of variables between the two groups<br>SPSS26 was used to make descriptive statistical analysis, including mean and SD<br>Coh-Metrix 3.0 was used to analyse the language characteristics, and Spss26 was used to analyse the mean and SD of the indicators, and <i>t</i> - test for analysing the differences of the two groups |
|                                                                                                   | RQ4:Is there any improvement progress in students' oral English performance in class after using WBFAT over time?                                                               | (1)Records of class observation<br>(2)Scripts of instructor interviews                                                                                                  | Use NVIVO 11 to conduct coding node, theme induction and coding frequency matrix analysis<br>Nvivo11 was used to summarise and analyse the records of instructor interviews                                                                                                                                                                                                                                                                                                                             |
| Theme3:<br>Students' perceptions of WBFAT                                                         | RQ5:What are students' perceptions of the satisfaction with WBFAT?<br><br>RQ6:What are students' perceptions of the importance of WBFAT?                                        | (1)Answers to the questionnaire<br>(2)Scripts of a focus group discussion                                                                                               | spss26 was used for descriptive statistical analysis (including mean, SD), T-test analysis and IPA analysis<br>Nvivo11 was used to analyse the data by word cloud analysis, coding frequency analysis and matrix coding query results                                                                                                                                                                                                                                                                   |

## 4.7 Chapter Summary

In this chapter, the research process and findings were reviewed and critically reflected upon, particularly with regard to the outcomes of the experimental study. The study demonstrated how the use of WBFAT influenced the development of oral English proficiency in EFL learners, providing valuable insights into both learner performance and satisfaction. However, the reflection also reveals several areas that require improvement, suggesting directions for future research.

The research began with a pilot study that successfully tested the feasibility of the proposed research design. This initial phase proved useful in identifying potential challenges related to data collection and participant recruitment, thus improving the quality of the main study. However, external factors, particularly the Covid-19 pandemic, caused a significant delay in the start of the experimental phase. While the study eventually resumed, the disruption may have affected students' engagement and their overall performance during the research period. The experimental study showed that the experimental group, which utilised WBFAT, made noticeable progress in oral English proficiency, particularly in areas related to fluency, accuracy, and syntactic complexity. Quantitative analyses using SPSS indicated statistically significant differences between the experimental and control groups, confirming that the intervention had a positive impact on learners' oral skills. Additionally, linguistic analyses conducted with Coh-Metrix 3.0 revealed improvements in language cohesion and syntactic sophistication among learners in the experimental group.

Furthermore, learner satisfaction with WBFAT was generally positive, as reflected in the questionnaire and focus group discussions. Learners particularly appreciated the platform's ability to provide immediate feedback and its support for autonomous learning. However, some participants reported difficulties with the platform's usability, particularly in terms of navigation and the user interface. This feedback suggests that future iterations of WBFAT should focus on improving the users' experience, making the platform more intuitive and accessible.

## Chapter 5 Findings

In this chapter, the researcher will analyse each of the eight sets of data according to the type and characteristics of the data collected in Chapter 4, using different analytical methods, and the results are reported in detail.

### 5.1 Students' Attitudes Towards Using WBFAT

As explained in Chapter 4, the initial survey of students' attitude towards WBFAT was carried out through a questionnaire that included 5 demographic questions and 21 five-level Likert scale items.

#### 5.1.1 Demographic Description of the Participants

As shown in Table 20, the average gender value among participants is 1.87, close to 2, indicating a higher proportion of female respondents, with a mode of 2 confirming this majority. The low standard deviation ( $SD = 0.33$ ) suggests a concentrated distribution around female respondents. The mean for the major variable is 1.85 ( $SD = 1.15$ ), indicating a broader range of fields, with most participants in major 1 (Education). A mean grade of 1.41, with a mode of 1 and a standard deviation of 0.68, suggests most respondents are first-year students, though grades are slightly dispersed among first- and second-years. All respondents have experience with WBLT (mean = 1,  $SD = 0.00$ ). For WBFAT usage, the mean is 1.64 (mode = 1,  $SD = 0.48$ ), indicating some variation, with most students lacking WBFAT experience.

*Table 20 Demographic Statistics of the Scales in the Survey of Students' Attitudes Towards Using WBFAT*

|                           | Mean | Mode | SD   |
|---------------------------|------|------|------|
| Gender                    | 1.87 | 2    | 0.33 |
| Major                     | 1.85 | 1    | 1.15 |
| Grade                     | 1.41 | 1    | 0.68 |
| Experience of using WBLT  | 1    | 1    | 0    |
| Experience of using WBFAT | 1.64 | 2    | 0.48 |

#### 5.1.2 Reliability and Validity of the Scale Items

Table 21 shows that the overall Cronbach's Alpha value for the 21 items in the questionnaire is 0.793, indicating good internal consistency and a high degree of correlation among the items. For 19 of the items, the Cronbach's Alpha if Item Deleted is noticeably lower than this value, while for two items (Q4 and Q22), the Cronbach's Alpha if Item Deleted is 0.793. However, this still suggests that the questionnaire exhibits strong

internal reliability. The Cronbach's Alpha values for each item after deletion can be found in Appendix 2.

*Table 21 Reliability Test of the Scales in the Survey of Students' Attitudes Towards Using WBFAT*

| Cronbach's Alpha | N of Items |
|------------------|------------|
| .793             | 21         |

As seen in Table 22, the KMO value for the 21 questionnaire items is 0.699, close to 0.7, indicating that the data is suitable, though not ideal, for factor analysis. The Approximate Chi-Square is 727.169 ( $df = 210$ ,  $Sig. = 0.000$ ), showing a significant Bartlett's test of sphericity and rejecting the null hypothesis that the correlation matrix is an identity matrix. This suggests correlations among variables, supporting factor analysis suitability.

*Table 22 KMO and Bartlett's Test of the Scales in the Survey of Students' Attitudes Towards Using WBFAT*

|                                                         |                    |         |
|---------------------------------------------------------|--------------------|---------|
| <b>Kaiser-Meyer-Olkin Measure of Sampling Adequacy.</b> |                    | .699    |
| <b>Bartlett's Test of Sphericity</b>                    | Approx. Chi-Square | 727.169 |
|                                                         | df                 | 210     |
|                                                         | Sig.               | .000    |

### 5.1.3 Statistical Analysis of 21 Scale Items in the Questionnaire

Overall, the statistical data in Table 23 indicates that respondents generally hold a *positive attitude towards WBFAT across most questions.*

*Table 23 Descriptive Analysis of the Scales in the Survey of Students' Attitudes Towards Using WBFAT*

|     | Min. | Max. | Mode | Mean | Std. |
|-----|------|------|------|------|------|
| Q1  | 2    | 4    | 3    | 3.21 | 0.69 |
| Q2  | 3    | 5    | 4    | 3.81 | 0.64 |
| Q3  | 3    | 5    | 4    | 3.72 | 0.54 |
| Q4  | 2    | 5    | 4    | 3.60 | 0.57 |
| Q5  | 3    | 5    | 4    | 3.81 | 0.56 |
| Q6  | 2    | 4    | 3    | 3.15 | 0.53 |
| Q7  | 2    | 4    | 3    | 3.29 | 0.64 |
| Q8  | 2    | 5    | 4    | 3.66 | 0.68 |
| Q9  | 2    | 5    | 4    | 3.50 | 0.53 |
| Q10 | 2    | 5    | 4    | 3.56 | 0.61 |
| Q11 | 2    | 5    | 3    | 3.34 | 0.54 |
| Q13 | 2    | 5    | 3    | 3.33 | 0.74 |
| Q14 | 2    | 5    | 4    | 3.38 | 0.76 |
| Q15 | 2    | 4    | 3    | 3.24 | 0.57 |
| Q16 | 2    | 5    | 4    | 3.50 | 0.58 |
| Q17 | 3    | 5    | 4    | 3.82 | 0.56 |
| Q18 | 2    | 5    | 4    | 3.70 | 0.62 |
| Q19 | 2    | 5    | 3    | 3.39 | 0.63 |
| Q20 | 2    | 5    | 4    | 3.66 | 0.66 |
| Q21 | 2    | 5    | 4    | 3.62 | 0.60 |
| Q22 | 3    | 5    | 4    | 3.69 | 0.56 |

Students affirmed the effects of WBFAT on the assist of their oral English learning as they hold generally positive attitudes, particularly regarding items related to learning outcomes (e.g., S8, S10, S16), which exhibit relatively high mean values. This suggests that respondents perceive WBFAT as effective in enhancing the learning experience and convenience. However, views on personalised learning (e.g., S14) appear somewhat neutral, possibly indicating that the benefits of the personalisation features have not been fully realised. Additionally, the standard deviations are generally small, indicating that responses to most items are relatively consistent, reflecting a shared consensus among respondents regarding the functionality of WBFAT.

***Strong positive items: Q2, Q5, Q17, and Q18***

Items such as Q2 (using WBFAT when needed), Q5 (using WBFAT in oral English learning), Q17 (improving learning convenience), and Q18 (hoping to gain access easily) generally have mean values close to or above 3.7 with a mode of 4 or 5, reflecting respondents' positive views of WBFAT in these areas. These questions focus on the practical use scenarios and learning aids of the tool, showing that WBFAT has certain practicability and positive influence in their minds.

***Slightly neutral items: Q6, Q14***

Items such as Q6 (using WBFAT in other courses) and Q15 (personalising experiences to enhance learning) have a lower mean, closer to 3.2, with a mode of 3, indicating that respondents' experiences or perceptions of these dimensions were more neutral, or that the role of WBFAT in these dimensions was not fully developed or understood. Personalised learning experiences (Q15), in particular, suggest that WBFAT's design for personalised learning support could be further improved.

***Standard deviation shows general consensus existed in EG and CG except for Q13 and Q14***

In terms of standard deviation, most items have relatively low values (mostly between 0.5 and 0.7), indicating that respondents' ratings were fairly concentrated, with little divergence. This also suggests that there was a certain level of consensus among respondents regarding their attitudes and expectations towards the use of WBFAT.

However, a few items show slightly higher standard deviations, exceeding 0.7, such as Q13 (increased remote interaction with peers) and Q14 (increased remote interaction

with teachers). This indicates greater variability in respondents' views on these interactive features and reflects that some believed WBFAT had room for improvement in remote interaction or that individuals had differing needs for these functionalities.

***Mode demonstrates most respondents affirmed the positive role of WBFAT in learning***

The modal analysis shows that most questions have a mode of 4 (suggesting a more positive response), which is consistent with the mean ratings across the majority of items. Responses to questions concerning practical learning applications, such as Q8 (enhancing participation in independent learning), Q9 (increasing interest in learning), and Q17 (improving the convenience of learning), are particularly concentrated around mode 4. This concentration further underscores respondents' positive perception of WBFAT's role in supporting learning.

***Special items mean students' requirements of remote interaction function in EG and CG vary greatly***

Q6 (use of WBFAT in other courses) and Q15 (personalised experiences to enhance learning) have relatively low means and modes and small standard deviations, suggesting that most respondents are more neutral about this. Personalised learning experiences, in particular, suggest that respondents' needs in this area are not being fully met and may require WBFAT to develop stronger capabilities in this area. Large standard deviations in Q13 (increased remote interaction with peers) and Q14 (increased remote interaction with teachers) mean that different respondents' experiences or needs for these features vary considerably. This may have a lot to do with the learning environment of the respondents, the need for remote interaction, or the specific use of WBFAT.

***Correlation among items reflects positive correlation exists among items and dimensions***

Based on the descriptive statistical analysis in Table 23, there is a positive correlation among the three questionnaire dimensions: willingness to use, motivation to use, and expectation to use. Overall, the three dimensions show a positive correlation, with higher willingness often paired with greater motivation and expectations. However, personalised and interactive features exhibit divergent needs, suggesting areas for WBFAT's improvement.

Firstly, a positive correlation exists between willingness to use and motivation to use. Higher mean scores and concentrated responses in willingness (e.g., Q1, Q5, Q7) and

motivation items (e.g., Q8, Q9, Q10) suggest that respondents who are more willing to use WBFAT are also more motivated, especially in terms of fostering autonomous learning and interest. Secondly, positive correlation also exists between motivation and expectation. Items in motivation and expectation dimensions, particularly those related to convenience and autonomous learning (e.g., Q17, Q21), show positive correlations, indicating that higher motivation often aligns with greater expectations for WBFAT's features. However, we can find divergence existing in the items related to the needs of personalisation and interaction of WBFAT. Scores for items on personalisation and remote interaction (e.g., Q14, Q15) are more varied, suggesting differing needs among respondents, which weakens the correlation between motivation and expectation in these areas.

## **5.2 Students' Engagement in OEAL & FA**

### **5.2.1 Descriptive Analysis of Students' Engagement in OEAL & FA**

Based on the analysis of the maximum, minimum, mean, mode, and standard deviation for the 14 variables involved in students' participation in independent English-speaking learning and formative assessment, the data reveals the following characteristics:

#### ***Significant performance differences***

From the standard deviation and range of several variables (such as time invested, tasks completed, and feedback received), it is evident that students exhibit substantial variation in their independent learning performance. Some students show a high level of engagement in terms of time spent, task completion, and feedback acquisition, while others demonstrate lower levels of engagement, with some students not completing certain tasks. This disparity reflects individual differences in learning motivation, learning habits, and learning strategies.

#### ***Prevalence of extreme values***

Many variables (such as the number of fluency tasks completed, coherence tasks, and appropriateness tasks) exhibit negative kurtosis, indicating a flatter distribution with a prevalence of extreme values. For instance, some students completed a large number of tasks, while others did not complete any. This suggests that, although the average performance reflects overall participation, the actual level of student involvement varies significantly.

### ***Consistent performance in certain tasks, poor performance in specific tasks***

Students generally performed well on more straightforward tasks (such as time investment, overall task completion, and pronunciation accuracy), with higher average completion rates and lower standard deviations, indicating more consistent performance across these tasks. However, for more challenging tasks requiring greater strategic adjustments (such as fluency, coherence, and appropriateness tasks), there was a marked difference in completion rates, with many students either not participating or completing fewer tasks. This may be related to task difficulty and students' understanding of the task requirements.

### ***Feedback acquisition and strategy adjustment***

There were also significant differences in students' performance with respect to feedback acquisition and strategy adjustment. Some students received extensive feedback and used it to continually adjust and improve their learning performance, while others received less feedback and adjusted their strategies less frequently. This highlights the importance of feedback in independent learning, as the timeliness and frequency of feedback can greatly influence learning outcomes.

### ***Low engagement in peer interaction and peer assessment activities***

In terms of peer interaction and peer assessment tasks, most students were relatively disengaged, particularly in giving and receiving peer ratings, where the mode was 0. This indicates that the majority of students did not participate in these activities, possibly suggesting a lack of emphasis on peer interaction and assessment, or that the task design did not effectively motivate students to engage.

### ***Time investment could affect, but does not determine learning outcomes***

As shown in Table 24, there appears to be some correlation between time investment and learning outcomes, with students who invested more time generally performing better on overall tasks and some individual tasks. However, time investment alone is not the determining factor for success. In more challenging tasks, learning strategies, the timeliness of feedback, and adjustments to learning methods play an equally crucial role.

Table 24 Descriptive Statistics of the Variables in Students' Engagement in OEAL & FA

|             | Min.   | Max.    | Mode   | Mean    | Std.   | Kurtosis |
|-------------|--------|---------|--------|---------|--------|----------|
| Variable 1  | 800.00 | 1764.00 | 840.00 | 1295.82 | 275.77 | -0.84    |
| Variable 2  | 384.00 | 821.00  | 663.00 | 640.70  | 111.14 | -0.42    |
| Variable 3  | 21.00  | 49.00   | 22.00  | 30.00   | 8.63   | -0.83    |
| Variable 4  | 21.00  | 64.00   | 37.00  | 42.43   | 9.64   | 0.03     |
| Variable 5  | 0.00   | 42.00   | 0.00   | 14.70   | 14.73  | -1.38    |
| Variable 6  | 15.00  | 51.00   | 22.00  | 31.09   | 10.36  | -1.03    |
| Variable 7  | 0.00   | 29.00   | 0.00   | 9.50    | 8.86   | -1.07    |
| Variable 8  | 0.00   | 23.00   | 0.00   | 8.34    | 8.12   | -1.25    |
| Variable 9  | 20.00  | 213.00  | 22.00  | 87.30   | 67.94  | -1.82    |
| Variable 10 | 68.00  | 157.00  | 76.00  | 106.57  | 22.95  | 0.07     |
| Variable 11 | 20.00  | 149.00  | 22.00  | 61.00   | 41.83  | -1.57    |
| Variable 12 | 0.00   | 3.00    | 0.00   | 1.14    | 1.25   | -1.66    |
| Variable 13 | 0.00   | 4.00    | 0.00   | 1.16    | 1.27   | -1.20    |
| Variable 14 | 0.00   | 103.00  | 0.00   | 29.05   | 26.24  | 0.34     |

\*Variable 1: Time spending; Variable 2: Time of submitted recording; Variable 3: Number of comprehensive tasks completed; Variable 4: The number of single tasks completed with pronunciation accuracy; Variable 5: Number of fluency tasks completed; Variable 6: The number of lexical and grammar tasks completed; Variable 7: Number of tasks completed in practising coherence of oral English; Variable 8: Number of tasks completed in practising; Variable 9: The number of times you get instant feedback; Variable 10: Number of repeated tasks; Variable 11: Number of policy adjustments; Variable 12: The number of times you rate other students; Variable 13: The number of times you were graded by other students; Variable 14: number of peer interactions.

The above descriptive analysis highlights significant variation in students' independent learning. There are notable differences in aspects such as time spent on learning, task completion, feedback received, and strategy adjustments, which underscores the importance of personalised learning in the context of independent study. While some students excel in routine tasks and engage well with feedback mechanisms, others struggle with more challenging tasks and interactive activities. To enhance overall learning outcomes, it may be necessary to provide more tailored support and incentives that address these differences, ensuring that each student can achieve optimal results in their independent learning journey.

### 5.2.2 Analysis of Differences between the EG and CG

As seen from Table 25, the independent samples *t*-test results comparing the experimental and control groups across 14 variables reveal significant differences in several key learning indicators.

#### ***Significant improvement in learning engagement and task participation***

The *t*-test results for time invested ( $t = 7.70$ ,  $p < 0.05$ ) and total recording duration ( $t = 3.46$ ,  $p < 0.05$ ) demonstrate that the experimental group showed significantly higher levels of engagement in terms of time spent and participation in speaking practice compared to the control group. This indicates that the intervention effectively enhanced

students' learning engagement and time management skills, encouraging them to spend more time on autonomous learning and task completion.

### ***Differences in task completion***

In terms of task completion, the experimental group significantly outperformed the control group in fluency tasks ( $t = 14.80$ ,  $p < 0.05$ ), coherence tasks ( $t = 7.76$ ,  $p < 0.005$ ), and appropriateness tasks ( $t = 9.23$ ,  $p < 0.05$ ). This suggests that the intervention significantly improved students' performance in these areas, particularly with regard to fluency, coherence, and appropriateness in language tasks. However, in integrated tasks ( $t = -3.75$ ,  $p < 0.05$ ) and pronunciation accuracy tasks ( $t = -2.69$ ,  $p < 0.05$ ), the control group performed better, implying that the control group may have excelled in these tasks, or that the task design may have better suited the control group's focus.

### ***Significant increase in feedback and strategy adjustment***

The experimental group showed significantly higher instances of instant feedback ( $t = 29.21$ ,  $p < 0.05$ ) and strategy adjustments ( $t = 19.67$ ,  $p < 0.05$ ) than the control group. This suggests that the intervention effectively increased the frequency of feedback received by students, prompting them to adjust and optimise their learning strategies more frequently. These findings reflect the intervention's success in strengthening the feedback mechanism and fostering reflective behaviour in the autonomous learning process.

### ***Significant increase in peer interaction***

The experimental group had significantly more peer interactions ( $t = 9.28$ ,  $p < 0.05$ ) than the control group, indicating that the intervention greatly enhanced the frequency of peer interaction and collaborative learning. This demonstrates that the intervention not only improved individual learning outcomes but also fostered social interaction throughout the learning process.

### ***No significant difference in peer assessment tasks***

There was no significant difference between the experimental and control groups in the number of times students assigned scores to their peers ( $t = 0.00$ ,  $p > 0.05$ ) or received scores from others ( $t = 0.12$ ,  $p > 0.05$ ). Both groups showed relatively low levels of engagement in these peer assessment tasks, suggesting that the intervention did not have a significant effect in this area.

### ***Minimal differences in task repetition***

The *t*-test result for task repetition ( $t = -2.03$ ,  $p < 0.05$ ) approached significance, indicating a small difference between the experimental and control groups. The experimental group appeared to be more inclined towards repeating tasks, which suggests that while the intervention had a limited effect in encouraging task repetition and self-improvement, there was still some impact.

Overall, these *t*-test results demonstrate that the intervention had a significant positive effect on several learning outcomes, particularly in terms of learning engagement, task performance, and feedback strategy adjustment. However, the control group performed better in certain tasks, possibly due to differences in task design or individual student characteristics.

*Table 25 Independent Sample T-test of the 14 Variables in Students' Engagement in OEAL & FA*

|             | <i>t</i> | <i>p</i> |
|-------------|----------|----------|
| Variable 1  | 7.70     | 0.00     |
| Variable 2  | 3.46     | 0.00     |
| Variable 3  | -3.75    | 0.00     |
| Variable 4  | -2.69    | 0.01     |
| Variable 5  | 14.80    | 0.00     |
| Variable 6  | -1.23    | 0.23     |
| Variable 7  | 7.76     | 0.00     |
| Variable 8  | 9.23     | 0.00     |
| Variable 9  | 29.21    | 0.00     |
| Variable 10 | -2.03    | 0.05     |
| Variable 11 | 19.67    | 0.00     |
| Variable 12 | 0.00     | 1.00     |
| Variable 13 | 0.12     | 0.91     |
| Variable 14 | 9.28     | 0.00     |

\* Significant ( $p < 0.05$ )

On the whole, experimental intervention has a positive and significant impact on students' autonomous learning behaviour, especially in time management, task completion, feedback acquisition, strategy adjustment, and peer communication. The experimental group's performance is significantly better than that of the control group. The experimental intervention prompted students to participate more actively in autonomous learning tasks and to show significant improvements in key language skills such as fluency and coherence. Although the control group performed better on some tasks (such as integrated tasks and pronunciation accuracy), overall the experimental intervention effectively improved the quality of students' autonomous learning by strengthening the feedback mechanism, promoting the adjustment of learning strategies, and enhancing peer

interaction. These results provide important empirical support for further optimisation of instructional intervention design for autonomous learning.

### 5.3 Students' Progress in Scores of the Five Tests of Oral English

#### 5.3.1 Descriptive Analysis of the Five Tests

The researcher conducted an analysis of the mean and standard deviation of the average scores from the three assessors (FiF, Martin, and the researcher) for the five oral English tests in both the experimental group and the control group (see Table 26).

*Table 26 Descriptive Analysis of Students' Average Scores on the Five Oral English Tests*

|    | Test1_Avg |      | Test2_Avg |      | Test3_Avg |      | Test4_Avg |       | Test5_Avg |      |
|----|-----------|------|-----------|------|-----------|------|-----------|-------|-----------|------|
|    | Mean      | SD   | Mean      | SD   | Mean      | SD   | Mean      | SD    | Mean      | SD   |
| EG | 83.55     | 5.42 | 82.23     | 4.39 | 86.42     | 5.79 | 85.89     | 12.01 | 89.15     | 4.91 |
| CG | 82.98     | 3.73 | 82.62     | 4.21 | 84        | 3.62 | 85.73     | 3.82  | 86.79     | 2.21 |

In Test 1, the mean for the experimental group was 83.55, with a standard deviation of 5.42, while the control group had a mean of 82.98 and a standard deviation of 3.73. This indicates that, in the first test, the mean scores of the two groups were quite similar, suggesting that their initial performance was comparable. However, the larger standard deviation in the experimental group suggests a more dispersed performance among its students, with some potentially excelling or underperforming. In contrast, the control group had a smaller standard deviation, indicating that most students' scores were more clustered around the mean, with less variation.

In Test 2, the mean scores for the experimental and control groups were very close (82.23 and 82.62, respectively), indicating that the students' performance at this stage remained highly similar. The standard deviations were also very close (EG: SD = 4.39, CG: SD = 4.21), suggesting that individual differences within each group were comparable, and no significant trend of score dispersion or concentration was observed.

Beginning from Test 3, the mean score of the experimental group (mean = 86.42) increased significantly, surpassing the control group (mean = 84.00) by over two points. This indicates that students in the experimental group demonstrated greater progress at this stage, which could be attributed to the use of the FiF, potentially enhancing their oral English skills. Additionally, the standard deviation for the experimental group also increased markedly (SD = 5.79), suggesting that the performance gap between students widened, with some making substantial progress while others may have exhibited more

average results. In contrast, the control group's standard deviation remained relatively low (SD = 3.62), indicating more consistent performance with less individual variation.

In Test 4, the experimental group's mean (mean = 88.32) continued to rise, again surpassing the control group's mean (mean = 85.73) by more than two points, showing sustained improvement in overall performance for the experimental group. Meanwhile, the standard deviation for the experimental group decreased to 4.68, indicating that the differences between students' scores had narrowed, with individual performance becoming more aligned. Although the control group's mean also improved, its standard deviation remained relatively small (SD = 3.82), indicating stable performance and minimal individual variation.

In the final test (Test 5), the experimental group's mean reached its highest point at 89.15, significantly higher than the control group's mean of 86.79. This suggests that the experimental group made substantial progress, with the learning effects becoming more pronounced in the later stages of the study, especially with the use of the FiF. At the same time, the experimental group's standard deviation remained relatively high (SD = 4.91), indicating that while some students performed well above average, others showed more moderate improvement, leading to considerable individual variation. Conversely, the control group's standard deviation dropped significantly to 2.21, showing that most students' scores were closely grouped, with very little individual variation and relatively uniform progress across the group.

From the mean scores of both groups, it is evident that from Test 3 onwards, the experimental group consistently outperformed the control group, with the gap between the two groups widening over time. This suggests that the experimental group experienced superior learning outcomes compared to the control group, particularly in terms of improving oral English proficiency through the use of the FiF.

The analysis of the standard deviations for both groups reveals that the experimental group exhibited more fluctuation in its standard deviation throughout the testing process, particularly in Test 3 and Test 5, where the larger standard deviation indicates that while some students made remarkable progress, others advanced more slowly. In contrast, the control group maintained a consistently lower standard deviation,

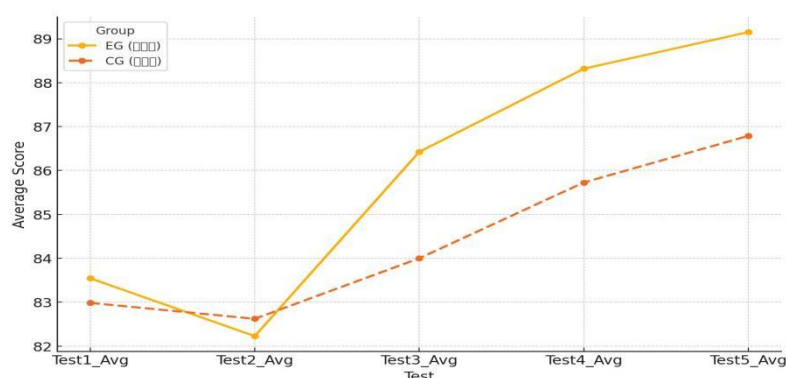
indicating that students' performance was more concentrated, with less individual variation, but the overall progress was less pronounced compared to the experimental group.

These findings suggest that the use of WBFAT to support autonomous learning and formative assessment of English oral English proficiency had varied effects on different students in the experimental group. In contrast, the traditional learning and assessment methods used by the control group appeared to have more uniform effects across all students but without leading to significant overall progress.

### 5.3.2 The trend of scores of two groups of students in five oral tests

As shown in Figure 11, the experimental group exhibited significant growth in performance, particularly in the later tests, where students consistently outperformed those in the control group. This indicates that the FiF had a positive impact on the development of oral English proficiency. In contrast, the control group demonstrated more stable but less pronounced progress, suggesting that the traditional learning methods did not result in the same level of improvement as seen in the experimental group.

*Figure 11 The Trend of Student's Average Scores of Oral English Tests by Group*



From the overall trend, it is evident that the gap between the experimental and control groups widened as the tests progressed, highlighting the increasingly apparent advantages of the learning approach used by the experimental group, particularly in the later stages of the study.

## 5.4 Students' Progress in Linguistic Features From Their Audio Transcripts

### 5.4.1 Descriptive Analysis of Students' Five Audio Transcripts

After the analysis of linguistic characteristics of students' audio transcripts in the five oral English examinations via Coh-Metrix 3.0, the research employed SPSS 26 to analyse the values of the 20 indicators, as shown in Table 27.

Table 27 Descriptive Analysis of the Values of the 20 Indicators of Linguistic Characteristics in Students' Five Audio Transcripts

|         | Test1  |       |        |       | Test2  |       |        |       | Test3  |       |        |       | Test4  |       |        |       | Test5  |       |        |       |
|---------|--------|-------|--------|-------|--------|-------|--------|-------|--------|-------|--------|-------|--------|-------|--------|-------|--------|-------|--------|-------|
|         | EG     | EG    | CG     | CG    | EG     | EG    | CG     | CG    | EG     | EG    | CG     | CG    | EG     | EG    | CG     | CG    | EG     | EG    | CG     | CG    |
|         | Mean   | Std.  | Mean   | Std.  | Mean   | Std.  | Mean   | Std.  | Mean   | Std.  | Mean   | Std.  | Mean   | Std.  | Mean   | Std.  | Mean   | Std.  | Mean   | Std.  |
| DESWC   | 176.82 | 45.15 | 163.41 | 22.19 | 172.55 | 22.24 | 170.68 | 16.33 | 244.73 | 26.68 | 187.23 | 24.75 | 225.86 | 26.67 | 153.95 | 22.50 | 213.77 | 21.01 | 172.59 | 23.03 |
| DESSL   | 9.65   | 3.52  | 9.67   | 3.82  | 9.24   | 3.15  | 10.24  | 3.81  | 14.55  | 2.79  | 11.51  | 3.92  | 13.23  | 3.43  | 11.95  | 4.51  | 17.08  | 1.98  | 14.77  | 2.75  |
| PCNARp  | 30.04  | 11.07 | 34.72  | 7.81  | 10.02  | 8.48  | 10.40  | 7.10  | 33.00  | 8.76  | 28.19  | 15.52 | 23.40  | 10.55 | 15.75  | 10.36 | 65.77  | 12.77 | 54.33  | 18.55 |
| PCSYNp  | 59.55  | 23.93 | 65.01  | 22.17 | 81.61  | 13.31 | 77.79  | 19.42 | 71.88  | 16.52 | 59.20  | 18.08 | 76.56  | 21.00 | 55.12  | 27.70 | 93.72  | 12.26 | 45.07  | 22.88 |
| PCCNCp  | 87.44  | 14.48 | 89.29  | 11.06 | 68.59  | 17.63 | 68.73  | 20.23 | 82.33  | 15.09 | 75.99  | 14.37 | 86.52  | 12.12 | 82.00  | 19.44 | 19.72  | 15.47 | 26.25  | 18.25 |
| PCREFp  | 85.62  | 16.52 | 88.25  | 20.49 | 64.33  | 17.07 | 70.03  | 19.07 | 64.86  | 23.87 | 66.49  | 22.98 | 68.25  | 27.56 | 70.71  | 33.66 | 30.91  | 28.30 | 51.91  | 37.36 |
| PCDCp   | 28.26  | 14.31 | 37.79  | 20.29 | 76.91  | 15.05 | 71.31  | 12.00 | 89.08  | 12.57 | 84.98  | 19.60 | 55.66  | 14.40 | 55.88  | 22.14 | 80.59  | 18.63 | 85.44  | 15.32 |
| CRFAOa  | 0.17   | 0.11  | 0.18   | 0.11  | 0.10   | 0.06  | 0.14   | 0.10  | 0.17   | 0.09  | 0.19   | 0.10  | 0.21   | 0.14  | 0.24   | 0.15  | 0.36   | 0.11  | 0.36   | 0.14  |
| CRFCWO1 | 0.06   | 0.05  | 0.06   | 0.03  | 0.04   | 0.02  | 0.05   | 0.03  | 0.03   | 0.03  | 0.03   | 0.03  | 0.05   | 0.04  | 0.06   | 0.03  | 0.10   | 0.03  | 0.10   | 0.03  |
| LSASS1  | 0.10   | 0.06  | 0.11   | 0.06  | 0.08   | 0.04  | 0.11   | 0.07  | 0.07   | 0.03  | 0.08   | 0.04  | 0.15   | 0.07  | 0.17   | 0.08  | 0.17   | 0.05  | 0.18   | 0.05  |
| LDTTRc  | 0.76   | 0.07  | 0.77   | 0.08  | 0.86   | 0.04  | 0.86   | 0.05  | 0.82   | 0.04  | 0.82   | 0.04  | 0.81   | 0.05  | 0.79   | 0.06  | 0.80   | 0.04  | 0.80   | 0.05  |
| LDTTRa  | 0.59   | 0.07  | 0.58   | 0.06  | 0.67   | 0.06  | 0.67   | 0.05  | 0.65   | 0.04  | 0.65   | 0.04  | 0.63   | 0.05  | 0.63   | 0.05  | 0.62   | 0.03  | 0.63   | 0.04  |
| SMTEMP  | 0.88   | 0.10  | 0.86   | 0.07  | 0.84   | 0.09  | 0.84   | 0.08  | 0.79   | 0.11  | 0.77   | 0.15  | 0.91   | 0.05  | 0.89   | 0.07  | 0.80   | 0.07  | 0.79   | 0.09  |
| WRDFAMc | 576.76 | 3.74  | 579.07 | 6.30  | 567.21 | 4.79  | 567.20 | 5.72  | 596.35 | 4.23  | 599.05 | 3.51  | 595.09 | 4.02  | 561.59 | 5.09  | 617.13 | 2.75  | 575.50 | 3.71  |
| WRDCNCc | 396.23 | 11.37 | 400.49 | 16.34 | 363.52 | 15.22 | 363.10 | 11.45 | 370.70 | 9.39  | 370.46 | 6.50  | 415.77 | 11.40 | 402.62 | 16.24 | 408.09 | 7.55  | 340.24 | 7.40  |
| WRDHYPn | 6.53   | 0.19  | 6.46   | 0.18  | 6.77   | 0.28  | 6.71   | 0.30  | 7.06   | 0.21  | 7.01   | 0.21  | 6.52   | 0.26  | 6.55   | 0.22  | 6.82   | 0.21  | 6.81   | 0.29  |
| WRDHYPv | 1.76   | 0.18  | 1.77   | 0.16  | 1.43   | 0.11  | 1.45   | 0.11  | 1.50   | 0.13  | 1.48   | 0.15  | 1.77   | 0.13  | 1.78   | 0.14  | 1.51   | 0.10  | 1.50   | 0.10  |
| RDFRE   | 66.07  | 7.47  | 68.12  | 8.47  | 53.22  | 8.77  | 51.61  | 7.77  | 75.07  | 4.86  | 75.37  | 5.47  | 61.79  | 6.74  | 59.14  | 6.72  | 64.29  | 3.04  | 66.99  | 6.00  |
| RDFKGL  | 6.44   | 1.79  | 6.16   | 1.97  | 8.13   | 1.84  | 8.61   | 1.94  | 5.56   | 1.24  | 5.61   | 1.67  | 7.35   | 1.74  | 7.98   | 1.99  | 8.16   | 0.77  | 7.56   | 1.42  |
| RDL2    | 16.23  | 4.14  | 18.51  | 4.18  | 11.54  | 4.19  | 12.12  | 4.32  | 16.06  | 3.02  | 16.83  | 3.14  | 6.50   | 3.85  | 6.84   | 3.96  | 18.29  | 2.38  | 19.44  | 2.39  |

### ***EG achieved a faster improvement in vocabulary and sentence length***

The experimental group exhibited a faster rate of improvement in vocabulary and sentence length than the control group. In the first two tests, mean scores for DESWC and DESSL revealed no notable difference between the two groups (EG: mean<sub>1,2</sub> = 176.82, 172.55; mean<sub>1,2</sub> = 9.65, 9.24; CG: mean<sub>1,2</sub> = 163.41, 170.68; mean<sub>1,2</sub> = 9.67, 10.24). However, from the third test onwards, the experimental group demonstrated a marked increase, with both the mean scores and growth rate exceeding those of the control group (EG: mean<sub>3,4,5</sub> = 244.73, 225.86, 213.77; mean<sub>3,4,5</sub> = 14.55, 13.23, 17.08; CG: mean<sub>3,4,5</sub> = 187.23, 153.95, 172.59; mean<sub>3,4,5</sub> = 11.51, 11.95, 14.77). This suggested that the experimental group's oral expressive ability and complexity began to improve from the third test.

### ***No significant differences between groups in text readability indicators except for PCSYNp***

Readability indicators included PCNARp, PCSYNp, PCREFp, and PCDCp. In the five tests, both EG and CG showed considerable fluctuations in PCNARp, with the experimental group displaying more pronounced variation between tests. Although the experimental group's PCNARp reached its peak (mean = 65.77) in the fifth test, further differentiating it from the control group (mean = 54.33), their performance was otherwise similar. PCSYNp showed a continuous increase in the experimental group (mean<sub>1</sub> = 59.55, mean<sub>5</sub> = 93.72), diverging significantly from the control group (mean<sub>1</sub> = 65.01, mean<sub>5</sub> = 45.07), indicating consistent improvement in syntactic complexity. Initial PCREFp and PCDCp mean scores for the experimental group were notably higher, yet by the fifth test, they had decreased significantly, while the control group's scores rose, eventually surpassing the experimental group.

### ***The experimental group excelled in WRDFAMc but declined in SMTEMP, while both groups showed low and stable coherence scores.***

Indicators of coherence and semantic relevance assessed students' consistency and logical flow in oral expression. Both groups maintained low average scores on CRFAOa, CRFCWO1, and LSASS1 (typically between 0 and 0.2), with minor fluctuations across tests. SMTEMP mean scores were significantly higher in the experimental group but declined progressively during the tests, while the control group's scores remained

relatively stable. In WRDFAMc, the experimental group consistently outperformed the control group, showing steady increases across tests and peaking in the final test, suggesting enhanced performance in specific domains.

***The experimental group excelled in lexical diversity but saw declines in PCCNCp, with limited impact on WRDHYPn across groups.***

Lexical concreteness (PCCNCp, WRDCNCc) indicated whether vocabulary was concrete and easily understood, while lexical diversity (LDTTRc, LDTTRa) reflected the range of vocabulary used. The experimental group consistently scored higher than the control group across all tests for LDTTRc and LDTTRa, achieving peaks followed by slight declines, demonstrating an ongoing advantage in vocabulary diversity and complexity. Mean WRDCNCc scores in the experimental group remained high throughout, while the control group's mean scores were significantly lower. This suggested that the experimental tool may have aided in improving complex vocabulary use. PCCNCp means in the experimental group started high (mean1 = 87.44) but dropped in the final test (mean5 = 19.72), aligning with the control group's level (mean5 = 26.25). Mean WRDHYPn scores remained relatively stable across tests for both groups, showing limited impact of the intervention on syntactic complexity and lexical depth.

***No significant differences for indicators of readability and comprehension difficulty between groups***

RDFRE (Flesch Reading Ease) and RDFKGL (Flesch-Kincaid Grade Level) indicated text readability, while RDL2 reflected second language readability. Experimental group RDFRE mean scores fluctuated between 53.22 and 75.07, generally maintaining high readability levels. The control group's means closely aligned initially but varied, peaking at 75.37 in the third test and ending at 66.99 in the final test, slightly below the experimental group (64.29), indicating comparable readability levels. RDFKGL means in the experimental group ranged from 5.56 to 8.16, slightly higher than those of the control group, suggesting marginally higher readability difficulty. While the control group had lower RDFKGL means in the initial tests, a marked increase occurred in the final test, closing the gap with the experimental group. Experimental group RDL2 means varied significantly (6.50 to 18.29), while the control group maintained a low level (3-4) until the

fifth test, where it rose sharply to 19.44, surpassing the experimental group's 18.29, indicating notable late-stage progress in linguistic complexity.

#### 5.4.2 *T*-test of 20 Indicators' Mean of Language Features in the Five Tests

As shown in Table 28, the independent sample *t*-tests on the 20 linguistic feature indicators across five oral tests in both groups show that certain variables, such as DESWC, DESSL, PCSYNp, and WRDFAMc, began to show differences between the two groups from the third test onwards. Other variables, such as WRDCNCc, displayed significant differences starting from the fourth test, while variables like PCNARp, PCREFp, RDFRE, and RDFKGL only revealed significant differences in the fifth test. Meanwhile, some variables, including CRFAOa, CRFCWO1, LSASS1, LDTTRc, LDTTRa, SMTEMP, WRDHYPn, and WRDHYPv, consistently showed no significant differences across all five tests.

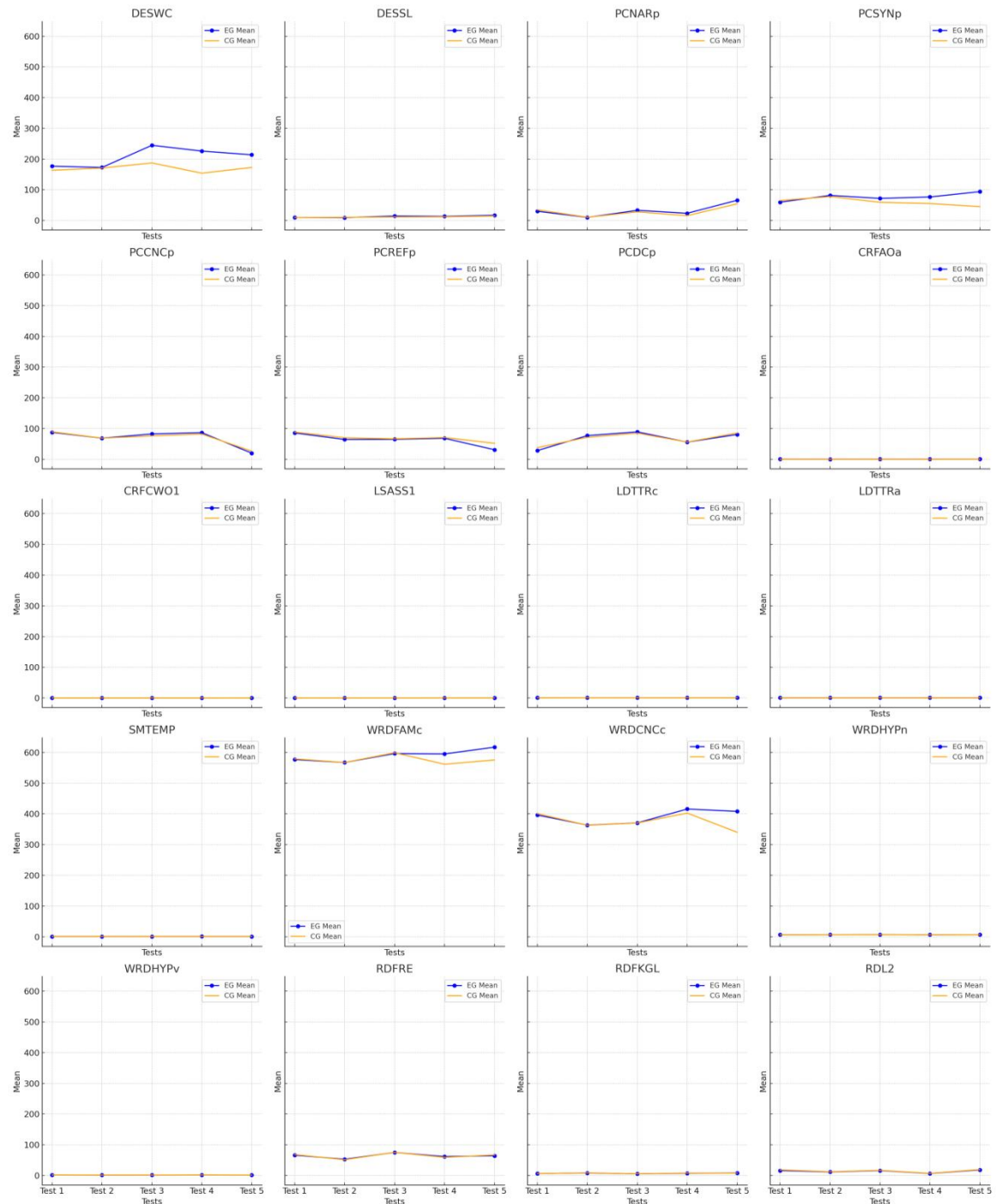
Table 28 *T*-tests of the 20 Linguistic Characteristic Indicators' Mean in Students' Five Audio Transcripts

|         | Test 1   |          | Test 2   |          | Test 3   |          | Test 4   |          | Test 5   |          |
|---------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
|         | <i>t</i> | <i>p</i> | <i>t</i> | <i>p</i> | <i>t</i> | <i>p</i> | <i>t</i> | <i>p</i> | <i>t</i> | <i>p</i> |
| DESWC   | 1.46     | 0.15     | 0.37     | 0.71     | 8.65     | 0.00     | 11.29    | 0.00     | 7.24     | 0.00     |
| DESSL   | -0.02    | 0.98     | -1.11    | 0.27     | 3.46     | 0.00     | 1.24     | 0.22     | 3.74     | 0.00     |
| PCNARp  | -1.89    | 0.06     | -0.19    | 0.85     | 1.48     | 0.14     | 2.84     | 0.01     | 2.78     | 0.01     |
| PCSYNp  | -0.92    | 0.36     | 0.89     | 0.38     | 2.84     | 0.01     | 3.38     | 0.00     | 10.26    | 0.00     |
| PCCNCp  | -0.56    | 0.58     | -0.03    | 0.98     | 1.67     | 0.10     | 1.08     | 0.28     | -1.50    | 0.14     |
| PCREFp  | -0.55    | 0.59     | -1.22    | 0.23     | -0.27    | 0.79     | -0.31    | 0.76     | -2.45    | 0.02     |
| PCDCp   | -2.10    | 0.04     | 1.59     | 0.12     | 0.96     | 0.34     | -0.05    | 0.96     | -1.10    | 0.27     |
| CRFAOa  | -0.46    | 0.65     | -1.85    | 0.07     | -0.91    | 0.37     | -0.98    | 0.33     | -0.21    | 0.83     |
| CRFCWO1 | -0.29    | 0.77     | -1.34    | 0.18     | -0.62    | 0.54     | -1.01    | 0.31     | 0.17     | 0.87     |
| LSASS1  | -0.23    | 0.82     | -1.84    | 0.07     | -0.47    | 0.64     | -1.10    | 0.27     | -0.47    | 0.64     |
| LDTTRc  | -0.20    | 0.84     | 0.72     | 0.48     | 0.51     | 0.61     | 0.97     | 0.33     | 0.00     | 1.00     |
| LDTTRa  | 0.30     | 0.76     | 0.21     | 0.84     | -0.31    | 0.76     | 0.11     | 0.92     | -1.28    | 0.21     |
| SMTEMP  | 1.01     | 0.32     | -0.29    | 0.77     | 0.57     | 0.57     | 0.81     | 0.42     | 0.56     | 0.58     |
| WRDFAMc | -1.72    | 0.09     | 0.01     | 0.99     | -2.69    | 0.01     | 28.27    | 0.00     | 49.35    | 0.00     |
| WRDCNCc | -1.17    | 0.25     | 0.12     | 0.90     | 0.12     | 0.91     | 3.63     | 0.00     | 35.16    | 0.00     |
| WRDHYPn | 1.30     | 0.20     | 0.83     | 0.41     | 0.95     | 0.35     | -0.48    | 0.63     | 0.13     | 0.89     |
| WRDHYPv | -0.16    | 0.87     | -0.81    | 0.42     | 0.68     | 0.50     | -0.07    | 0.95     | 0.20     | 0.84     |
| RDFRE   | -0.99    | 0.32     | 0.75     | 0.45     | -0.22    | 0.83     | 1.53     | 0.13     | -2.19    | 0.03     |
| RDFKGL  | 0.57     | 0.57     | -0.97    | 0.34     | -0.12    | 0.91     | -1.30    | 0.20     | 2.03     | 0.05     |
| RDL2    | -2.12    | 0.04     | -0.53    | 0.60     | -0.97    | 0.34     | -0.34    | 0.73     | -1.87    | 0.07     |

The development trend of students' oral English proficiency according to the 20 indicators of linguistic features across the five tests is shown in Figure 12.

Figure 12 Trends of the Values of the 20 Indicators of Linguistic Characteristics in Students' Five Audio

### Transcripts



***The variables with significant differences are mainly concentrated in the indices of text readability and readability and some indicators of basic lexical ability***

Table 28 demonstrates that from the third test onwards, DESWC ( $t_{3,4,5} = 8.65, 11.29, 7.24; p_{3,4,5} = 0.00$ ) and PCSYNp consistently presented significant p-values ( $t_{3,4,5} = 2.84, 3.38, 10.26; p_{3,4,5} = 0.01, 0.00, 0.00$ ). DESSL achieved significance in both the third ( $t_3 = 3.46; p_3 = 0.00$ ) and fifth tests ( $t_5 = 3.74; p_5 = 0.00$ ). Furthermore, PCNARp

showed significant differences beginning from the fourth test ( $t_{4,5} = 2.84, 2.78$ ;  $p_{4,5} = 0.01$ ). These results indicate that the experimental group demonstrated significant improvements in text length, the use of simple, familiar structures, and narrative abilities in spoken expression, particularly from the latter stages of the experiment (i.e., the third or fourth tests) onwards.

From the third test, WRDFAMc displayed significant p-values ( $t_{3,4,5} = -2.69, 28.27, 49.35$ ;  $p_{3,4,5} = 0.01, 0.00, 0.00$ ), indicating a sustained, significant improvement in the use of familiar vocabulary. WRDCNCc, starting from the fourth test, also became significant ( $t_{4,5} = 3.63, 35.16$ ;  $p_{4,5} = 0.01, 0.00$ ), suggesting a notable enhancement in lexical concreteness.

Moreover, during the final (the fifth) test, RDFRE ( $t_5 = -2.19$ ;  $p_5 = 0.03$ ), RDFKGL ( $t_5 = 2.03$ ;  $p_5 = 0.05$ ), and PCREFp ( $t_5 = -2.45$ ;  $p_5 = 0.02$ ) exhibited significant p values, highlighting emerging differences in reading ease, readability grade level, and referential cohesion between the groups.

***Variables with no significant differences focus on the indicators of cohesion, latent semantics, and advanced lexical ability***

Variables including PCCNCp, CRFAOa, CRFCWO1, LSASS1, LDTTRc, LDTTRa, SMTEMP, WRDHYPn, WRDHYPv, PCDCp, and RDL2 did not reach significance at any point across the five tests ( $p > 0.05$  for all tests). This indicates that students in the experimental group made limited progress in these areas, with no significant differences observed between the two groups.

## **5.5 Student's Performance of Oral English in Class Observation**

### **5.5.1 EG Showed Higher Classroom Engagement, Oral Accuracy, Fluency, and Coherence than the CG, with Notable Improvements after Using FiF**

From the analysis of node coding frequency in Table 29, we can observe that the experimental group recorded 15 instances of classroom participation, slightly higher than the control group's 12 instances. Although the two groups displayed similar levels of classroom participation, the experimental group exhibited marginally higher engagement, suggesting that they were more interactive during lessons. This was particularly evident after the use of the FiF, where students in the experimental group became more proactive in participating in classroom tasks.

*Table 29 The Node Coding Frequency of the Six Class Observations*

|    | <b>Classroom engagement coding</b> | <b>Oral accuracy coding</b> | <b>Oral fluency Coding</b> | <b>Oral Coherence and integrity coding</b> |
|----|------------------------------------|-----------------------------|----------------------------|--------------------------------------------|
| EG | 15                                 | 20                          | 18                         | 17                                         |
| CG | 12                                 | 15                          | 14                         | 12                                         |

For the accuracy of oral English, the experimental group recorded 20 instances, significantly higher than the control group's 15. The experimental group performed better in oral accuracy, including pronunciation, vocabulary usage, and grammatical correctness. This difference was especially pronounced in the later tests, where the FiF may have contributed more effectively to improving the students' accuracy.

In the fluency, the experimental group recorded 18 instances, compared to 14 for the control group. The experimental group demonstrated greater improvement in reducing pauses and meaningless repetition and increasing speech speed. Although the control group also showed some improvement in fluency, the degree of progress was less substantial.

For the oral coherence and completeness, the experimental group recorded 17 instances, while the control group recorded 12. Students in the experimental group showed enhanced coherence and completeness in expressing complex ideas, with stronger logical connections between sentences and more complete transmission of information. On the other hand, control group students continued to face difficulties when conveying more complex content.

### **5.5.2 EG Showed Continuous and Significant Improvements in Oral English Proficiency Over Time With the Use of FiF**

As can be seen from Table 30, from Observation 1 to Observation 6, the coding frequency of the experimental group gradually increased, especially in Observation 5 and Observation 6, where the number of coding times increased significantly. This showed that as the experiment progressed, students in the experimental group showed continuous improvement in oral accuracy, fluency, and coherence, especially in the later tests, where students' performance was more stable and significant. However, the coding frequency of the control group was relatively low, and the increase rate between Observation 1 and Observation 6 was not as significant as that of the experimental group. Although there is a certain improvement from Observation 1 to Observation 4, the overall progress is relatively slow. Especially in Observation 5 and Observation 6, the number of coding times of the experimental group reached 8 and 9, while the number of coding times of the control

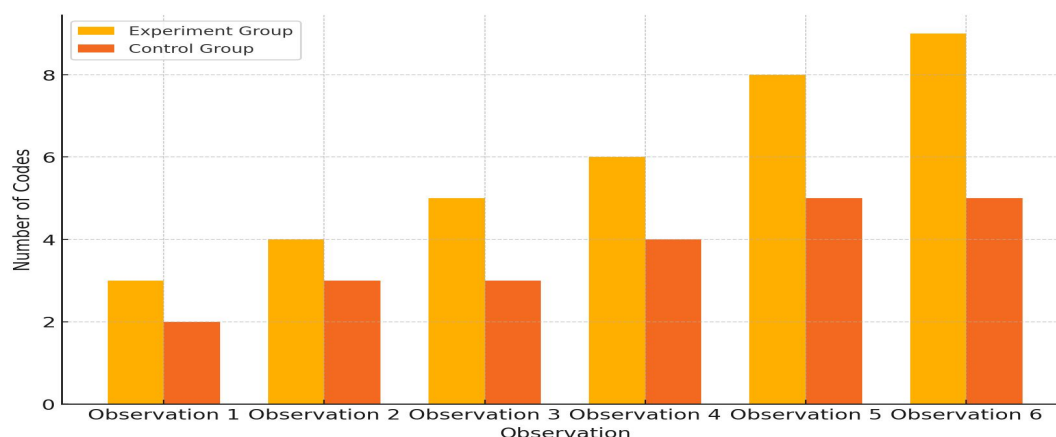
group stayed at 5, indicating that the progress of oral performance of the control group had stalled and failed to achieve continuous improvement like the experimental group.

*Table 30 Query Results of The Matrix Coding of the Six Class Observations*

|    | Observation<br>1 | Observation<br>2 | Observation<br>3 | Observation<br>4 | Observation<br>5 | Observation<br>6 |
|----|------------------|------------------|------------------|------------------|------------------|------------------|
| EG | 3                | 4                | 5                | 6                | 8                | 9                |
| CG | 2                | 3                | 3                | 4                | 5                | 5                |

In Figure 13, the matrix chart illustrates the coding frequencies for the experimental group and control group across Observation 1 to Observation 6, reflecting the trends in the students' classroom performance, such as oral accuracy, fluency, and other related aspects. The performance growth curve for the experimental group is noticeably steeper, particularly in the later observations, where students significantly improved their oral abilities through the use of the FiF. In contrast, the growth curve for the control group is relatively flat, indicating that traditional teaching methods did not yield the same level of progress as seen in the experimental group.

*Figure 13 The Matrix chart: Number of Codes by Group in Class Observation*



The above findings indicate that the experimental group demonstrated steady improvements in classroom participation, oral accuracy, fluency, and coherence, with particularly notable progress in the later observations. This trend is closely linked to the use of the FiF, suggesting that the tool had a significant impact on enhancing the students' oral English proficiency. Although the control group also made some progress, it was relatively slower, especially in Observation 5 and Observation 6, where the improvements were less pronounced compared to the experimental group. This suggests that traditional learning methods have limited effectiveness in improving oral English proficiency, particularly in more complex language expression and fluency.

## 5.6 Themes of the Instructor Interviews

The researcher collated the four interviews with the instructor, Martin and then used NVivo 11 to summarise the themes as follows.

#### **5.6.1 WBFAT Improved Students' Participation and Initiative in Oral English in Class**

Students in the experimental group demonstrated a significant improvement in their participation and initiative in oral English in classes. This change began to emerge approximately eight weeks after starting WBFAT. Martin observed that these students gradually shifted from passively participating to actively expressing their views in English. They started engaging in group discussions and interactions, going beyond merely providing brief responses to teacher prompts ( Martin, Interview 1). By the 11th week, the experimental group's initiative had further developed. Not only were they participating more frequently, but they also began taking leadership roles in group discussions and encouraging their peers to contribute. Martin attributed this improvement to WBFAT's consistent provision of language practice and feedback, which helped students build confidence and establish a habit of actively using English during classroom interactions (Martin, Interview 2). In contrast, the control group showed relatively modest progress. Although some students began participating due to an improved classroom atmosphere, many remained reliant on teacher prompts and displayed limited initiative, indicating a persistent gap in engagement between the two groups (Martin, Interview 3).

#### **5.6.2 WBFAT Enhanced Confidence in Oral English Expression Among Higher Vocational College Students**

The experimental group began to exhibit a marked increase in confidence around the 11th week of WBFAT usage. Martin noted that these students gradually became more fluent and composed when expressing their views, even in challenging tasks. Notably, they continued speaking despite occasional errors, rather than hesitating or stopping out of fear of making mistakes (Martin, Interview 2). By the 17th week, this confidence had significantly improved. Students not only appeared more self-assured during classroom activities but also actively volunteered for speaking tasks and contributed proactively during group discussions. This growth in confidence was closely linked to WBFAT's instant feedback mechanism, which provided a structured and supportive environment for language practice, allowing students to overcome their anxiety about errors through steady

progress (interview 3, Martin, page 9). Conversely, the control group's confidence improvements were limited to simpler tasks, such as reading aloud or responding to direct questions. In more complex situations, such as impromptu speaking tasks, many students in the control group displayed noticeable hesitation, frequent pauses, or repetitive statements, and some even abandoned their responses entirely (Martin, Interview 4).

### **5.6.3 WBFAT Promoted Students' Pronunciation Accuracy in Oral English**

The experimental group showed notable improvements in pronunciation, with progress becoming evident around the 11th week of WBFAT usage. Martin highlighted significant enhancements in students' articulation of vowels, clarity of final consonants, and accuracy in word stress. These improvements were largely attributed to WBFAT's pronunciation correction feature, which allowed students to practise repeatedly and internalise correct pronunciation patterns (Martin, Interview 2). By the 17th week, students in the experimental group demonstrated clearer and more natural pronunciation. They also showed improved sentence-level intonation and rhythm, which Martin observed as enhancing their ability to convey meaning and emotion more effectively during classroom interactions (Martin, Interview 3). In contrast, the control group's progress in pronunciation was less pronounced. Persistent issues such as unclear articulation of final consonants, misplaced word stress, and monotone intonation limited the effectiveness of their spoken communication (Martin, Interview 4).

### **5.6.4 WBFAT Helped Higher Vocational College Students Improve Vocabulary Accuracy, Specificity, and Grammatical Accuracy**

Students in the experimental group demonstrated a clear advantage in vocabulary usage and grammatical accuracy by the 17th week. Martin noted that WBFAT provided diverse language input and topic-based exercises, enabling students to move beyond basic vocabulary to adopt more precise and contextually appropriate expressions during classroom discussions (interview 3, Martin, page 8).

In grammar, experimental group students developed greater proficiency in complex structures such as conditionals and subordinate clauses. Although occasional errors remained, their overall grammatical performance was noticeably more fluid and accurate. Martin emphasised that these improvements were particularly evident in group discussions and responses to open-ended questions. By contrast, the control group's grammar usage

remained largely limited to simple sentences, with higher error rates that hindered their ability to articulate more sophisticated ideas (Martin, Interview 4).

#### **5.6.5 WBFAT Had Positive Impacts on Fluency and Coherence in Oral English**

The experimental group exhibited significant progress in fluency and coherence by the 17th week of WBFAT usage. Martin observed that these students displayed fewer hesitations and repetitions, and their sentence transitions became smoother, resulting in more natural and fluid oral expression (interview 3, Martin, page 10). By the 20th week, students in the experimental group were able to construct more comprehensive and well-structured responses during open-ended discussions. They moved beyond basic statements, analysing topics from multiple perspectives and demonstrating greater logical reasoning and coherence. In contrast, the control group's progress in fluency and coherence remained limited. Many of their responses to complex tasks were fragmented and lacked logical flow, with frequent interruptions and inconsistencies (Martin, Interview 4).

#### **5.6.6 Higher Vocational College Students Still Face Challenges in the Development of Oral English Proficiency**

Despite the experimental group's significant progress in pronunciation, vocabulary specificity, grammatical accuracy, fluency, and coherence, both groups faced common challenges in mastering advanced oral English competence. Firstly, students struggled with constructing clear and logical responses in complex tasks, such as debates or impromptu speeches. Martin observed that many responses lacked depth, with students confined to surface-level observations without systematically exploring causes, implications, or solutions (Martin, Interview 4). Secondly, students demonstrated limited critical thinking and analytical depth. In open-ended tasks, many responses were superficial and lacked nuanced arguments or detailed evaluations. Martin suggested that this issue went beyond linguistic ability, reflecting a need to develop students' cognitive and reasoning skills alongside their language proficiency (Martin, Interview 4).

These findings highlight the importance of targeted interventions to address logical reasoning, critical thinking, and advanced oral communication skills, particularly for higher vocational college students striving for more sophisticated English expression.

### **5.7 Students' Perceptions of WBFAT**

### 5.7.1 Reliability and Validity Test of the Scales Used in the Survey on Students’

#### Perceptions of the Satisfaction and Importance of FiF

##### *Reliability analysis*

It can be seen from Table 31 that the reliability coefficient of the scale is high, so, the survey data is considered to be reliable.

*Table 31 Reliability Test of the Scales in the Survey of Students’ Perceptions of WBFAT*

|              | <b>Cronbach’s Alpha</b> | <b>N of Items</b> |
|--------------|-------------------------|-------------------|
| Satisfaction | 0.849                   | 20                |
| Importance   | 0.903                   | 20                |

##### *Validity analysis*

Table 32 indicates that in the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy, the KMO value for ‘Satisfaction’ is 0.840, and for ‘Importance’, the value of KMO is 0.842. This suggests that the sample has sufficient common variance to justify the use of factor analysis. In Bartlett’s Test of Sphericity, the Approximate Chi-Square values are 1727.922 for ‘Satisfaction’ and 1476.389 for ‘Importance’, with corresponding degrees of freedom (*df*) of 190. The significance level (*Sig.*) is 0.000, indicating that the test result is significant, meaning the null hypothesis (that the correlation matrix is an identity matrix) is rejected.

*Table 32 KMO and Bartlett’s Test of the Scales in the Survey of Students’ Perceptions of WBFAT*

|                                                         | <b>Satisfaction</b> | <b>Importance</b> |
|---------------------------------------------------------|---------------------|-------------------|
| <b>Kaiser-Meyer-Olkin Measure of Sampling Adequacy.</b> | 0.840               | 0.842             |
| <b>Bartlett’s Test of Sphericity</b>                    |                     |                   |
| Approx. Chi-Square                                      | 1727.922            | 1476.389          |
| df                                                      | 190                 | 190               |
| Sig.                                                    | 0.000               | 0.000             |

Therefore, the KMO and Bartlett’s test results both indicate that the scales used in the survey on students’ perceptions of the satisfaction and importance of FiF are highly suitable for factor analysis across both dimensions. This suggests that the relationships between the variables can be effectively explained.

### 5.7.2 T-test Highlights the Need for Improvement in Users’ Satisfaction

As shown in Table 33, the research data reveals discrepancies between users’ evaluations of the importance and satisfaction levels regarding various features of the FiF platform. Firstly, regarding ease of use, users rated its importance (mean = 3.595, SD = 0.740) significantly higher than their satisfaction (mean = 3.248, SD = 0.938), with a

difference of 0.347,  $t = 3.774$ , and  $p < 0.001$ . This indicates that, although the platform is relatively user-friendly, users' expectations surpass their actual experience.

Secondly, in terms of resource support, the importance score (mean = 3.523, SD = 0.903) was significantly higher than the satisfaction score (mean = 2.760, SD = 1.059), with a difference of 0.763,  $t = 7.460$ , and  $p < 0.001$ . This result suggests that the platform's resources may not fully meet users' expectations, particularly regarding the diversity or practicality of the resources provided.

*Table 33 T-tests of the Four Dimensions of the Survey of Students' Perceptions of WBFAT*

|                                                                                                    | Mean  | SD    | Diff. | $t$    | p     |
|----------------------------------------------------------------------------------------------------|-------|-------|-------|--------|-------|
| The importance of the ease of use of WBFAT                                                         | 3.595 | 0.740 | 0.347 | 3.774  | 0.000 |
| The satisfaction with the ease of use of WBFAT                                                     | 3.248 | 0.938 |       |        |       |
| The importance of the resources provided by WBFAT                                                  | 3.523 | 0.903 | 0.763 | 7.460  | 0.000 |
| The satisfaction with the resources provided by WBFAT                                              | 2.760 | 1.059 |       |        |       |
| The importance of WBFAT's support for the development of their oral English proficiency            | 3.886 | 0.741 | 1.162 | 11.640 | 0.000 |
| The satisfaction with the platform's support for the development of their oral English proficiency | 2.724 | 1.009 |       |        |       |
| The importance of WBFAT's evaluation, feedback, and statistical analysis functions                 | 3.808 | 0.730 | 0.665 | 6.871  | 0.000 |
| The satisfaction with WBFAT's evaluation, feedback, and statistical analysis functions             | 3.143 | 0.901 |       |        |       |
| Overall importance                                                                                 | 3.703 | 0.560 | 0.734 | 18.672 | 0.000 |
| Overall satisfaction                                                                               | 2.969 | 0.574 |       |        |       |

For support in developing speaking skills, the importance score was the highest among all features (mean = 3.886, SD = 0.741), while the satisfaction score was the lowest (mean = 2.724, SD = 1.009). The difference of 1.162,  $t = 11.64$ , and  $p < 0.001$  highlights a significant gap, implying that, despite users placing great value on this feature, the platform falls short of meeting their high expectations, potentially due to limitations in the timeliness of feedback or the effectiveness of support strategies.

In terms of evaluation and feedback functionality, users also demonstrated a high level of appreciation for its importance (mean = 3.808, SD = 0.730), yet their satisfaction (mean = 3.143, SD = 0.901) was lower, with a difference of 0.665,  $t = 6.871$ , and  $p < 0.001$ . This indicates a generally favourable view of this feature but highlights room for further improvement.

Finally, from an overall perspective, the platform's importance score (mean = 3.703, SD = 0.560) was significantly higher than the satisfaction score (mean = 2.969, SD = 0.574), with a difference of 0.734,  $t = 18.672$ , and  $p < 0.001$ . This suggests that, although the platform performs adequately in some respects, there remains a substantial gap between users' expectations and their actual experience.

Through further  $t$ -tests for each item, it can be seen in Table 33 that the significance of items 2, 5, 6 and 8 is greater than 0.001, indicating that students had similar perceptions of the importance and satisfaction in these four items. For other items, satisfaction is significantly lower than the importance, indicating that they need to be improved.

### **5.7.3 Importance-Performance Analysis Highlights Areas to Prioritise for Improvement**

Using SPSS 26 to process and analyse the available data, we obtained the distribution of various indicators within the IPA framework, as shown in figure 14. The mean importance is 4.173, and the overall mean satisfaction is 3.245.

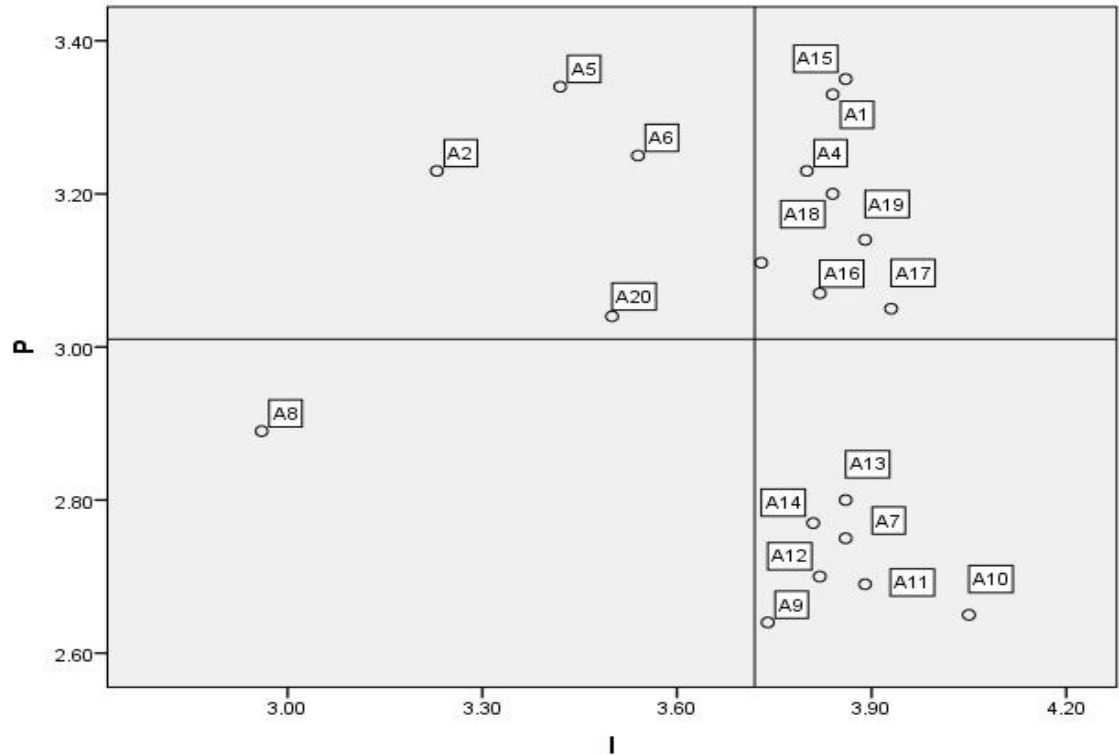
From the figure, it is evident that the indicators in the first quadrant (Keep up the good work), such as A1, A3, A4, A15, A16, A17, A18, and A19, fall into the high importance and high satisfaction area. These should be maintained at their current level of performance.

The indicators in the second quadrant (possible overkill), such as A2, A5, A6, and A20, are not considered very important by users. For these indicators, it is sufficient to provide basic maintenance, while attention should be directed to other areas.

The indicators in the third quadrant (low priority), such as A8, are of relatively low importance to users, and their performance is also poor. While these need improvement, identifying the specific issues and causes behind their underperformance is essential. However, as these indicators are of lower importance than those in the fourth quadrant, they should not be prioritised for improvement.

As seen in Figure 14, the indicators in the fourth quadrant (concentrate here), such as A7, A9, A10, A11, A12, A13, and A14, are highly important to users but underperforming, resulting in low satisfaction. These areas should be the focus of improvement, with efforts made to identify and address the underlying problems.

Figure 14 Quadrants of the Scales in the Survey of Students' Perceptions of WBFAT



## 5.8 Findings From the Student Focus Group Discussion

### 5.8.1 Student Focus on WBFAT Satisfaction and Importance

Through word frequency analysis of the transcripts of the student focus group discussion, word cloud maps of focus group discussions were generated (see Figure 15). As can be seen from the word cloud map, words or phrases such as ‘feedback’, ‘scoring’, ‘interaction’, ‘technical issues’, ‘user-friendly interface’, and ‘improvement’ appear very frequently.

Figure 15 The Word Cloud of the Student Focus Group Discussion



***Feedback: one of the focal points of students' attention regarding WBFAT.***

Feedback emerged as one of the most frequently discussed features of FiF. Students valued its ability to deliver immediate responses after practice, recognising this as a useful tool for timely self-assessment. However, they expressed dissatisfaction with the feedback's lack of personalisation and detailed explanations. Many students found the feedback overly generic, feeling that it did not adequately address their specific errors or provide clear guidance on how to improve. For example, rather than receiving tailored comments on pronunciation, grammar, or vocabulary, they often encountered vague statements that left them unsure about the precise areas needing improvement. Students suggested that incorporating more specific, actionable feedback would make FiF a far more effective learning tool, helping them to progress with greater confidence and clarity.

***Scoring: a major concern, but an aspect questioned by students***

Scoring accuracy was another major concern among students. Several students noted that FiF's scoring system often failed to reflect their actual performance, particularly in areas like pronunciation. Many reported feeling that their scores were either disproportionately low or inconsistent with their real abilities, creating frustration and confusion. For instance, some students highlighted instances where they received unexpectedly low marks for pronunciation, despite perceiving noticeable improvement in their speech clarity and accuracy. This perceived discrepancy left them questioning the reliability of the scoring, as well as its usefulness for tracking their progress. Students suggested that a more accurate and transparent scoring system, ideally with criteria aligned more closely to individual performance aspects, would enhance their trust in the tool and provide clearer guidance for improvement.

***Interaction: various forms of interaction were expected.***

While FiF offered a platform for interactive learning, many students felt it lacked the immediacy and effectiveness of face-to-face communication. They observed that the current level of interaction within FiF, although somewhat beneficial, did not fully replicate the engaging and responsive experience of in-person discussions and immediate feedback from peers and tutors. Some students suggested that FiF could be enhanced by incorporating more robust interactive features, such as live chat rooms, group discussion forums, and direct messaging with both classmates and tutors. These additions, they

believed, would foster a greater sense of community and support, helping students feel more connected and encouraged in their learning journey. Students highlighted that such features would not only enhance engagement but also provide them with timely guidance and support, thus more closely mirroring the advantages of traditional classroom interactions.

***Technical issues: system crashes and network instability frustrated students***

Frequent technical issues, such as system crashes and network instability, were repeatedly highlighted by students as significant barriers to effective learning. These disruptions not only interrupted study sessions but also led to frustration and reluctance among students to use FiF consistently. Many reported that these issues affected their ability to concentrate, as sudden crashes or connectivity problems often forced them to restart activities or lose progress. As a result, students suggested that addressing these technical challenges would be crucial for enhancing the tool's reliability and improving their overall learning experience with FiF.

***User-friendly interface: lack of clarity and accessibility***

While the interface was generally viewed as intuitive and easy to use, some students felt it lacked clarity in certain areas, particularly when trying to locate specific functions. They reported challenges in navigating to key features, which often led to frustration and disrupted their learning process. For instance, self-assessment and feedback functions were not always immediately visible or easily accessible. Students emphasised the importance of a more streamlined layout or clearer categorisation of features, suggesting that enhancing the user interface would improve the tool's overall usability and effectiveness.

***Improvement: personalised feedback and adaptive content were most expected***

Many students offered constructive suggestions to improve FiF, emphasising the need for more personalised feedback, guidance on effective learning strategies, and adaptable content that aligns with individual progress. For instance, they suggested implementing feedback that directly addresses their specific strengths and weaknesses, helping them to target areas needing improvement more precisely. Additionally, students expressed a desire for tips on study techniques and strategies to maximise their learning efficiency. They also recommended that FiF's content adjust dynamically to reflect their

evolving skills, allowing for a more tailored and supportive learning experience. These enhancements, they felt, would make FiF more responsive to their personal learning needs and promote a deeper, more engaging educational experience.

### **5.8.2 Students were most concerned about feedback and grading functions of WBFAT**

From Table 34, it can be seen that students are most concerned about feedback and scoring functions, and the shortcomings of these functions are frequently mentioned. Technical issues and interactive features also led to more discussion, reflecting the need for improvements in system performance and learning experiences.

*Table 34 The Node Coding Frequency of the Student Focus Group Discussion*

|                  | <b>Coding Frequency</b> |
|------------------|-------------------------|
| Feedback         | 20                      |
| Scoring          | 15                      |
| Interaction      | 10                      |
| Technical Issues | 12                      |
| User Interface   | 8                       |
| Improvements     | 14                      |

#### ***Feedback (20 mentions)***

Students repeatedly emphasised that feedback is a critical feature but expressed dissatisfaction with its lack of specificity and personalisation. For instance, they felt that the feedback seemed generic and wanted more targeted suggestions. They suggested that FiF's feedback system could provide more immediate and detailed responses, particularly for improving complex sentences.

#### ***Scoring (15 mentions)***

The scoring system was a key concern, with multiple students noting inaccuracies. Some students expressed frustration that their pronunciation was often scored lower than expected despite their efforts. They found that the system did not offer enough detailed analysis to understand where they went wrong, pointing to a need for smarter algorithms to enhance accuracy.

#### ***Interaction (10 mentions)***

While interaction was generally appreciated, many felt that the system's interaction features were limited compared to face-to-face practice. Students appreciated the interaction function for facilitating learning but some of them noted it was still not as effective as real-life communication. The limited flexibility in real-time discussion was also seen as a shortcoming.

### ***Technical Issues (12 mentions)***

Students frequently discussed technical problems, including system crashes and slow responses, which hindered their practice. For example, several students expressed frustration about system crashes affecting motivation and highlighted how these issues slowed their progress.

### ***User Interface (8 mentions)***

User interface feedback was mixed, with some appreciating its simplicity, while others found certain functions harder to locate, which detracted from the user experience. They suggested clearer navigation would improve the tool's usability.

### ***Improvements (14 mentions)***

Students offered numerous suggestions for improvements, such as adding more personalised learning paths and improving the feedback system, they wanted faster feedback and more real-world conversation exercises.

## **5.8.3 Students Were Dissatisfied with the WBFAT's Functions of Feedback and Scoring**

As can be seen from Matrix Coding Query Results in Table 35, students have more negative comments on the feedback and grading functions, indicating that there is great dissatisfaction with these two functions. Relatively speaking, the user interface received the most positive reviews, while the interactive features were more balanced, indicating that while the interactive features have advantages, there is still room for improvement.

*Table 35 The Matrix Coding Query Results of the Student Focus Group Discussion*

|                  | <b>Positive Comments</b> | <b>Negative Comments</b> |
|------------------|--------------------------|--------------------------|
| Feedback         | 5                        | 15                       |
| Scoring          | 3                        | 12                       |
| Interaction      | 6                        | 4                        |
| Technical Issues | 2                        | 10                       |
| User Interface   | 7                        | 1                        |
| Improvements     | 4                        | 10                       |

### ***Feedback***

With 15 negative comments, it's clear that students were often dissatisfied with the feedback, especially regarding its lack of personalisation. Some students pointed out that feedback often felt repetitive and didn't address individual needs. However, some students valued its immediacy, which allowed them to correct mistakes quickly.

### ***Scoring***

The scoring received 12 negative comments, reflecting widespread frustration. Some students noted that the system didn't provide scores that reflected their actual performance, which left them feeling demotivated.

### ***Interaction***

Interaction saw a more balanced view, with 6 positive comments. Most of the students appreciated the chance to practice with classmates and valued this feature for building confidence. However, technical issues and limited real-time features diminished the overall effectiveness.

### ***Technical Issues***

With 10 negative comments, technical issues were a significant pain point. Students mentioned that these problems directly affected their confidence in using the tool, leading to a negative learning experience.

### ***User Interface***

The user interface was mostly praised, with 7 positive comments. For instance, some of them appreciated the tool's design for its simplicity compared to other tools on the market.

### ***Improvements***

Improvements of WBFAT received both positive and negative feedback. Suggestions included adding personalised feedback and improving the scoring system to better reflect students' abilities.

## **5.9 Chapter Summary**

The findings in this chapter underscore the significant impact of WBFAT on the experimental group, showing enhanced engagement, oral accuracy, fluency, and confidence compared to the control group. Descriptive and inferential analyses reveal that students in the experimental group outperformed their control counterparts in most areas, particularly in later oral tests, where the effectiveness of the WBFAT became more pronounced. While the WBFAT demonstrated effectiveness in enhancing student engagement and oral English proficiency, several limitations emerged that warrant further discussion. Firstly, both groups showed minimal progress in advanced linguistic aspects such as deep cohesion and lexical concreteness, indicating that WBFAT may need additional functionalities to address these complex language features. The generic nature of feedback was frequently criticised

by students, who expressed a desire for more tailored, specific responses to better address individual errors and weaknesses in areas such as pronunciation and grammar.

The study also identified technical issues, such as system crashes and network instability that hindered consistent usage and impacted the learning experience. Furthermore, while students valued the interactive elements of WBFAT, many felt they did not match the effectiveness of face-to-face interactions. Improvements in the tool's adaptability and personalisation are necessary to cater to the diverse learning needs and preferences of students, as highlighted in focus group discussions and the Importance-Performance Analysis.

In sum, while WBFAT proved valuable in supporting autonomous learning and improving oral English proficiency, enhancing the feedback system, resolving technical issues, and refining interactive functionalities could further elevate its effectiveness and meet students' expectations for personalised and engaging language learning. These challenges give rise to further concerns about WBFAT's impact on developing EFL learners' oral English proficiency at a Chinese higher vocational college. In the next chapter, the researcher will build upon these findings to contextualise the research themes and research questions in the context of Chinese high vocation education while integrating insights from the literature review to conduct further in-depth discussion.

## **Chapter 6 Discussion**

This chapter builds upon the research aims and questions of the study, focusing on an in-depth discussion of the findings presented in Chapter 5. First, the researcher revisits the significance of the research themes and questions. This study centres on exploring the role of the WBFAT in facilitating English oral learning among Chinese vocational college students, with three key research themes: students' attitudes towards using WBFAT, the impact of WBFAT on their oral English proficiency, and their satisfaction with and perceptions of the tool's importance. Through further exploration of these themes, this chapter seeks to uncover the theoretical and practical implications of the study's findings.

Subsequently, the chapter delves into students' attitudes towards using WBFAT, highlighting the observed disparity between their high expectations of the tool and their relatively low willingness to use it. This gap is discussed in relation to the extent to which WBFAT's design aligns with students' intrinsic and extrinsic motivational needs. The chapter then examines the impact of WBFAT on students' oral English proficiency, focusing on improvements in autonomous learning engagement, oral test performance, and classroom participation, while also addressing the tool's limitations in enhancing complex linguistic abilities.

Finally, the chapter explores students' satisfaction with WBFAT and their perceptions of its importance, noting their increasing demand for dynamically updated and personalised learning resources. Furthermore, issues such as the insufficient integration of the tool with teaching practices and the temporal gap between learning outcomes and psychological expectations are discussed as potential factors influencing satisfaction levels. The discussion in this chapter provides a theoretical foundation for the further optimisation of WBFAT and offers practical recommendations for its application in vocational education.

### **6.1 Revisiting the Research Aims and Questions**

This is an exploratory study aiming at achieving three objectives through a mixed research methodology with the combination of literature review, local sample surveys, and small-scale experimental sampling. First, it seeks to ascertain the attitudes of Chinese higher vocational students towards the use of WBFAT. Second, it aims to explore the impact of WBFAT on their autonomous oral English learning and oral English proficiency

development, which includes examining the influence of WBFAT on their autonomous learning processes, oral English proficiency improvement, and classroom speaking performance. Third, the study investigates users' satisfaction with WBFAT and their perceptions of its importance, with the ultimate goal of identifying areas for improvement in WBFAT tailored to the needs of Chinese vocational college students, which also helps to understand how these learners can better adapt to the ever-emerging technology-enhanced learning and assessment models.

To achieve these objectives, three corresponding research themes were established, they are 'students' attitudes towards using WBFAT', 'students' performance and progress in oral English proficiency with the use of WBFAT', and 'students' perceptions of WBFAT', encompassing six research questions. The first research question focuses on vocational students' attitudes towards using WBFAT, aligning with the first research objective and the first theme. The second research question investigates students' performance in autonomous oral English learning, while the third examines whether WBFAT significantly enhances their oral English proficiency. The fourth research question considers whether WBFAT positively influences their classroom speaking performance. These three questions align with the second research objective and the second theme. The fifth research question addresses students' satisfaction with WBFAT, and the sixth explores their perceptions of its importance, both of which correspond to the third research objective and the third theme.

Based on the findings outlined in Chapter 5, this discussion will further elaborate on these research themes and questions.

## **6.2 Discussion on Theme 1**

The statistical analysis of data from the questionnaire on students' attitudes on WBFAT (see Chapter 5) shows that students generally held positive attitudes towards the WBFAT. The mean values of the three dimensions indicate that respondents had the highest expectations for WBFAT, followed by willingness to use, and lastly, motivation. While respondents have high hopes for WBFAT, their actual willingness and motivation to use it are lower, possibly due to uncertainties about its effectiveness and suitability. The high expectations may reflect their desire for WBFAT to address certain learning needs, but the tool may not yet fully attract some users. The low standard deviations across the three

dimensions suggest that respondents' ratings were relatively consistent, indicating minimal disagreement. Particularly in terms of expectations, most respondents had a positive attitude towards WBFAT, reflecting their hope that it could benefit their learning. However, the slightly higher variability in motivation suggests that some respondents may lack sufficient drive to use the tool, likely due to differences in personal learning needs and habits.

### **6.2.1 High Value of Expectation of Using WBFAT Indicates a Strong Demand for and Trust in WBFAT in Their Learning**

The mean of students' expectations of using WBFAT is higher than those for other dimensions, reflecting their acceptance of technology (Davis, 1989) and their trust in the potential usefulness of technology-empowered learning. This phenomenon can be traced back to the context of Chinese higher vocational education and the psychological characteristics of Chinese vocational college students. As discussed in the background section of this study (see Chapter 2), Chinese higher vocational education has experienced rapid development over the past two decades. However, its growth has been marked by significant disparities across different regions, particularly between the eastern, central, and western areas, as well as between urban and rural regions. These disparities extend to institutional resources, structural hierarchies, teacher-student ratios, and quality outcomes (Xu et al., 2022; Yang and Liu, 2016). Vocational colleges in more developed regions possess greater financial capacity to purchase or lease advanced, feature-rich platforms for learning and assessment that can support students' educational needs. Conversely, institutions in less developed areas often lack the resources to invest in such technologies. The vocational college selected for this study is located in one of these underdeveloped regions and is emblematic of the latter situation. Under these conditions, students tend to have higher expectations for technology platforms that can significantly enhance learning efficiency and outcomes. Platforms like FiF, an advanced WBFAT with diverse functionalities, are perceived by Chinese university students as essential resources to address their learning and evaluation needs. In language learning, the provision of authentic and comprehensible input through information and communication technologies (ICT) to engage learners in teaching and learning processes is critical (Krashen, 1998).

Furthermore, the typically large class sizes in Chinese vocational colleges—often exceeding 40 students per class—limit opportunities for individual practice during lessons. This is particularly problematic in skills-based language learning, where students require substantial practice to improve their listening and speaking abilities. However, due to an imbalanced teacher-student ratio, teachers are unable to provide personalised guidance to each student in class. Consequently, autonomous learning outside the classroom becomes an essential avenue for students to consolidate classroom content and enhance their skills. The effectiveness of autonomous learning hinges on students' access to abundant learning resources and timely evaluation and feedback. A review of the literature highlights the critical role of technology-supported learning tools in facilitating this process. Such tools can provide learners with a wealth of language materials and interactive activities, create additional opportunities for practice, and monitor students' learning progress in real time, generating feedback reports that help them identify their strengths and weaknesses (Ghanizadeh et al., 2015; Egbert, 1999; Salaberry, 2001). Evaluation and feedback are indispensable in autonomous learning; tools that deliver clear learning suggestions and guidance can significantly enhance learning outcomes (García-Jiménez et al., 2015; Lin et al., 2024).

### **6.2.2 The Relatively Low Mean of Willingness of Using WBFAT Reflects a Significant Gap Between Students' Cognition of WBFAT and Their Actual Behaviour**

Although students demonstrate high expectations for WBFAT, their actual usage intention scores are relatively low, revealing a pronounced gap between their perception and behaviour. This disparity stems from the characteristics of Chinese vocational college students and the traditional classroom-centred learning model prevalent in China. Excluding the vocational undergraduate institutions (four-year program) and a limited number of 'Dual-High Vocational Colleges' (three-year program), the majority of common vocational institutions admit students, particularly those enrolled through independent recruitment pathways, who generally exhibit weak learning autonomy, self-discipline, and study skills. These students often lack a solid educational foundation and established habits of independent learning, despite some displaying specific professional knowledge strengths or notable problem-solving abilities (Guo & Tan, 2015; Liang, 2015). Vocational students are accustomed to a 'teacher-centred and passive-receptive' learning approach, with a

general deficiency in independent learning awareness (Xue & Su, 2024; Xiang, 2020; Wei, 2015).

The usage of WBFAT inherently promotes a autonomous learning approach, requiring students to manage their own learning processes with the support of the tool. This includes setting learning goals, planning study content, maintaining consistent engagement, and continuously adjusting their learning strategies to develop a unique, personalised independent learning model. However, such a methodology conflicts with the traditional learning habits of many students. A significant number of students may lack sufficient time management skills and intrinsic motivation to explore and utilise the tool proactively, which directly limits their willingness to use WBFAT.

From the perspective of the Technology Acceptance Model (TAM), perceived ease of use (PEOU) and perceived usefulness (PU) are critical factors influencing the adoption of technology (Davis, 1989). If students perceive the tool as overly complex or unintuitive during its initial usage, or if the perceived learning benefits do not align with their expectations, their intention to engage with the tool diminishes. Moreover, within the context of Chinese vocational education, the adoption of technological tools faces challenges due to disparities in technical support and students' digital literacy, which may hinder WBFAT from fulfilling its intended purpose of being 'easy to use.'

Additionally, according to motivation theory (Ryan & Deci, 2000), the willingness to use WBFAT may also be constrained by a lack of extrinsic motivation. For instance, if the use of WBFAT is not incorporated into assessment criteria or directly linked to exam-related tasks, students' extrinsic drive to engage with the tool may weaken. On an intrinsic level, some students might lack interest or confidence in using technology-based tools, particularly those accustomed to traditional rote-learning methods. These students might perceive WBFAT as unfamiliar or even daunting, further reducing their inclination to adopt it. This interplay of mismatched expectations, insufficient motivation, and conflicts with entrenched learning practices highlights the challenges in encouraging the use of WBFAT among Chinese vocational students. Overcoming these barriers requires not only technical improvements but also better alignment with students' educational contexts and learning behaviours.

### **6.2.3 The Variability in Motivation Necessitates that WBFAT Design Addresses both Intrinsic and Extrinsic Motivational Needs of Students**

Data analysis indicates significant individual differences in students' motivations for using WBFAT, reflecting the diversity of their learning needs. This variation in motivation manifests in the usage patterns: students with strong intrinsic motivation are more inclined to explore the diverse functionalities of the tool to support their learning processes, while those driven by extrinsic motivation tend to focus primarily on tasks related to examinations. Self-Determination Theory (SDT, see Chapter 3) underscores the central role of intrinsic motivation in autonomous learning (Deci & Ryan, 2000). Intrinsic motivation drives students to engage in learning activities for the inherent satisfaction or interest derived from them, rather than for external rewards such as grades or certificates (Ryan & Deci, 2000).

On the one hand, some students at vocational colleges exhibit considerable interest in English learning. For instance, students who achieved higher scores in English during the college entrance examination often possess strong intrinsic motivation. This compels them to actively seek out methods, including leveraging WBFAT, to achieve more advanced English learning goals. These students tend to explore the various features of WBFAT proactively to support their learning. On the other hand, for many vocational college students in China, foreign language learning (primarily English) is heavily influenced by extrinsic factors. More specifically, their motivation is shaped by external demands such as end-of-term assessments or future workplace requirements. Their motivation is often geared towards passing exams, obtaining professional qualifications, or ensuring proficiency for future employment (Li, 2021; Lian & Ji, 2017; Zeng, 2003). Such external motivation is typically short-term and instrumental, leading students to adopt a cost-effective approach focused on measurable outcomes. Consequently, they tend to use WBFAT only when it is directly aligned with their extrinsic goals and, even then, limit their attention to the functions that serve these specific purposes.

Therefore, the task design of WBFAT should cater to the intrinsic motivational needs of learners, encouraging them to experience the joy and inherent value of learning. At the same time, it's also necessary to pay particular attention to the sources of extrinsic motivation prevalent among vocational students. For example, incorporating simulation

tasks relevant to real-world professional scenarios, providing targeted exercises for entrance exams, or implementing reward systems such as points and feedback visualisation can better align with their expectations. Additionally, offering clear learning pathways, accurate assessment, and instant feedback for learning and practising can help students perceive a sense of incremental achievement and attainable goals, thereby enhancing their motivation to engage with the tool.

## **6.3 Discussion on Theme 2**

### **6.3.1 WBFAT Enhanced Students' Motivation and Engagement in OEAL & FA**

To gain a detailed understanding of the engagement of both the control group and the experimental group in OEAL & FA, the researcher identified 14 variables (see Chapter 4) to observe the performance of the two groups during the 20-week experimental phase. The analysis of these variables prompts the following reflections.

#### ***Significant differences in engagement reveal stratified autonomous learning abilities among vocational students***

The data analysis of the 14 variables shows that students performed relatively well in basic tasks (e.g., time investment, integrated tasks, pronunciation accuracy), with higher average completion rates and lower standard deviations, indicating more consistent performance in these areas. However, in more challenging tasks that required greater strategic adjustment (e.g., fluency, coherence, appropriateness, peer assessment), significant differences were observed in task completion rates. While some students managed to complete these tasks and actively seek feedback, many either failed to engage or showed minimal completion.

This contrasts sharply with findings from previous research, which highlighted that undergraduates generally exhibit stronger goal-setting abilities, higher learning motivation, and greater planning and self-regulation when completing complex tasks (Li, 2006; Gong et al., 2017; Lei & Huang, 2024). The results also reflect the pronounced variability in autonomous learning abilities among vocational students in China. This variability stems from differences in their academic backgrounds, learning objectives, and vocational aspirations.

A segment of vocational students, driven by strong intrinsic motivation, possess clear goal-setting abilities and actively engage in autonomous learning, often seeking

support from technological tools. In contrast, other students, influenced by insufficient learning motivation and a history of exam-orientated education, are accustomed to passively completing tasks and lack self-regulation skills (Wang et al., 2018; Wang, 2018; Zhang & Ya, 2024). When faced with complex tasks, such as fluency or coherence exercises, this latter group may abandon the tasks due to their perceived difficulty, further widening the gap in capabilities among students.

### ***Intergroup differences reflect the support of technology in autonomous learning***

The experimental group demonstrated significantly better performance than the control group in terms of time investment, task completion, and particularly in fluency tasks, coherence tasks, appropriateness tasks, feedback utilisation, strategy adjustment, and peer interaction. This indicates that the intervention had a positive and significant impact on students' autonomous learning behaviours, further supporting the efficacy of WBFAT in promoting autonomous learning (Sharma et al., 2023; Robertson et al., 2019; Elmahdi et al., 2018; Sun, 2011).

Learner engagement, as Dörnyei (2001) suggests, is closely tied to the authenticity of tasks and their relevance to learning objectives. As previously noted, the learning goals of most vocational college students are strongly aligned with their career needs. WBFAT addresses this by designing multimodal learning tasks and augmented reality-based vocational training scenarios that simulate real-world working environments. These tasks create a contextualised learning setting that integrates language acquisition with the students' professional goals, enabling them to recognise the practical value of language skills while simultaneously enhancing their interest and motivation to participate. In contrast, the tasks designed for the control group, operating within a traditional learning model, were limited in scope, lacking simulated contexts and confined to grammar exercises or fixed dialogue templates. Such tasks, being disconnected from students' real-world needs, made it difficult for learners to associate them with their professional aspirations or everyday life, resulting in a lower perceived value of the tasks. This lack of relevance diminished their motivation and engagement. The disparity highlights the critical role of task design in influencing learning motivation: only when tasks address students' real-life needs and professional aspirations can they fully engage learners' intrinsic drive and encourage active participation and sustained effort.

Moreover, technology-supported formative assessment tools that provide learners with immediate feedback not only help them correct errors in a timely manner but also enhance their motivation through reinforcement (Shute, 2008). This is particularly crucial for vocational college students in China. In the context of large-class teaching, where each class often comprises 40 to 50 students, it is challenging for teachers to offer timely, personalised guidance for every task. This challenge is even more pronounced in vocational institutions, where students typically exhibit lower levels of intrinsic motivation. Without prompt feedback, students are more likely to abandon tasks, further weakening their interest and engagement. Large-class teaching also limits opportunities for interaction and practical participation, hindering skill development and learning outcomes. Thus, leveraging technology or innovative teaching approaches to address these limitations is necessary and urgent.

### **6.3.2 Long-term Use of WBFAT Contributes to Improving the Test Scores of Students' Oral English**

Statistical analysis of the scores from the five monthly English oral tests conducted with the two groups of students (see Chapter 5) reveals that, starting from the third test, the experimental group's average scores began to significantly surpass those of the control group. This advantage became even more pronounced in the fourth and fifth tests. In contrast, while the control group also showed some improvement, their overall progress was more modest and remained relatively steady. These findings underscore the significant impact of WBFAT in improving vocational college students' performance in English oral examinations.

It has almost become a consensus that Chinese vocational college students generally perform reasonably well in English written tests but face significant challenges in oral expression, a phenomenon commonly referred to as 'Mute English.' This phenomenon highlights a pervasive lack of speaking practice in their language learning and is further exacerbated by the absence of authentic English-speaking environments. Additionally, due to the constraints of large-class teaching and limited teaching resources, students rarely receive timely, personalised evaluation and feedback during their learning and practice. Consequently, they often fail to identify errors in their oral expression, understand the reasons for those errors, or determine how to improve. Against this backdrop,

technology-supported formative assessment tools address these challenges by creating authentic language exchange environments through multimodal technology. For instance, AR and VR technologies can simulate everyday conversations or vocational scenarios, enabling students to practise language expression in specific contexts and thus compensate for the lack of real-world communication opportunities. Moreover, WBFAT provides immediate and personalised feedback, which is often unattainable in traditional teaching models. Using speech recognition and intelligent analysis technologies, WBFAT can automatically evaluate students' oral English learning and generate detailed feedback, covering aspects such as pronunciation accuracy, grammatical errors, and vocabulary use. Furthermore, it identifies specific areas of weakness and offers concrete suggestions for improvement.

In this study, FiF used by the experimental group not only automatically scored each practice task or test but also delivered clear and detailed real-time feedback. This feedback included error identification as well as suggestions for refinement. By analysing and reflecting on the feedback provided by FiF, learners were able to comprehensively evaluate their learning process, recognise both strengths and weaknesses, and adjust their learning strategies accordingly. This iterative process fostered a positive cycle of 'learn – feedback – improve', helping students enhance their English learning efficiency and outcomes, which has been widely evidenced in previous studies on technology-aided evaluation of Chinese learners' in English writing and speaking (Luo et al., 2018; Lu, 2015; Li and Yan, 2012; Qiao et al., 2012; Wang, 2024; Bai and Wang, 2019; Jiang and Chen, 2021).

### **6.3.3 WBFAT Has a Positive Effect on Improving Students' Oral English Performance in Class**

Based on the analysis of the instructor's responses across four interviews, it was affirmed that students in the experimental group demonstrated significant progress in their oral English performance in class compared to the control group after using FiF for a period of time. Noticeable improvements were observed in classroom participation, accuracy, fluency, and confidence in oral expression. This aligns with the perspective that 'technology-enhanced learning environments promote students' knowledge construction through contextualised tasks' (Sharma et al., 2023; Vygotsky, 1978; Piaget, 1973). The

tasks designed in WBFAT, such as simulated dialogues, role-playing in situational contexts, or vocational scenario training, encourage students to apply their linguistic knowledge in real-life or quasi-authentic contexts, actively engage in communication, and solve problems. This ‘task-driven learning’ approach enables students to transform fragmented linguistic knowledge into practical communicative competence during task completion, thereby enhancing their ability to use language effectively in the classroom.

Constructivist learning theory posits that learning is achieved through individuals actively constructing knowledge, interacting with their environment, and reflecting on their own cognitive processes. WBFAT provides a dynamic and interactive learning environment for experimental group students, where authentic task contexts and personalised feedback stimulate their intrinsic motivation to learn and foster their capacity for knowledge construction.

Additionally, confidence plays a critical role in classroom participation and language performance. WBFAT significantly boosted students’ confidence through immediate feedback and tailored guidance. Shute (2008) observes that instant feedback not only enhances learners’ performance but also significantly strengthens their confidence and motivation. Chinese students often experience language anxiety in English oral classes, with a fear of making mistakes that undermines classroom participation (Horwitz et al., 1986). WBFAT addressed this issue by offering a wide range of learning and practice materials designed in a progressive format, moving from simpler to more complex tasks. This scaffolding allowed students to experience incremental success, particularly in tasks that were achievable with reasonable effort. The accumulation of such success experiences bolstered their confidence in oral expression, encouraging them to engage more actively in classroom activities. Moreover, by providing automatic scoring and immediate feedback, WBFAT enabled students to repeatedly practise and correct errors independently, thereby gradually reducing their anxiety about making mistakes. As their anxiety diminished, students became more willing to speak up and respond to questions in class.

Furthermore, this improvement in classroom performance can be attributed to the ‘spillover effect’ of WBFAT’s positive impact on students’ autonomous learning. According to self-regulated learning theory (Zimmerman, 2002), skills and strategies developed during autonomous learning can transfer to other learning contexts. In the

process of using WBFAT, students were required to engage in self-management, such as planning their learning schedules, setting goals, and evaluating their progress based on system-generated tasks. These self-regulation skills extended to classroom contexts, manifesting as higher levels of participation and greater task completion willingness (Luo et al., 2018; Robertson et al., 2019).

#### **6.3.4 WBFAT Provided Positive Support for the Development of Oral English Proficiency in Higher Vocational College Students**

Findings from teacher interviews and classroom observation records reveal that approximately 11 weeks after using FiF, students in the experimental group demonstrated significantly better performance in oral English pronunciation accuracy, vocabulary accuracy and specificity, grammatical accuracy, fluency, and completeness compared to the control group. Moreover, as time progressed, the gap between the two groups in these areas continued to widen. This strongly suggests that WBFAT plays a positive role in supporting the development of oral English proficiency among higher vocational college students (Stockwell, 2022; Yesilyurt, 2023; Shadiev & Yang, 2020; Levy & Hubbard, 2005).

Additionally, this phenomenon is closely related to the typical tendencies of Chinese higher vocational students in language learning. Analysis of 14 variables from the students' autonomous learning during the experimental phase indicates that their tasks related to oral English proficiency were most concentrated on pronunciation accuracy, followed by vocabulary and grammar exercises, and then fluency practice. However, tasks targeting coherence and appropriateness were completed significantly less frequently. This trend reflects the characteristic learning preferences of Chinese vocational students. For them, pronunciation and fluency practice are relatively straightforward and easy to manage. By following the simple pattern of 'listening, repeating, and practising,' students can quickly achieve visible improvements. The immediate feedback and apparent results motivate students to allocate more time to these aspects. At the same time, vocabulary and grammar instruction has historically been prioritised in English education in China, particularly at the beginner levels. As fundamental components of exam-focused education, vocabulary and grammar mastery are regarded as the foundation of language learning. This emphasis creates a habitual preference for vocabulary and grammar-related tasks during independent study (Dörnyei, 2014). While this inclination enhances the use and application

of vocabulary and grammar, it often diverts attention from the more demanding aspects of language learning, such as coherence and appropriateness.

### **6.3.5 WBFAT Had Limited Impact on the Development of Students' Complex Language Skills**

Despite the experimental group outperforming the control group in descriptive indicators (DESWC, DESSL), lexical diversity (LDTTRc, LDTTRa), and content word concreteness (WRDCNCc), they showed no significant advantage over the control group in the remaining 15 variables. This indicates that FiF has a limited impact on the development of complex language skills in higher vocational students' oral English proficiency.

Firstly, the evaluation of autonomous learning often relies on indicators such as learning frequency, time invested, the number of tasks completed, and the quality of task completion. Stockwell (2022) notes that mobile-assisted language learning platforms provide students with flexible learning opportunities, where high-quality task completion correlates positively with language improvement. Specifically, the higher the quality of task completion, the greater the improvement in language ability.

However, analysis of the 14 observational variables from the autonomous learning phase revealed that both experimental and control group students concentrated primarily on relatively simple tasks, such as pronunciation, vocabulary, grammar, and fluency. Tasks involving more complex language skills, such as coherence and appropriateness, were completed far less frequently (with a mean of 14.91 for the EG and only 2.77 for the CG).

According to Ellis's (2003, 2017) task-based language teaching (TBLT) theory, language learning is most effective when driven by tasks, particularly those with high cognitive demands. Such complex tasks encourage comprehensive language output and communication skills. Simple tasks, in contrast, often lack contextualised input and output, making it difficult for students to develop the ability to organise language coherently, adapt to contextual demands, and engage in meaningful communication. As a result, even when students acquire grammatical rules and vocabulary, they often struggle to express themselves fluently and confidently in real-world communication scenarios. Without sufficient exposure to complex, context-rich tasks, their ability to use language effectively in spontaneous interactions remains underdeveloped.

## **6.4 Discussion on Theme 3**

Despite the findings from interviews, oral tests, classroom observations, and focus group discussions indicating that the WBFAT platform has positively contributed to improving vocational students' oral English proficiency, the survey on users' satisfaction revealed a low overall satisfaction with WBFAT. Students also expressed dissatisfaction with the two features they considered most important: providing resources and supporting the development of their oral English proficiency. This raises a thought-provoking issue.

#### **6.4.1 Learners' Growing Demand for Personalised and Dynamically Updated Learning Resources**

Students acknowledged that FiF provides a large quantity of learning materials covering diverse topics and learning needs, including simulated dialogues, workplace English, and situational exercises. However, the lack of an intelligent recommendation feature makes it difficult for learners to swiftly locate content that best suits their individual needs. Vocational students' English learning goals are often highly personalised: some students aim to enhance their workplace English skills, while others focus on improving their daily conversational ability. The absence of functionality to recommend relevant content based on learners' progress, interests, or career objectives forces them to spend significant time navigating resources. This mismatch in content can reduce learners' enthusiasm for using the platform.

Another critical factor affecting satisfaction is the insufficient dynamic updating of learning resources. If students frequently encounter outdated materials that are disconnected from real-world language usage, they are more likely to experience boredom and frustration. The 'ageing' of learning resources not only contributes to negative perceptions of the platform but may also diminish learning outcomes. Regular updates to learning resources can effectively sustain learners' interest and help them maintain high levels of motivation (Mishra & Koehler, 2006).

Additionally, the current resource delivery method of many platforms tends to be one-directional, lacking support for learners to autonomously construct their own resource libraries. During their learning process, students may have specific interests or needs. Allowing them to upload and share customised learning materials or training content, such as recorded speaking tasks, industry-specific vocabulary lists, or peer-designed collaborative tasks, could significantly enhance their engagement and sense of ownership.

Furthermore, self-built resource libraries could better address the personalised needs and practical scenarios of vocational students, thereby complementing the platform's existing resources and creating a more targeted learning ecosystem.

#### **6.4.2 The Importance of Integrating Tools Like WBFAT Into Teaching**

The effectiveness of language learning tools largely depends on the extent to which they are integrated into teaching practices. For Chinese vocational students, this integration is particularly crucial as their learning methods are heavily reliant on teacher guidance and assigned tasks. While tools such as WBFAT offer a wealth of functional support, these functionalities may fail to realise their intended potential if not closely aligned with course objectives and teaching content. This disconnect can prevent students from recognising the practical value of the tools in their learning processes.

Moreover, insufficient integration into teaching can leave students without a clear sense of direction or purpose in using the platform. The effectiveness of educational technology tools is not only determined by their technical design but also by how teachers incorporate them into learning contexts. If teachers do not systematically embed the platform's functionalities into teaching activities—such as assigning pre-class tasks, demonstrating usage in class, or providing follow-up feedback—students may only engage with the tools superficially, failing to unlock their full potential. This lack of integration may further lead to low utilisation rates, ultimately affecting students' learning outcomes and satisfaction with the tools.

#### **6.4.3 Students' Satisfaction May Be Influenced by the Gap Between Expectations and the Time Needed for Learning Outcomes**

Students' satisfaction with learning tools is often influenced by the discrepancy between their expectations and the actual time required to achieve learning outcomes. Vocational students typically aim for rapid improvement in their practical language skills. The conflict between these short-term goals and the long-term nature of language proficiency development may lead to disappointment. For instance, although test results and classroom observations demonstrated that WBFAT positively impacts students' speaking proficiency, students often find it difficult to perceive these improvements over a short period, especially in areas such as fluency and grammatical accuracy, which tend to show incremental progress. This delay in perceiving progress might lead students to

believe the platform's effectiveness falls short of expectations, thereby lowering their satisfaction.

#### **6.4.4 Inherent Flaws in the Tool Can Undermine Users' Expectations and Satisfaction**

During focus group discussions, several students highlighted concerns about the reliability of the platform's automated scoring system. Some students reported receiving high scores despite making obvious mistakes, which raised doubts about the accuracy and rigour of the scoring system. This undermined their trust in the platform and, consequently, their satisfaction.

Additionally, the lack of specificity in feedback further reduced the practical value of the tool. Vocational students' learning needs are often highly specific and practical, but the platform's feedback was often too general, such as merely indicating 'grammar errors' or 'pronunciation needs improvement,' without providing detailed error analysis or clear directions for improvement. This insufficiently targeted feedback left students feeling confused or at a loss, fostering negative attitudes towards the tool. Feedback is an essential function of assessment, designed to provide students with clear and specific information about their learning performance. Strengths and weaknesses are identified, and improvement is guided through such feedback (Hattie & Timperley, 2007). To be effective, feedback is required to exhibit timeliness, specificity, and constructiveness. By providing feedback promptly, adjustments to learning strategies can be made efficiently. Through specific feedback, areas requiring improvement are pinpointed accurately, while constructive feedback ensures actionable solutions are offered to address identified challenges. Within the context of Chinese higher vocational colleges, the provision of feedback should be closely aligned with the diverse learning foundations of students, who are often characterised by strong practical skills but weaker theoretical knowledge. In skills-based courses, operational details can be optimised through the use of immediate feedback by instructors. In theoretical courses, a gradual and systematic approach is recommended, with feedback directed towards addressing knowledge gaps and reinforcing foundational concepts. Furthermore, since practicality is prioritised by vocational students, the inclusion of real-world examples and professional contexts in feedback is advised to stimulate interest and enhance relevance to their career development.

#### **6.5 Chapter Summary**

This chapter, through an in-depth discussion of the research findings, highlights both the effectiveness and the limitations of WBFAT in facilitating oral English learning among vocational college students. The findings reveal significant motivational differences among students when using WBFAT, reflecting not only their expectations regarding the tool's functionality and design but also its shortcomings in addressing the diverse learning needs of students. Furthermore, WBFAT has demonstrated notable effectiveness in certain aspects of oral English proficiency, particularly in enhancing learning engagement, increasing classroom participation, and improving the accuracy and fluency of expression. However, its limited impact on the development of complex linguistic abilities underscores the need to explore the potential of technology-supported tools in fostering advanced language skills.

Based on the discussion, the study raises the following considerations. Firstly, there is a need to better balance intrinsic and extrinsic motivational needs in the design of such tools to enhance students' acceptance and actual willingness to use them. Secondly, the tool's limited effectiveness in fostering complex linguistic abilities indicates a need for further development, such as integrating deeper corpus analysis or advanced language structure practice. Finally, the effective integration of WBFAT with traditional teaching methods remains a critical challenge. Future efforts could focus on exploring dynamic resource updates and personalised support to establish a collaborative model between technology and classroom teaching, thereby achieving a dual enhancement of teaching efficiency and learning outcomes.

## **Chapter 7 Conclusion**

In this chapter, I will provide a summary of the entire study, exploring its academic contributions, practical significance, research limitations, and future research directions. First, we reviewed the purpose and research methodology of this study and analysed the impact of WBFAT on the oral English proficiency of EFL students in Chinese higher vocational colleges. Secondly, based on the research findings, we clarified the academic contributions of this study, particularly its role in filling the research gap in the context of EFL students in higher vocational education, as well as deepening learning theories and technology acceptance theories. The chapter then discusses the practical significance of the study, highlighting the application of findings in teaching practice and the enhancement of students' autonomous learning abilities. Finally, the limitations of the research are acknowledged, and several directions for future research are proposed, including personalised learning, the integration of emotional factors, cross-cultural adaptability, and the exploration of long-term impacts.

### **7.1 Summary of the Study**

This research focuses on exploring the impact of WBFAT on the development of oral English proficiency among students in Chinese higher vocational colleges. The aim is to comprehensively investigate students' attitudes towards using WBFAT, analyse its role in improving the oral English skills of vocational students, and gain an in-depth understanding of their satisfaction and perceptions of the tool's importance after its use. The findings of this study not only aim to provide valuable references for the construction of autonomous learning models and the effective implementation of e-learning strategies for EFL learners in Chinese higher vocational colleges but also to promote the optimisation and improvement of WBFAT by developers. This includes enhancing its interface design, content structure, and functionality to better cater to the unique learning characteristics and practical needs of Chinese vocational college students.

To achieve these research objectives, the researcher employed a mixed-methods approach guided by an explanatory sequential design. The study involved pilot research, a survey, and an experimental study. The pilot research played a key role in the research process by serving as a precursor and calibration stage. It helped to initially test and optimise the research tools, methods, and processes, enabling the identification of potential

issues—such as the design of the questionnaire or the clarity of interview outlines—in order to make timely adjustments and ensure the scientific rigour and effectiveness of the subsequent large-scale study. In the survey phase, the researcher used questionnaires to comprehensively understand vocational students' attitudes towards using WBFAT, laying the foundation for the subsequent experimental study. Over the course of 20 weeks, the researcher conducted a quasi-experimental study to observe and record data on the experimental and control groups of students regarding their involvement and performance in English oral autonomous learning and formative assessment, their performance in five oral English tests, and their classroom speaking performance. This provided firsthand evidence for determining whether WBFAT had a positive effect on students' oral English proficiency development. In addition, teacher interviews, surveys on the satisfaction and perceived importance of FiF, and focus group discussions were used to provide comprehensive feedback and to assess the experimental process and outcomes from multiple perspectives. This allowed for a thorough evaluation of changes in students' oral English skills, classroom performance, and perceptions of the tool's satisfaction and importance.

Through these research procedures and methods, the study addressed three main themes and answered six research questions. First, survey data from 119 vocational students revealed their attitudes towards using WBFAT in their learning, highlighting their willingness to use, motivation for use, and expectations for use. Second, the 20-week experimental study provided a comparative analysis of the experimental group (22 students) and the control group (22 students) regarding their performance in autonomous oral English learning, their scores in five oral English tests, language proficiency characteristics, and their use of oral English in class, providing a comprehensive evaluation of WBFAT's effect on the improvement of oral English proficiency for vocational EFL learners. Third, through questionnaires and focus group discussions, the study gathered insights on the experimental group students' satisfaction with WBFAT and their understanding of its importance after 20 weeks of use, identifying the areas of concern and needed improvements.

The findings revealed that, in the preliminary attitude survey, participants generally held positive views towards WBFAT, although they had a neutral stance regarding its role

in personalised learning. Vocational EFL students exhibited a complex pattern of participation in autonomous learning and formative assessment, with significant individual differences. The experimental and control groups displayed varying performance in terms of task completion, feedback and strategy adjustment, and peer interaction. Regarding oral test scores, the experimental group showed notable improvement from the third test onwards, with fluctuating results, while the control group had slower but more stable progress. In terms of language features, the experimental group outperformed in certain areas, such as vocabulary and sentence length, with t-tests confirming significant differences in these indicators. Classroom observations showed that the experimental group's speaking performance improved over time. Interviews indicated that WBFAT had a positive impact on various aspects of students' oral English proficiency, but challenges remained. In terms of satisfaction and perceived importance, t-tests and IPA analysis identified key concerns and areas for improvement in the tool. The focus group report highlighted students' primary concerns and dissatisfaction with aspects like feedback and scoring features, offering a comprehensive presentation of students' experiences with WBFAT.

These results demonstrate that the design of the study, implementation, and methods were scientifically sound and feasible. The mixed-methods approach, integrating both qualitative and quantitative research, ensured the reliability of the findings. The combination of large-scale surveys, standardised tests, in-depth interviews, and focus group discussions provided a rich collection of evidence and allowed for a multidimensional assessment of WBFAT's effects. The study answers all of the research questions and achieves its goal of exploring the impact of WBFAT on the oral English proficiency of vocational students. Its results have the potential for replication and application in similar educational research settings.

## **7.2 Academic Contributions**

This study makes several significant academic contributions. First, it fills a critical gap in existing research by focusing on EFL learners in Chinese higher vocational colleges, a group that has been largely overlooked in studies on the integration of teaching, learning, and technology in China's education. The literature review in Chapter 3 showed that previous studies have mainly concentrated on university undergraduates and primary or

secondary school students, with relatively few studies on the growing cohort of students in Chinese vocational colleges, particularly in relation to the use of technology-based tools for supporting autonomous learning.

Second, from a theoretical standpoint, the study provides an in-depth analysis of the complex cognitive, behavioural, and ability development processes of vocational EFL students in oral English learning under the intervention of WBFAT. It sheds light on how students gradually construct strategies for autonomous speaking learning and the development of oral English proficiency through technology, thereby offering new directions for the advancement of theories on autonomous learning and language learning, especially in the context of vocational education.

Third, in terms of technology acceptance and application theory, the study enriches the theoretical understanding of the psychological and behavioural driving forces behind vocational EFL students' acceptance of WBFAT. It examines the multi-dimensional factors such as perceived functionality, ease of use, and learning outcomes, thereby filling a gap in research on the use of educational technology in vocational EFL teaching and autonomous learning.

Fourth, in terms of research methodology, this study employs a mixed-methods approach, combining both qualitative and quantitative research to gather evidence from multiple perspectives. The study uses large-scale surveys, standardised tests, in-depth interviews, and focus group discussions to collect comprehensive data. The use of multiple data collection and analysis methods ensures the reliability of the conclusions. This approach opens up new perspectives for research on the use of technology in EFL oral autonomous learning and classroom teaching evaluation, contributing to the methods used in educational technology research and providing innovative ideas and methodological references for future studies.

### **7.3 Practical Implications**

First, the findings of this study provide targeted guidance for teachers to optimise their teaching practices in web-based teaching and learning activities. Teachers can design differentiated teaching plans based on the differences in students' performance with WBFAT. For students with weaker foundational skills, additional practice modules can be implemented, while for more advanced students, higher-level tasks can be provided. For

example, based on students' performance in oral tasks, teachers may assign pronunciation drills and vocabulary reinforcement exercises to lower-level students and debate or oral report tasks to more advanced learners. Throughout the teaching process, teachers should focus on guiding students to effectively use feedback to adjust their learning strategies and foster autonomous learning skills.

Second, this study provides practical guidance for vocational college EFL students in carrying out autonomous oral English learning. It encourages students to fully utilise technology-based tools, thereby enhancing their autonomous learning skills and improving their speaking performance. By deepening students' understanding of WBFAT, they will be able to select the most appropriate technological tools for their specific learning goals and content. For instance, if students aim to improve their pronunciation, they can use software with accurate speech recognition and correction functions, while those focused on expanding vocabulary can make use of vocabulary memory tools. Additionally, the study presents behaviour patterns and learning paths of successful learners, guiding students to set reasonable learning plans, follow through with their plans, regularly evaluate their progress, and flexibly adjust their learning strategies based on feedback from the technological tools. This process helps students transition from passive knowledge receivers to proactive knowledge seekers, enhancing their language foundation for future professional development and personal growth and strengthening their competitiveness and adaptability in a globalised job market.

Third, in terms of student assessment, this study encourages evaluators to create more scientific, dynamic, and comprehensive assessment systems that support learning through web-based technologies. By focusing on students' evolving abilities and shifting strategies throughout the learning process, it aims to improve the quality of oral English instruction and student autonomous learning outcomes in vocational colleges.

#### **7.4 Limitations of the Study**

Although this study provides valuable theoretical guidance and empirical evidence for the application of WBFAT in autonomous oral English learning in Chinese higher vocational colleges, it still has certain limitations. Firstly, the sample selection in this study is limited to a general higher vocational college in an inland province of China and was not extended to a broader range of vocational colleges. This limitation may affect the external

validity of the research results. Given the significant regional, cultural, and resource differences in Chinese higher vocational education and the varying degrees of implementation of IT, AIT, and autonomous learning across different institutions, the findings may not fully represent the situation in all vocational colleges. Secondly, during the questionnaire phase, although the survey targeted all students in the school, only 119 students participated, resulting in a relatively small sample size. This may affect the generalizability and representativeness of the data. Thirdly, the experimental and control groups consisted of only 44 students (22 per group), and the experimental period was 20 weeks. This duration may be insufficient for the development of language skills, especially for long-term development and steady improvement in speaking proficiency. The 20 weeks may not fully reflect the changes in the development of students' oral English proficiency. Lastly, this study only selected FiF as the experimental tool, which is a WBFAT focused on oral English learning. Whether other subjects or different types of WBFAT would have similar effects still requires further investigation. Therefore, while the results of this study have certain theoretical and practical value, the limitations must be carefully considered in future research.

### **7.5 Future Research Directions and Prospects**

With the continuous development of educational technology, the potential application of WBFAT in education has become increasingly significant. Future research is expected to deepen the exploration in multiple dimensions, further promoting the innovation and practice of WBFAT in language teaching and in supporting students' autonomous learning in language.

Firstly, personalised learning and intelligent feedback will be key areas for the development of WBFAT in the future. Existing WBFAT tools often provide standardised feedback mechanisms, which have not yet been refined to meet the individual needs of students. Future research could focus on how to leverage big data and artificial intelligence (AI) technologies to achieve more precise learning assessments and personalised learning path recommendations (Renz & Hilbig, 2020; Kakungulu, 2024; Owan et al., 2023). For instance, AI could adjust the assessment content and feedback strategies based on students' learning progress, error patterns, and learning styles, thereby helping students to master language skills more efficiently. The introduction of such intelligent feedback mechanisms

would significantly enhance students' motivation and language proficiency (Kuddus; Schmidt & Strasser, 2022).

Secondly, the impact of emotional factors on learning is gaining increasing attention. Existing WBFAT tools have yet to fully account for students' emotional states and how these influence learning. Future research could explore how affective computing technology can be used to monitor and adjust students' emotional responses in real time, providing learning support that better aligns with students' emotional needs. Research has shown that emotional factors play a crucial role in language learning, particularly when students experience anxiety or stress, as these feelings can negatively impact their learning outcomes (D'Mello & Graesser, 2014). Thus, future WBFAT should incorporate emotional regulation features in their design to reduce students' anxiety, enhance their learning interest and confidence, and ultimately improve learning outcomes.

Additionally, research on cross-cultural adaptability and globalisation is also an important direction for the future development of WBFAT. With the continual changes in the global educational environment, the cross-cultural adaptability of WBFAT has become particularly significant. In different cultural contexts, students' acceptance of and approaches to using technological tools may vary. Therefore, WBFAT should be designed with localisation in mind, according to different cultural and linguistic backgrounds, in order to maximise their effectiveness (Shonfeld et al.; Shadiev et al., 2021). Future research should explore how to design WBFAT tools that are better suited to the educational needs of different countries and regions, thus promoting their global application.

Finally, the long-term effects of WBFAT in language learning require further investigation (Lin et al.; Zhang et al., 2023). Current research mainly focuses on short-term effects, with insufficient in-depth analysis of long-term outcomes. Future studies could track long-term data to examine the lasting impact of WBFAT on students' oral English proficiency, autonomous learning abilities, and learning attitudes. This would provide a theoretical foundation for the continuous optimisation and practical application of WBFAT tools.

## **7.6 Chapter Summary**

This chapter recaps the research purpose, process, methodology, and findings of this study. It also elaborates on the theoretical contributions and practical significance of the research while addressing the limitations and shortcomings of the study and proposing future research directions. This comprehensive review not only provides an overall assessment of the study but also offers guidance for improving subsequent related research.

The study has completed all research tasks but leaves room for further reflection. In terms of research design, although a mixed-method approach was employed to balance qualitative and quantitative methods, the regional limitations of the sample selection have somewhat diminished the generalisability of the findings. Future research should expand the geographical coverage of the sample, incorporating vocational colleges from regions with varying levels of economic development and significant differences in educational resources to ensure a more representative study design. Regarding the research process and methodology, while the pilot study helped optimise the process in advance, issues such as low participation rates in the questionnaire survey, the small sample size of both the experimental and control groups, and the limited duration of the experiment indicate that future research should refine the questionnaire distribution strategy to encourage greater student participation. Additionally, the sample size should be appropriately expanded, and the experimental period should be extended to capture more comprehensive and long-term changes in language learning dynamics.

From the perspective of research findings, while I am pleased to have filled certain academic gaps and provided practical guidance, the results also reveal several areas for improvement in the current WBFAT, such as the lack of personalised feedback mechanisms, insufficient emotional engagement, and inadequate cross-cultural adaptation. These insights highlight the potential for further research using big data, artificial intelligence, and emotional computing technologies. They guide the future refinement of research paradigms and inspire continuous improvement in the field of EFL teaching and learning in Chinese vocational education, contributing to its ongoing development.

## Bibliography

- Akinyode, B. F., & Khan, T. H. (2018) 'Step by step approach for qualitative data analysis', *International Journal of built environment and sustainability*, 5(3).
- Alam, M. K. (2021) 'A systematic qualitative case study: questions, data collection, NVivo analysis and saturation', *Qualitative Research in Organisations and Management: An International Journal*, 16(1), pp. 1-31.
- Alam, A. (2022) 'Mapping a sustainable future through conceptualisation of transformative learning framework, education for sustainable development, critical reflection, and responsible citizenship: an exploration of pedagogies for twenty-first century learning', *ECS Transactions*, 107(1), pp. 9827-9840.
- Alamri, H., Lowell, V., Watson, W. & Watson, S. L. (2020) 'Using personalized learning as an instructional approach to motivate learners in online higher education: Learner self-determination and intrinsic motivation', *Journal of Research on Technology in Education*, 52(3), pp. 322-352.
- Alhazmi, A. A., & Kaufmann, A. (2022) 'Phenomenological qualitative methods applied to the analysis of cross-cultural experience in novel educational social contexts', *Frontiers in Psychology*, 13, 785134.
- Allen, L. K., Snow, E. L., & McNamara, D. S. (2014) 'Complexity, Readability, and Cohesion in Texts: Theoretical and Practical Considerations', *Behavior Research Methods*, 46(3), pp.848-860.
- Almoussa, N. A. (2023) 'The Impact of Using E-Formative Assessment on Critical Thinking', *Amman Arab University Journal*, 8(3).
- Anderson, T. (2008) *The Theory and Practice of Online Learning*. Athabasca University Press.
- Anney, V. N. (2014) 'Ensuring the quality of the findings of qualitative research: Looking at trustworthiness criteria', *Journal of emerging trends in educational research and policy studies (jeteraps)*, 5(2), pp. 272-281.
- Anthony, J., & Padmanabhan, S. (2010) 'Digital divide and equity in education: A Rawlsian analysis', *Journal of Information Technology Case and Application Research*, 12(4), pp. 37-62.
- Anwar, M. (2015) 'Mixed-methods research in pharmacy practice and its implications', *Archives of Pharmacy Practice*, 6(1). <https://doi.org/10.4103/2045-080x.151279>
- Avery, K.N., Williamson, P.R., Gamble, C., Francischetto, E.O.C., Metcalfe, C., Davidson, P., Williams, H. and Blazeby, J.M. (2017) 'Informing efficient randomised controlled trials: exploration of challenges in developing progression criteria for internal pilot studies', *BMJ open*, 7(2), p.e013537.
- Awidi, I. T. & Paynter, M. (2016) 'The impact of a flipped classroom approach on student learning experience', *Computers & Education*, 128, pp. 269-283.
- Aydemir, M. and Karaman, S. (2017) 'Üstbılışsel etkinliklerin uzaktan eğitim öğrencilerinin üstbılışsel seviyeleri ve ders çalışma süreçleri açısından incelenmesi', *Eğitim Teknolojisi Kuram Ve Uygulama*, 7(2), 18-18. <https://doi.org/10.17943/etku.288730>
- Azevedo, R. & Gašević, D. (2019) 'Analysing multimodal multichannel data about self-regulated learning with advanced learning technologies: Issues and challenges', *Computers in Human Behavior*, 96, pp. 207-210.
- Bagozzi, R. P. (2007) 'The Legacy of the Technology Acceptance Model and a Proposal for a Paradigm Shift', *Journal of the Association for Information Systems*, 8(4), 244-254.
- Bai, L. and Wang, J. (2019) 'A review of the effectiveness of automated feedback on English writing over the past 20 years', *Foreign Languages Research*, 36(1), pp. 65–71, 88.
- Bakar, M. A. A. and Ismail, N. (2020) 'Exploring metacognitive regulation and students' interaction in mathematics learning: an analysis of needs to enhance students' mastery', *Humanities & Social Sciences Reviews*, 8(2), 67-74. <https://doi.org/10.18510/hssr.2020.82e07>
- Baker, R. S., & Yacef, K. (2009) 'The state of educational data mining in 2009: A review and future

visions', *Journal of educational data mining*, 1(1), pp. 3-17.

Ban, J. W. (2020) 'Teaching and Educating is an Important Moral Mission of Teachers in the New Era'. *People's Education*, 2020(19), pp. 30 - 32.

Bandura, A. (1997) *Self-efficacy: The exercise of control* (Vol. 604). Freeman.

Bao, W. P., Lu, L., & Lou, J. (2022) 'The Distinction Between University Internal Governance and It's Internal Management', *Heilongjiang Researches on Higher Education*, 2022(40)(05), pp. 51 - 57.

Bárkányi, Z. (2021), 'Motivation, self-efficacy beliefs, and speaking anxiety in language MOOCs', *ReCALL*, 33(2), pp.143-160.

Barrot, J.S., Llenares, I.I. & del Rosario, L.S. (2021) 'Students' online learning challenges during the pandemic and how they cope with them: the case of the Philippines', *Educational Information Technology*, 26, pp.7321–7338. <https://doi.org/10.1007/s10639-021-10589-x>.

Bashir, N. (2020) 'The qualitative researcher: the flip side of the research encounter with vulnerable people', *Qualitative Research*, 20(5), 667-683. <https://doi.org/10.1177/1468794119884805>

Bazeley, P. (2012) 'Integrative analysis strategies for mixed data sources', *American Behavioral Scientist*, 56(6), pp.814-828.

Bećirović, S. (2023) 'Challenges and barriers for effective integration of technologies into teaching and learning', In *Digital Pedagogy: The Use of Digital Technologies in Contemporary Education*, pp. 123-133. Singapore: Springer Nature Singapore.

Behrens, J. T. (1997) 'Principles and procedures of exploratory data analysis', *Psychological methods*, 2(2), p. 131.

Belouafa, S., Habti, F., Benhar, S., Belafkih, B., Tayane, S., Hamdouch, S., Bennamara, A. and Abourriche, A. (2017) 'Statistical tools and approaches to validate analytical methods: methodology and practical examples', *International Journal of Metrology and Quality Engineering*, 8, p.9.

Bennett, R. E. (2011) 'Formative assessment: A critical review', *Assessment in Education: Principles, Policy & Practice*, 18(1), pp. 5–25.

Berger, R. (2015) 'Now I see it, now I don't: researcher's position and reflexivity in qualitative research', *Qualitative Research*, 15(2), 219-234. <https://doi.org/10.1177/1468794112468475>

Bi, Z. W. (2022) 'Analysis of the Development Path of College English Knowledge Courses and Skills Teaching — Review of “New Exploration of College English Teaching Method Theory and Practice”', *Technology Enhanced Foreign Language Education*, 2022(02), p. 109.

Black, P., & Wiliam, D. (1998) 'Assessment and classroom learning', *Assessment in Education: Principles, Policy & Practice*, 5(1), pp.7-74.

Black, P. & Wiliam, D. (1998) 'Inside the black box: Raising standards through classroom assessment', *Phi Delta Kappan*, 80(2), pp.139-148.

Blasbalg, U., Hershkowitz, I., Lamb, M.E., Karni-Visel, Y. and Ahern, E.C. (2019) 'Is interviewer support associated with the reduced reluctance and enhanced informativeness of alleged child abuse victims?', *Law and Human Behavior*, 43(2), p.156.

Bode, B. H. (1938) *Progressive Education at the Crossroads*. New York and Chicago: Newson & Company, pp: 107, 113.

Bogi, R., Badu, S., Kojo Rockson, D., & Asebiga, E. (2023) 'Coping with motherhood and education: revealing strategies of student nursing mothers in distance education programs in the upper west region, ghana', *Journal of Advanced Research in Women's Studies*, 1(2), pp. 1-18. <https://doi.org/10.33422/jarws.v1i2.465>

Boud, D., & Falchikov, N. (2007) *Rethinking assessment in higher education*, pp. 191-207. London: Routledge.

Bozkurt, A., Jung, I., Xiao, J., Vladimirsch, V., Schuwer, R., Egorov, G., Lambert, S., Al-Freih, M., Pete, J., Olcott Jr, D. and Rodes, V. (2020) 'A global outlook to the interruption of education due to

COVID-19 pandemic: Navigating in a time of uncertainty and crisis', *Asian Journal of Distance Education*, 15(1), pp.1-126.

Bradbury-Jones C (2007) 'Enhancing rigor in qualitative health research: exploring subjectivity through Peshkin's I's', *Journal of Advanced Nursing*, 59, pp. 290–298.

Broadbent, J. and 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, pp. 1–13.

Broadbent, J. (2018) 'Comparing online and blended learner self-regulated learning strategies and academic performance', *The Internet and Higher Education*, 36, pp. 24–32.

Brookhart, S.M. (2017). *How to give effective feedback to your students*. 2nd ed. Alexandria, VA: ASCD.

Brophy, J. (2004) *Motivating students to learn*. LEA.

Brown, A. L. (1987) 'Metacognition, Executive Control, Self-Regulation, and Other More Mysterious Mechanisms', *Metacognition, Motivation, and Understanding*, pp. 65–116.

Brown, H. D. (2014) *Principles of language learning and teaching: A course in second language acquisition*. Pearson.

Bryman, A. (1984) 'The debate about quantitative and qualitative research: a question of method or epistemology?', *British journal of Sociology*, pp.75-92.

Bryman, A. (2016) *Social research methods*. 5th ed. Oxford: Oxford University Press.

Budiman, A., 2020 'ICT and foreign language learning: An overview', *Tarling: Journal of Language Education*, 3(2), pp.245-267.

Bulut, İ. (2018) 'The levels of classroom and pre-school teachers' metacognitive awareness', *Universal Journal of Educational Research*, 6(12), pp. 2697-2706.  
<https://doi.org/10.13189/ujer.2018.061201>

Buragohain, D., Punpeng, G., Jaratjarungkiat, S. & Chaudhary, S. (2023) 'Impact of e-learning activities on English as a second language proficiency among engineering cohorts of Malaysian higher education: A 7-month longitudinal study', *Informatics*, 10(1), p. 31. MDPI.

Çakiroğlu, Ü., Kokoç, M. & Atabay, M. (2024) 'Online learners' self-regulated learning skills regarding LMS interactions: A profiling study', *Journal of Computing in Higher Education*, 36(1), pp. 220-241.

Carless, D. (2007) 'Learning-oriented assessment: Conceptual bases and practical implications', *Innovations in Education and Teaching International*, 44(1), pp.57-66.

Carter, N., Bryant-Lukosius, D., DiCenso, A., Blythe, J. & Neville, A.J. (2014) 'The use of triangulation in qualitative research', *Oncology Nursing Forum*, 41(5), pp.545-547. DOI: 10.1188/14.ONF.545-547.

Central Committee of the Communist Party of China and the State Council (1999) *Decision on deepening education reform and comprehensively promoting Quality education*. available at: <https://www.waizi.org.cn/law/5058.html> (Accessed: 12 December 2024)

Çekiç, A., & Bakla, A. (2021) 'A review of digital formative assessment tools: Features and future directions', *International Online Journal of Education and Teaching*, 8(3), pp. 1459-1485.

Central People's Government of the People's Republic of China (2023) *Promoting the Steady and Long-term Development of Building a Powerful Country in Education—A Review of the Reform and Development of China's Educational Undertaking in the Past Five Years*, Available at: [https://www.gov.cn/yaowen/liebiao/202309/content\\_6902771.htm](https://www.gov.cn/yaowen/liebiao/202309/content_6902771.htm) (Accessed: November 19, 2023).

Chapelle, C. A. (2006) *English language learning and technology: Lectures on applied linguistics in the age of information and communication technology*. John Benjamins Publishing Company.

- Chen, C., & Bao, L. (2021) 'Reform to Broaden and Deepen the Content of College Entrance Examination: Transform Teaching-to-the-Test into Fruitful Learning', *Journal of China Examinations*, 2021(06), pp. 26 - 36.
- Chen, D. et al. (2023) 'Use of technology-based assessments: a systematic review covering over 30 countries', *Assessment in Education: Principles, Policy & Practice*, 30(5–6), pp. 396–428.  
Doi: 10.1080/0969594X.2023.2270181
- Chen, F. (2022) 'Thoughts on the Learning Pain Points of Higher Vocational College Students', *Popular Literature and Art*, 2022(18), pp. 120 - 122.
- Chen, J. (2022) 'The Complexity Challenges and Path Optimisation for the Balanced Development of Urban and Rural Preschool Education in the New Era—Based on the Complexity Theory', *Social Science Front*, No. 5, pp. 235 - 241.
- Chen, L., & Chen, Y. H. (2020) 'Chinese Experience in Technological Transformation of Education since the Reform and Opening up', *China Educational Technology*, 2020(4), pp. 51 - 59.
- Chen, L., Jiang, R., Mao, W. X., Wen, Y. Y., & Zhang, G. F. (2022) 'On the Starting Point and Development Stages of Educational informatisation in China', *Distant education in China*, 2022(1), pp. 37 - 51.
- Chen, Q. Z. (2018) *History of China's Education*. Arranged by Dong, Z. F. Central China Normal University Press.
- Chen, S. (2022) 'Exploring the Value Orientation of Educational Policies in the Qing Dynasty from the History of the Qing Dynasty—A Review of 'Impression·Chinese History (Qing Dynasty Volume): The Dynasty on the Eve of Modern Times'', *Journal of the Chinese Society of Education*, 2022(08), p. 148.
- Cheung, A., et al. (2021) 'The impact of formative assessment with digital technologies on student learning: A meta-analysis', *Educational Technology Research and Development*, 69(3), pp. 1493-1515.
- Cheng, G. Q., & Niu, M. G. (2021) *History of China's Education*. Nanjing University Press.
- Cheng, R. (2022) 'Reform and Innovation of Legal English Translation Teaching Method in the Context of Smart Education — Review of "New Exploration of Legal English Translation in the New Era"', *Technology Enhanced Foreign Language Education*, 2022(03), p. 98.
- Cheung, R., & Vogel, D. (2013) 'Predicting user acceptance of collaborative technologies: An extension of the technology acceptance model for e-learning', *Computers & education*, 63, pp. 160-175.
- Cheung, S. T., Wang, F. L., & Kwok, L. F. (2021) 'The continuous pursuit of smart learning', *Australasian Journal of Educational Technology*, 37(2), pp. 1-6.
- Cheng, Y. M. (2011) 'Antecedents and consequences of e-learning acceptance', *Information Systems Journal*, 21(3), pp. 269-299.
- Chick, R.C., Clifton, G.T., Peace, K.M., Propper, B.W., Hale, D.F., Alseidi, A.A. and Vreeland, T.J. (2020) 'Using technology to maintain the education of residents during the COVID-19 pandemic', *Journal of surgical education*, 77(4), pp.729-732.
- Cladera, M. (2021) 'An application of importance-performance analysis to students' evaluation of teaching', *Educational Assessment, Evaluation and Accountability*, 33(4), pp. 701-715.
- Clément, B., Sauzéon, H., Roy, D. & Oudeyer, P. Y. (2024) 'Improved performances and motivation in intelligent tutoring systems: Combining machine learning and learner choice', *arXiv preprint arXiv:2402.01669*
- Creswell, J. W. (2009) *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*. Thousand oaks, CA: sage.
- Creswell, J. W., & Plano Clark, V. L. (2011) *Designing and conducting mixed research* (2<sup>nd</sup> ed.). Thousand Oaks, CA: Sage.
- Creswell, J.W. (2014) *Research design: Qualitative, quantitative, and mixed methods approaches*. 4th ed. Thousand Oaks, CA: Sage.

- Creswell, J. W. and Creswell, J.D. (2017) *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.
- Creswell, J.W. and Plano Clark, V.L. (2017) *Designing and conducting mixed methods research*. 3rd ed. Thousand Oaks, CA: Sage Publications.
- Creswell, J. W., & Plano Clark, V. L. (2018) *Designing and Conducting Mixed Methods Research (3rd ed.)*. Sage Publications.
- Creswell, J.W. (2021) *A concise introduction to mixed methods research*. SAGE publications.
- Criollo-C, S., Guerrero-Arias, A., Jaramillo-Alcázar, Á. & Luján-Mora, S. (2021) 'Mobile learning technologies for education: Benefits and pending issues', *Applied Sciences*, 11(9), p.4111.
- Crossley, S. A., Greenfield, J., & McNamara, D. S. (2008) 'Assessing Text Readability Using Cognitively Based Indices', *TESOL Quarterly*, 42(3), pp.475-493.
- Crossley, S.A. and McNamara, D.S. (2012) 'Predicting second language writing proficiency: the roles of cohesion and linguistic sophistication', *Journal of Research in Reading*, 35(2), pp.115-135.
- Crossley, S. A., Allen, L. K., & McNamara, D. S. (2012b) 'Text Simplification and Cohesion: The Influence of Connectives and Lexical Diversity on Text Comprehension', in *Toward Automated Assessment of Language Proficiency*, eds. D. S. McNamara & S. Crossley.
- Cyberspace Administration of China (2022) '*The State Internet Information Office Releases the Digital China Development Report (2021)*'. Available at: [http://www.cac.gov.cn/2022 - 08/02/c\\_1661066515613922.htm](http://www.cac.gov.cn/2022 - 08/02/c_1661066515613922.htm) (Accessed: 30 March 2022).
- Dabbagh, N., & Kitsantas, A. (2012) 'Personal Learning Environments, Social Media, and Self-Regulated Learning: A Natural Formula for Connecting Formal and Informal Learning', *The Internet and Higher Education*, 15(1), pp. 3-8.
- Dağal, A. and Bayındır, D. (2016) 'The investigation of the relationship between the level of metacognitive awareness, self-directed learning readiness and academic achievement of preschool teacher candidates', *Universal Journal of Educational Research*, 4(11), pp. 2533-2540.  
<https://doi.org/10.13189/ujer.2016.041106>
- Dai, Y., & Chen, J. W. (2021) 'The Historical Evolution and Realistic Review of the Authority of Teachers in China', *Teacher Education Research*, 2021(33)(03), pp. 24 - 30.
- Davis, F. D. (1989) 'Perceived usefulness, perceived ease of use, and user acceptance of information technology', *MIS Quarterly*, 13(3), pp. 319-340.
- Dawadi, S., Shrestha, S. and Giri, R.A. (2021) 'Mixed-methods research: A discussion on its types, challenges, and criticisms', *Journal of Practical Studies in Education*, 2(2), pp.25-36.
- De Koning, R., Egiz, A., Kotecha, J., Ciuculete, A.C., Ooi, S.Z.Y., Bankole, N.D.A., Erhabor, J., Higginbotham, G., Khan, M., Dalle, D.U. and Sichimba, D. (2021), 'Survey fatigue during the COVID-19 pandemic: an analysis of neurosurgery survey response rates', *Frontiers in surgery*, 8, p.690680.
- De Villiers, B., & Werner, A. (2016) 'The relationship between student engagement and academic success', *Journal for New Generation Sciences*, 14(1), pp. 36-50.
- Deci, E. L. & Ryan, R. M. (1985) *Intrinsic motivation and self-determination in human behavior*. Springer Science & Business Media.
- Deci, E. L. and Ryan, R. M. (2000) 'The "what" and "why" of goal pursuits: human needs and the self-determination of behavior', *Psychological Inquiry*, 11(4), pp. 227-268.  
[https://doi.org/10.1207/s15327965pli1104\\_01](https://doi.org/10.1207/s15327965pli1104_01)
- Deeva, G., Bogdanova, D., Serral, E., Snoeck, M. & De Weerd, J. (2021) 'A review of automated feedback systems for learners: Classification framework, challenges and opportunities', *Computers & Education*, 162, p. 104094.
- Delve, Ho, L. and Limpaecher, A. (2023) *What is researcher triangulation in qualitative analysis?* Available at: <https://delvetool.com/blog/researcher-triangulation> (Accessed: 16 May 2024).

- Denscombe, M. (2017) *EBOOK: The good research guide: For small-scale social research projects*. McGraw-Hill Education (UK).p.3.
- Dewey, J. (1980) 'Democracy and Education', in Boydston, J. A. (ed.), *The Middle Works of John Dewey, 1899—1924* (vol. 9), Carbondale and Edwardsville: Southern Illinois University Press, p. 89.
- Dewey, J. (2014) *Problems of Men*. Translated by Fu, T. X. et al. Shanghai: Shanghai People's Publishing House, p. 29.
- Dey, I. (2003) *Qualitative data analysis: A user friendly guide for social scientists*. Routledge.
- Di, X., Ismail, W. M., Zailani, M. A., & Ismail, Z. B. (2022) 'A study of correlation between metacognitive strategy, self-efficacy and arabic achievements of university level students in malaysian public universities', *Al-Dad Journal*, 6(2), pp. 1-12. <https://doi.org/10.22452/aldad.vol6no2.1>
- Ding, Y. J. (2022) 'Community of Education and Training: An Exploration of the German Vocational Education and Continuing Education Alliance', *Education Research Monthly*, 2022(03), pp. 19 - 27.
- D'Mello, S. K., & Graesser, A. C. (2014) '31 Feeling, Thinking, and Computing with Affect-Aware Learning Technologies', *The Oxford handbook of affective computing*, p. 419.
- Dong, H. J. (2021) 'Home-school Collaborative Education under the Background of Double Reduction', *Journal of the Chinese Society of Education*, 2021(S2), pp. 196 - 199.
- Dong, G. W., Zhao, G. Q., Zheng, C., Song, L. B., Wang, X. H. (2021) 'Design and Practice of Virtual Simulation Experiment Teaching Mode Based on Rain Class', *Research and Exploration in Laboratory*, 2021(40)(10), pp. 215 - 218.
- Dong, L., Jamal Mohammed, S., Ahmed Abdel-Al Ibrahim, K. and Rezai, A. (2022) 'Fostering EFL learners' motivation, anxiety, and self-efficacy through computer-assisted language learning-and mobile-assisted language learning-based instructions', *Frontiers in psychology*, 13, p.899557.
- Dong, Y. (2010) 'A Brief Discussion on the Implementation Approaches of Social Education in Ancient Chinese Society', *Agricultural Archaeology*, 2010, (04), pp. 307 - 310.
- Dörnyei, Z. (2001). *Motivational strategies in the language classroom*. Cambridge University Press.
- Dörnyei, Z. (2014). *The psychology of the language learner: Individual differences in second language acquisition*. Routledge.
- Drane, C., Vernon, L. & O'Shea, S. (2020) 'The impact of 'learning at home' on the educational outcomes of vulnerable children in Australia during the COVID-19 pandemic', *Literature Review Prepared by the National Centre for Student Equity in Higher Education*. Curtin University, Australia.
- Drigas, A., Dede, D. E., & Dedes, S. (2020) 'Mobile and other applications for mental imagery to improve learning disabilities and mental health', *International Journal of Computer Science Issues (IJCSI)*, 17(4), pp. 18-23.
- 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), pp. 1929-1958.
- Duran, N. D., & McCarthy, P. M. (2010) 'Lexical Diversity and Automatic Essay Scoring: How Lexical Measures Contribute to Assessments of Text Quality', *Journal of Language and Social Psychology*, 29(4), pp.362-373.
- Editorial Board of Education Research (2022) '2021 Annual Report on the Frontiers and Hot Issues of Education Research in China', *China Educational Technology*, Vol. 2, No. 505, pp. 57 - 69.
- Egbert, J. (1999) 'CALL environments: Research, practice, and critical issues', *Teachers of English to Speakers of Other Languages, Inc.*
- Ellis, R. (2003). *Task-Based Language Learning and Teaching*. Oxford: Oxford University Press.
- Ellis, R. (2017) 'Task-based language teaching', in: Loewen, S. and Sato, M., eds. *The Routledge Handbook of Instructed Second Language Acquisition*. London: Routledge, pp. 108–125.
- Elmahdi, I., Al-Hattami, A. and Fawzi, H. (2018) 'Using Technology for Formative Assessment to

- Improve Students' Learning', *Turkish Online Journal of Educational Technology*, 17(2), pp. 182–188.
- Elmi, H., Ambiyar, A., Huda, Y., & Novaliendry, D. (2024) 'The Role of Information and Communication Technology in Interactive Learning', *Jurnal SAINTIKOM (Jurnal Sains Manajemen Informatika dan Komputer)*, 23(1), pp.193-203.
- Enosh, G., Tzafrir, S. S., & Stolovy, T. (2014) 'The development of client violence questionnaire (CVQ)', *Journal of Mixed Methods Research*, 9(3), 273–290. <https://doi.org/10.1177/1558689814525263>
- Eyadat, H.M.M. (2020) 'The challenges of using technology in vocational education and their impact on students' achievement from the teachers' point of view in Ramtha District schools in Jordan', *International Journal of Education and Practice*, 8(4), pp. 678-693.
- Fass-Holmes, B. (2022) 'Survey Fatigue--What Is Its Role in Undergraduates' Survey Participation and Response Rates?', *Journal of interdisciplinary Studies in Education*, 11(1), pp.56-73.
- Fauzi, S. S. M. and Hussain, R. M. R. (2016) 'Designing instruction for active and reflective learners in the flipped classroom', *Malaysian Journal of Learning and Instruction*, Vol. 13, 147-173. <https://doi.org/10.32890/mjli2016.13.2.6>
- Feijóo, C., Arenal, A., Armuña, C., Ramos, S., & Fernández, J. (2021) *Educational Technologies in China: Pre-and Post-Pandemic Lessons*, European Commission. Available from: EU Bookshop.
- Feng, J.J. (2019) 'Two Problems of China's Educational Philosophy Development in the Past 100 Years', *Journal of Higher Education*, 2019, 40(09), pp. 46 - 52.
- Feng, J. J. (2020) *Life and Education*. Educational Science Press.
- Fetters, M.D. and Freshwater, D. (2015) 'The 1+ 1= 3 Integration Challenge', *Journal of mixed methods research*, 9(2), pp.115-117.
- Fetters, M.D. (2016.) 'Haven't we always been doing mixed methods research? Lessons learned from the development of the horseless carriage', *Journal of Mixed Methods Research*, 10(1), pp.3-11. <https://doi.org/10.1177/1558689815620883>
- Flavell, J. H. (1979) 'Metacognition and Cognitive Monitoring: A New Area of Cognitive–Developmental Inquiry', *American Psychologist*, 34(10), pp. 906–911.
- Fontana, A., & Frey, J. H. (2005). 'The interview', *The Sage handbook of qualitative research*, 3(1), 695-727.
- Francom, G. M. (2019) 'Barriers to technology integration: A time-series survey study', *Journal of Research on Technology in Education*, 52(1), pp. 1–16. doi: 10.1080/15391523.2019.1679055.
- Gamla, W. (2021) 'Readability of Reading Passages in the General Aptitude Test (GAT) and English Textbooks for the Upper Secondary Level in Thailand by Using Coh-Metrix', *LEARN Journal: Language Education and Acquisition Research Network*, 14(1), pp.130-151.
- Gao, S.G. (2020) 'The Institutional Advantages of Socialist Education with Chinese Characteristics', *People's Education*, 2020, (01), pp. 23 - 27.
- García-Jiménez, E., Gallego-Noche, B. and Gómez-Ruiz, M.Á. (2015) 'Feedback and self-regulated learning: How feedback can contribute to increase students' autonomy as learners', in Peris-Ortiz, M. and Merigó-Lindahl, J.M. (eds.) *Sustainable learning in higher education: Developing competencies for the global marketplace*. Cham: Springer, pp. 113–130.
- Ge, S. L., Wang, N. X., Yin, J., & Li, W. C. (2022) 'Construction of Professional Talent Training Mode in Industry-characteristic Universities', *China Higher Education*, 2022(06), pp. 26 - 28.
- Ghanizadeh, A., Razavi, A., & Jahedizadeh, S. (2015) 'Technology-enhanced language learning (TELL): A review of resources and upshots', *International Letters of Chemistry, Physics and Astronomy*, 54, pp. 73-87.
- Gikandi, J. W., Morrow, D., & Davis, N. E. (2011) 'Online formative assessment in higher education: A review of the literature', *Computers & education*, 57(4), pp. 2333-2351.
- Goin Kono, K., & Taylor, S. (2021) 'Using an ethos of care to bridge the digital divide: Exploring

faculty narratives during a global pandemic', *Online Learning*. DOI: 10.24059/OLJ.V25I1.2484

González-Lloret, M. (2019) 'Task-based language teaching and L2 pragmatics: Benefits and challenges', *Annual Review of Applied Linguistics*, 39, pp. 137-152. Available at: Academia.edu.

Gonzalez, T., De La Rubia, M.A., Hincz, K.P., Comas-Lopez, M., Subirats, L., Fort, S. & Sacha, G.M. (2020) 'Influence of COVID-19 confinement on students' performance in higher education', *PLOS ONE*, 15(10), p.e0239490.

Glesne, C. (1999) *Becoming Qualitative researcher: An Introduction*, 2nd edition, New York: Longman.

Golonka, E.M., Bowles, A.R., Frank, V.M., Richardson, D.L. and Freynik, S. (2014) 'Technologies for foreign language learning: A review of technology types and their effectiveness', *Computer assisted language learning*, 27(1), pp.70-105.

Gong, X. & Xiang, C. (2022) 'Study on the effectiveness of college oral English teaching in independent colleges under the "3+1" blended teaching reform based on the FiF oral training platform', *Overseas English*, (12), pp.140-141+154.

Gong, X., Yu, X. & Ding, T. (2017). 'Analysis of differences in autonomous learning abilities among different types of undergraduates: A survey of 859 undergraduates in Jiangsu Province', *China University Teaching*, (9), pp. 91–96.

Gorski, P. (2005) 'Education equity and the digital divide', *AACE Review (Formerly AACE Journal)*, 13(1), pp. 3-45.

Γούδας, M., Dermitzaki, I., & Kolovelonis, A. (2015) 'Self-regulated learning and students' metacognitive feelings in physical education', *International Journal of Sport and Exercise Psychology*, 15(2), pp. 131-145. <https://doi.org/10.1080/1612197x.2015.1079791>

Graesser, A.C., McNamara, D.S., Louwerse, M.M. and Cai, Z. (2004) 'Coh-Metrix: Analysis of text on cohesion and language', *Behavior Research Methods, Instruments, and Computers*, 36, pp.193–202.

Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011) 'Coh-Metrix: Providing Multilevel Analyses of Text Characteristics', *Educational Researcher*, 40(5), pp.223-234.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989) 'Toward a conceptual framework for mixed-method evaluation designs', *Educational Evaluation and Policy Analysis*, 11(3), 255-274. <https://doi.org/10.3102/01623737011003255>.

Guerra Ayala, M. J., Reynosa Navarro, E., Durand Gómez, E. L., Florez Lucana, A., Chambi Catacora, M. A. d. P., Vargas Onofre, E., ... & Cari Checa, E. (2024) 'Enjoyment and oral english proficiency in future teachers', *Frontiers in Education*, 9. <https://doi.org/10.3389/feduc.2024.1306080>

Guillemin M, Gillam L (2004) 'Ethics, reflexivity and 'ethically important moments' in research', *Qualitative Inquiry*, 10, pp. 261–280.

Guan, H. Y. (2022) 'Innovation of Teaching Methods in Colleges and Universities under the Modern Educational Concept', *Teaching Reference of Middle School Politics*, 2022(01), p. 83.

Gunuc, S. (2014) 'The relationships between student engagement and their academic achievement', *International Journal on New Trends in Education and Their Implications*, 5(4), pp.216-231.

Guo, B. W. (2017) *History of the Evolution of Chinese Educational System*. The Commercial Press.

Guo, J. and Tan, X. (2015) 'Analysis of the changing trends in the quality of vocational college student sources and teaching strategy research', *Chinese Adult Education*, (15), pp. 71–73.

Guo, Y., & Lin, L. J. (2020) 'The Chinese Tradition of Respecting Teachers in the Sacrificial Culture of Ancient Academies', *The Teaching of Thought and Political Study*, 2020(01), pp. 8 - 11.

Gwet, K. (2001) *Handbook of inter-rater reliability*. Gaithersburg, MD: STATAXIS Publishing Company, pp.223-246.

Gwet, K.L. (2008) 'Computing inter-rater reliability and its variance in the presence of high agreement', *British Journal of Mathematical and Statistical Psychology*, 61(1), pp.29-48.

- Hampden-Turner, C. & Trompenaars, F. (2020) *Riding the waves of culture: Understanding diversity in global business*. Hachette UK.
- Han, F. & Ellis, R. A. (2021) 'Patterns of student collaborative learning in blended course designs based on their learning orientations: A student approaches to learning perspective', *International Journal of Educational Technology in Higher Education*, 18(1), p. 66.
- Han, X. Z., & Song, Z. Q. (2021) 'An Analysis of Xi Jinping's Exposition on Teaching and Educating in the New Era', *Studies in Ideological Education*, 2021(11), pp. 3 - 7.
- Hao, X., & Jiang, C. Y. (2020) 'The Origin, Connotation and Contemporary Enlightenment of Huang Yanpei's Vocational Education Thoughts', *Hebei Vocational Education*, 5, pp. 36 - 40.
- Harding, J., Patel, S., Davis, T., Patzer, R., McDonald, B., Walker-Williams, D. & Gander, J. (2022) 'Understanding racial disparities in covid-19-related complications: protocol for a mixed methods study', *Jmir Research Protocols*, 11(10), e38914. <https://doi.org/10.2196/38914>
- Harris, L. R. & Brown, G. T. (2018) *Using self-assessment to improve student learning*. Routledge. Available at: <https://doi.org/10.4324/9781351036979>
- Hart, C. (1998). *Doing a literature review: Releasing the social science research imagination*. London: Sage.
- Hartley, R.I. and Sturm, P. (1997) 'Triangulation', *Computer vision and image understanding*, 68(2), pp.146-157.
- Harzing, A.W. & Alakangas, S. (2016) 'Google Scholar, Scopus and the Web of Science: a longitudinal and cross-disciplinary comparison', *Scientometrics*, 106(2), pp.787–804.
- Hastings, J. T. and Madaus, G. F. (1971) *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill.
- Hatem, G., Zeidan, J., Goossens, M., & Moreira, C. (2022) 'Normality testing methods and the importance of skewness and kurtosis in statistical analysis', *BAU Journal-Science and Technology*, 3(2), p. 7.
- Hattie, J., & Timperley, H. (2007) 'The power of feedback', *Review of Educational Research*, 77(1), pp.81-112.
- Haywood, H. C., & Lidz, C. S. (2007) *Dynamic assessment in practice: Clinical and educational applications*. Cambridge: Cambridge University Press.
- Hazaymeh, W. A. (2021) 'EFL Students' Perceptions of Online Distance Learning for Enhancing English Language Learning during COVID-19 Pandemic', *International Journal of Instruction*, 14(3), pp. 501-518.
- Heale, R. and Forbes, D. (2013) 'Understanding triangulation in research', *Evidence-based nursing*, 16(4), pp.98-98.
- Hennink, M., Hutter, I., & Bailey, A. (2020) *Qualitative research methods*. Sage.
- Hew, K.F. and Brush, T. (2007) 'Integrating technology into K-12 teaching and learning: Current knowledge gaps and recommendations for future research', *Educational technology research and development*, 55, pp.223-252.
- Heyvaert, M., Hannes, K. and Onghena, P. (2016) *Using mixed methods research synthesis for literature reviews: the mixed methods research synthesis approach* (Vol. 4). Sage Publications.
- Higher Education Business Department of IFlytek Co., LTD. (2020) *FiF oral learning system construction scheme*.
- Holden, R. J. & Karsh, B. T. (2010) 'The Technology Acceptance Model: Its past and its future in health care', *Journal of Biomedical Informatics*, 43(1), pp. 159-172.
- Horwitz, E.K., Horwitz, M.B. and Cope, J. (1986) 'Foreign language classroom anxiety', *The Modern Language Journal*, 70(2), pp. 125–132.
- Hu, J. B. (2022) 'A Case Study of Paradigm Transformation with Student-centeredness in the Application-oriented University', *Journal of Higher Education*, 2022(42)(11), pp. 57 - 68.

- Hu, P. (2021) 'The Application of Ancient Chinese Educational Methods in Teaching Chinese as a Foreign Language', *International Chinese Culture Studies*, 2021(00), pp. 3 - 9.
- Hu, W. S., & Ye, L. (2022) 'Historical Evolution, Practical Basis and Development Trend of Integration of Industry and Education in Higher Vocational Education', *Modern Education Management*, 2022(10), pp. 82 - 89.
- Hu, Y., Wu, Q., & Lin, K. S. (2022) 'Practical Resistance and Promotion Path of the General and Vocational Diversion in the High School Entrance Examination', *Education and Vocation*, 2022(5), pp. 35 - 40.
- Hua, Y. H., Yang, M., & Lü, J. (2022) 'Current Situation and Countermeasures of General and Vocational Diversion from the Perspective of Educational Equity', *New West*, 2022(08), pp. 133 - 135.
- Huang, D. R. (2022) 'Reform of Higher Education Evaluation in the New Era: Characteristics and Reflections', *Journal of Hebei Normal University (Educational Science)*, 2022, 24(03), pp. 1 - 4 + 149.
- Hwang, G. J., & Chang, H. F. (2011) 'A formative assessment-based mobile learning approach to improving the learning attitudes and achievements of students', *Computers & Education*, 56(4), pp. 1023-1031.
- Huang, X. L. (2020) 'The Current Situation, Characteristics, Problems and Improvements of School-based Curriculum Learning Evaluation', *Journal of Teaching and Management*, 2020, (06), pp. 67 - 70.
- Huang, Y. (2020) 'Application of the FiF oral training system in the teaching of English listening and speaking for art and physical education students in colleges based on the input-output theory', *Overseas English*, (14), pp.131-132.
- Hwang, G. J. (2014) 'Definition, framework and research issues of smart learning environments - a context-aware ubiquitous learning perspective', *Smart Learning Environments*, 1(1).
- Illeris, K. (2018a) 'An overview of the history of learning theory', *European Journal of Education*, 53(1), pp. 86-101.
- Illeris, K. (2018b) 'The development of a comprehensive and coherent theory of learning', In *Learning, Development and Education* (pp. 21-32). Routledge.
- In J. (2017) 'Introduction of a pilot study', *Korean journal of anesthesiology*, 70(6), pp. 601–605. <https://doi.org/10.4097/kjae.2017.70.6.601>
- Jiang, C. (2022) 'Value, Premise and Risks of the Directed Cultivation of Rural Teachers', *Journal of Sichuan Normal University (Social Sciences Edition)*, 49 (03), pp. 114 - 121.
- Jiang, J. and Chen, D. (2021) 'Research on automated scoring of subjective items: Review, reflections and prospects', *Chinese Journal of Foreign Languages*, 18(6), pp. 58–64.
- Jiang, M. M. (2020) 'Talent Cultivation Returns to Its Original Orientation: Rethinking the Purpose and Function of Education in the New Era', *Studies on Mao Zedong and Deng Xiaoping Theories*, 6, pp. 32 - 39 + 108.
- Jin, S. H., Im, K., Yoo, M., Roll, I. & Seo, K. (2023) 'Supporting students' self-regulated learning in online learning using artificial intelligence applications', *International Journal of Educational Technology in Higher Education*, 20(1), p. 37.
- Jin, X. Z., & Sha, Q. B. (2022) 'The Application of "Project + Ladder + Inquiry" Teaching Method in the Teaching of "Fluid Pressure and Flow Velocity Relationship"', *Physics Teacher*, 43(09), pp. 43 - 45.
- Jonassen, D. H. & Rohrer-Murphy, L. (1999) 'Activity theory as a framework for designing constructivist learning environments', *Educational Technology Research and Development*, 47(1), pp. 61-79.
- Jones, J. M. (2017) 'Scaffolding self-regulated learning through student-generated quizzes', *Active Learning in Higher Education*, 20(2), pp. 115-126. <https://doi.org/10.1177/1469787417735610>
- Johnson, D. W. & Johnson, R. T. (1989) *Cooperation and competition: Theory and research*.

Interaction Book Company.

Kakungulu, S.J.(2024) 'The role of artificial intelligence in educational assessment', *EURASIAN EXPERIMENT JOURNAL OF SCIENTIFIC AND APPLIED RESEARCH (EEJSAR)*, 5(2), pp. 44-48.

Karagiozis, N. (2018) 'Complexities of the Researcher's Role', *The International Journal of Interdisciplinary Educational Studies*, 13(1), pp. 19-31.

Kaya-Capocci, S., O'Leary, M., & Costello, E. (2022) 'Towards a framework to support the implementation of digital formative assessment in higher education', *Education Sciences*, 12(11), p. 823.

Keshf, Z. and Khanum, S. (2021) *Career guidance and counseling needs in a developing country's context: a qualitative study*. Sage Open, 11(3), 215824402110401.  
<https://doi.org/10.1177/21582440211040119>

King, W. R., & He, J. (2006) 'A meta-analysis of the technology acceptance model', *Information & Management*, 43(6), pp. 740-755.

Kirschner, P. A., Sweller, J. & Clark, R. E. (2010) 'Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching', *Educational Psychologist*, 41(2), pp. 75-86.

Kistner, S., Rakoczy, K., Otto, B., Dignath, C., Büttner, G., & Klieme, E. (2010) 'Promotion of self-regulated learning in classrooms: investigating frequency, quality, and consequences for student performance', *Metacognition and Learning*, 5(2), pp. 157-171. <https://doi.org/10.1007/s11409-010-9055-3>

Kleijnen, J.P. (1999) 'Validation of models: statistical techniques and data availability',  
In *Proceedings of the 31st conference on Winter simulation: Simulation---a bridge to the future-Volume 1*, pp. 647-654.

Korstjens, I. and Moser, A. (2017) 'Series: practical guidance to qualitative research. part 4: trustworthiness and publishing', *European Journal of General Practice*, 24(1), pp. 120-124.  
<https://doi.org/10.1080/13814788.2017.1375092>

Krashen, S. (1998) *Second Language Acquisition and Second Language Learning*. Prentice-Hall International.

Kuddus, K. (2022) 'Artificial intelligence in language learning: Practices and prospects', *Advanced Analytics and Deep Learning Models*, pp. 1-17.

Kumar, H., Xiao, R., Lawson, B., Musabirov, I., Shi, J., Wang, X. & Liut, M. (2024) 'Supporting self-reflection at scale with large language models: Insights from randomized field experiments in classrooms', In *Proceedings of the Eleventh ACM Conference on Learning@ Scale*, pp. 86-97.

Kusz, E. (2023) 'Effects of self-imitation practice on L2 pronunciation with the use of a Golden Speaker Builder – an interactive tool for pronunciation training', *Pronunciation in Second Language Learning and Teaching Proceedings*, 13(1).

Lantolf, J.P., & Poehner, M.E. (2007) 'Dynamic assessment of L2 development: bringing the past into the future', *Journal of Applied Linguistics*, 1(1).

Lave, J. (1991) *Situated learning: Legitimate peripheral participation*. Cambridge University Press.

Lavrysh, Y., Leshchenko, M., & Tymchuk, L. (2023) 'Development of metacognitive skills through digital narratives in higher education', *Revista Electrónica De Investigación Educativa*, 25, pp. 1-15.  
<https://doi.org/10.24320/redie.2023.25.e07.5028>

Lavoie, J., Wyman, J., Crossman, A., & Talwar, V. (2021) 'Meta-analysis of the effects of two interviewing practices on children's disclosures of sensitive information: rapport practices and question type', *Child Abuse & Neglect*, 113, 104930. <https://doi.org/10.1016/j.chiabu.2021.104930>

Lawless, K.A. and Pellegrino, J.W. (2007) 'Professional development in integrating technology into teaching and learning: Knowns, unknowns, and ways to pursue better questions and answers', *Review of educational research*, 77(4), pp.575-614.

Lefever, S., Dal, M. and Matthíasdóttir, Á. (2007) 'Online data collection in academic research:

- advantages and limitations', *British journal of educational technology*, 38(4), pp.574-582.
- Legris, P., Ingham, J., & Collette, P. (2003) 'Why do people use information technology? A critical review of the technology acceptance model', *Information & Management*, 40(3), pp. 191-204.
- Lei, H. & Huang, M. (2024) 'Volitional control in self-regulated learning: Key factor affecting the learning outcomes of undergraduate students receiving national scholarships', *Research in Higher Engineering Education*, (4), pp. 159–165.
- Lei, H., Cui, Y., & Zhou, W. (2018) 'Relationships between student engagement and academic achievement: A meta-analysis', *Social Behavior and Personality: An International Journal*, 46(3), pp. 517-528.
- Levy, M. and Hubbard, P. (2005) 'Why call CALL "CALL"?', *Computer Assisted Language Learning*, 18(3), pp. 143-149.
- Li, B. Q., Luan, Z.Y. & Huang, Y.H. (2022) 'A Re-recognition of the Logic, Deduction of the Implication and Exploration of the Realm of Confucius' Heuristic Teaching', *Educational Science Research*, 2, pp. 5 - 11.
- 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, pp. 183-196.
- Li, F., Fan, S. & Wang, Y. (2022) 'Mobile-assisted language learning in the Chinese higher education context: A systematic review from the perspective of the situated learning theory', *Education and Information Technologies*, 27(7), pp. 9665-9688.
- Li, H.W. & Chen, L. (2015) 'The Historical Law of Chinese Ancient School Education's Inheriting and Innovating Chinese Culture', *Educational Research*, 2015, 36(10), pp. 119 - 125.
- Li, M. (2022) 'Research on College English Knowledge Courses and Skills Teaching — Review of "New Exploration of College English Teaching Method Theory and Practice"', *China University Science & Technology*, 6, pp. 109.
- Li, Q. R., & Liu, L. (2022) 'Research on the English Learning Motivation of Higher Vocational College Students in Yunnan Province', *Journal of Yunnan Agricultural University (Social Science)*, 16 (01), pp. 156 - 160.
- Li, S. L., & Wu, N. Z. (2021) 'Logical Starting Point, Formation Mechanism and Support Conditions of the Talent Training Mode of Simultaneous Education and Training in Vocational Education', *Education and Vocation*, 23, pp. 27 - 34.
- Li, W. (2021) 'A Study on the Current Situation of Autonomous Learning among Vocational College Students', *China Vocational and Technical Education*, (29), pp. 66-74.
- Li, X., & Teng, D. (2022) 'Construction of the Smart Education Service Ecological System in Higher Vocational Colleges in the Information 2.0 Era', *Education and Vocation*, 2022(11), pp. 28 - 34.
- Li, X. (2006) 'A study on the autonomous learning ability of English-major undergraduates', *Foreign Language Teaching*, (4), pp. 83–86.
- Li, Y.-l. and Yan, Y.-h. (2012) 'Research on automatic short answer scoring in spoken English tests based on multiple features', *Journal of Electronics & Information Technology*, 34(9), pp. 2097–2102.
- Li, Z. (2021) 'The Characteristics, Trends and Expectation of the Educational History in China', *Education Sciences in China (In Chinese and English)*, 2021, 4(02), pp. 55 - 65.
- Lian, X. & Ji, A. (2017) 'An Investigation into the Learning Motivation of Students in Higher Vocational Colleges: A Case Study of Jincheng Vo-Tech College', *Educational Theory and Practice*, 37(9), pp. 18-19.
- Liang, Z. (2015) 'Research on the characteristics and influencing factors of online learning among vocational college students', *Chinese Vocational and Technical Education*, (26), pp. 5–14.
- Lidz, C. S., & Gindis, B. (2003) 'Dynamic assessment of the evolving cognitive functions in

children’, in: A. Kozulin, B. Gindis, V. S. Ageyev, & S. M. Miller, eds. *Vygotsky's educational theory in cultural context*. Cambridge: Cambridge University Press, pp.99-116.

Lin, J. W., Tsai, C. W., & Hsu, C. C. (2023) ‘A comparison of computer-based and game-based formative assessments: a long-term experiment’, *Interactive Learning Environments*, 31(2), pp. 938-954.

Lin, X. B., Yu, J. B., & Wei, D. K. (2022) ‘Research on the High-quality Development of Industry-education Integration in Higher Vocational Colleges in the Yellow River Basin’, *Education and Vocation*, 18, pp. 13 - 19.

Lin, Y.-L., Wang, W.-T. and Hsieh, M.-J. (2024) ‘The effects of students’ self-efficacy, self-regulated learning strategy, perceived and actual learning effectiveness: A digital game-based learning system’, *Education and Information Technologies*, pp. 1–33.

Ling, M. Y. (2018) ‘The Symptoms of Exam-oriented Education and the Exploration of Countermeasures under the Background of Educational Reform’, *Journal of Teaching and Management*, 30, pp. 41 - 44.

Liu, C. X., & Liu, Y. (2021) ‘Research on the Development and Promotion Strategy of the Integration of Production and Education in Higher Vocational Education’, *Theory and Practice of Education*, 21, pp. 20 - 23.

Liu, D. J., Du, J., Jiang, N., & Huang, R. H. (2018) ‘Trends in Reshaping Education with Artificial Intelligence’, *Open Education Research*, 24 (04), pp. 33 - 42.

Liu, H. F. (2019) ‘Cultivating the Moral Foundation of Respecting Teachers and Valuing Education’, *People’s Tribune*, 19, pp. 140 - 141.

Liu, H. F., & Wei, H. F. (2021) ‘Drawing Lessons from Ancient Times in the Reform of College Admissions: A Historical Investigation into “Scores Only” and “College Enrollment Rates Only”’, *Educational Research*, 42 (05), pp. 86 - 100.

Liu, J. (2021) ‘Bridging digital divide amidst educational change for socially inclusive learning during the COVID-19 pandemic’, *Sage Open*, 11(4), 21582440211060810.

Liu, L. & Wang, Q. (2021) ‘Research on the Characteristics and Influencing Factors of Students’ Learning Engagement in Blended Teaching Model’, *Modern Educational Technology*, 31(11), pp. 80 - 86.

Liu, P. (2021) ‘On Parent Participation in School Education Supervision’. *Journal of Shanghai Educational Research*, 7, pp. 49 - 52.

Liu, T.F. (2022) ‘Into the Discourses of Life: The Study of the Analects of Confucius as a Life Climbing’, *University Education Science*, 5, pp. 23 - 30.

Liu, T.F. & Sun, Y.Y. (2020) ‘My way is consistent: Confucius's way of enlightenment and its contemporary significance’, *Journal of Jishou University (Social Sciences)*, 41 (05), pp. 88 - 96.

Liu, T. F., & Yang, L. P. (2017) ‘Study on the Scale Compilation and Characteristics of Learning Flow Experience of Higher Vocational College Students’, *Vocational and Technical Education*, 38 (13), pp. 56 - 62.

Liu, X. H. (2021) ‘How to Be a Teacher: The Changes and Experiences of Teachers' Authority’, *Forum on Contemporary Education*, 5, pp. 95 - 100.

Liu, Y. B., & Li, Y. (2022) ‘Revival of “Respecting Teachers and Education”: Contemporary Value and Its Path’, *Contemporary Education Sciences*, 5, pp. 56 - 62.

Liu, Z. J., & Xu, B. (2020) ‘Comprehensive Quality Evaluation: The Key to Eliminating “Scores-Only” Evaluation’, *Educational Research*, 41 (02), pp. 91 - 100.

Liu, Z. J., & Xu, B. (2022) ‘The Reform of Curriculum and Teaching Evaluation Methods Under the New Curriculum Standards: Challenges and Responses’, *Curriculum, Teaching Material and Method*, 42 (08), pp. 4-10 + 24.

Loeb, S., Dynarski, S., McFarland, D., Morris, P., Reardon, S., & Reber, S. (2017) ‘Descriptive Analysis in Education: A Guide for Researchers’, NCEE 2017-4023. *National Center for Education Evaluation and*

Lounis, M., Dendani, B. & Bahi, H. (2024) 'Mispronunciation detection and diagnosis using deep neural networks: A systematic review', *Multimedia Tools and Applications*, pp. 1-35.

Louwerse, M.M. (2001) 'An analytic and cognitive parameterisation of coherence relations', *Cognitive Linguistics*, 12, pp.291–315.

Lu, M. (2015) 'The exploration and practice of computerized automatic scoring in large-scale English oral tests', *Journal of China Examinations*, (10), pp. 51–57.

Lu, J. H. (2015) 'A Study about the Analysis of Nonintellectual Acting Factors and Strategies Improving of College Students' Self-regulated Learning', In *2nd Annual International Conference on Social Science and Contemporary Humanity Development* (pp. 239-243). Atlantis Press.

Luo, D., Xia, L., Zhang, C. and Wang, L. (2018) 'Automatic scoring of L2 English speech spoken by Chinese middle school students based on deep learning', *Journal of Shenzhen Institute of Information Technology*, 16(2), pp. 100–104.

Luo, F., Tian, X. T., Tu, Z. R., & Jiang, L. M. (2021) 'New Trend of Educational Assessment: A Research Overview of Intelligent Assessment', *Modern Distance Education Research*, 33 (05), pp. 42 - 52.

Luo, H. T. (2022) 'Be a Good Gardener in the "Jointing and Booting Period" of Students', *Leading Journal of Ideological & Theoretical Education*, 8, pp. 118 - 124.

Lytje, M. and Dyregrov, A. (2021) 'When young children grieve: supporting daycare children following bereavement—a parent's perspective', *OMEGA - Journal of Death and Dying*, 86(3), 980-1001. <https://doi.org/10.1177/0030222821997702>

Ma, N., Zhang, X. & Zhang, Y. (2015) 'The development of mobile learning in China's universities', in: Y. Zhang, ed. *Handbook of Mobile Teaching and Learning*. Berlin: Springer, pp. 47-1. [https://doi.org/10.1007/978-3-642-41981-2\\_47-1](https://doi.org/10.1007/978-3-642-41981-2_47-1)

Machi, L.A. & McEvoy, B.T. (2016) *The literature review: Six steps to success*. 3rd ed. Thousand Oaks, CA: Corwin Press.

Ma, A.L. (2002) 'On Strengthening Quality Education from the Perspective of the Drawbacks of the Ancient Education System', *Songliao Journal (Humanities & Social Sciences Edition)*, 3, pp. 49 - 51.

Ma, L. T. (2022) 'New Engineering, New Medicine, New Agriculture, New Liberal Arts —— From Educational Concepts to Paradigm Changes', *China Higher Education*, 12.

Malina, M.A., Nørreklit, H.S. and Selto, F.H. (2011) 'Lessons learned: advantages and disadvantages of mixed method research', *Qualitative Research in Accounting & Management*, 8(1), pp.59-71.

Martin, F., & Kumar, S. (2018) 'Frameworks for assessing and evaluating e-learning courses and programs', *Leading and managing e-learning: What the e-learning leader needs to know*, pp. 271-280.

Martins, R.M. and Gresse Von Wangenheim, C. (2023) 'Findings on teaching machine learning in high school: A ten-year systematic literature review', *Informatics in Education*, 22(3), pp.421-440.

Masoodi, M. (2018) 'Importance of promoting metacognitive awareness at university', *Vocational Training: Research and Realities*, 29(1), pp. 3-18. <https://doi.org/10.2478/vtrr-2018-0002>

Mayer, D.G. and Butler, D.G. (1993) 'Statistical validation', *Ecological modelling*, 68(1-2), pp.21-32.

McCarthy, P. M., Briner, S. W., Rus, V., & McNamara, D. S. (2007) 'Textual Cohesion', in *Handbook of Latent Semantic Analysis*, eds. T. K. Landauer, D. S. McNamara, S. Dennis, and W. Kintsch. Psychology Press.

McDonald, N., Schoenebeck, S. and Forte, A. (2019) 'Reliability and inter-rater reliability in qualitative research: Norms and guidelines for CSCW and HCI practice', *Proceedings of the ACM on human-computer interaction*, 3(CSCW), pp.1-23.

McNamara, D.S., Crossley, S.A. and McCarthy, P.M. (2010) 'Linguistic features of writing quality',

*Written Communication*, Vol. 27, pp.57–86.

McNamara, D. S., & Graesser, A. C. (2012) ‘Coh-Metrix: An Automated Tool for Theoretical and Applied Linguistics’, *Discourse Processes*, 49(4), pp.315-327.

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014) *Automated Evaluation of Text and Discourse with Coh-Metrix*. Cambridge University Press.

Meng, X., & Li, Z. (2008) ‘The Dilemmas and Solutions of Teachers' Participation in University Internal Governance — From the Perspective of Field’, *Modern University Education*, 6, pp. 82 - 88.

Miao, J., & Chang, J. M. (2022) ‘Research on a Meta-analysis of the Effects of Mobile Inquiry-based Learning’, *Journal of Northeastern University (Social Science)*, 24 (04), pp. 130 - 136.

Miller, K. A., Deci, E. L., & Ryan, R. M. (1988) ‘Intrinsic motivation and self-determination in human behavior’, *Contemporary Sociology*, 17(2), p. 253.

Ming, Q. H., & Zhao, P. (2022) ‘Rethinking Teachers as “Human Beings” and “Teachers”’, *Educational Science Research*, 4, pp. 12 - 18.

Ministry of Education of the People’s Republic of China (MOE) (2015) *Notice of the State Council on Printing and Issuing the Overall Plan for Comprehensively Promoting the Construction of World-Class Universities and First-class Disciplines*. Available at: [http://www.moe.gov.cn/jyb\\_xxgk/moe\\_1777/moe\\_1778/201511/t20151105\\_217823.html](http://www.moe.gov.cn/jyb_xxgk/moe_1777/moe_1778/201511/t20151105_217823.html) (Accessed: 19 November, 2022)

Ministry of Education of the People’s Republic of China (2021) *Notice of the General Office of the Ministry of Education on Doing a Good Job in the Enrollment Work of Secondary Vocational Schools in 2021*. Available at: [http://www.moe.gov.cn/srcsite/A07/moe\\_950/202104/t20210406\\_524618.html](http://www.moe.gov.cn/srcsite/A07/moe_950/202104/t20210406_524618.html)

Ministry of Education of the People’s Republic of China (2022a) *The 15th Session of the 1 + 1 Press Conference on Education in the Past Decade: Introducing the Achievements of China's Educational Reform and Development Since the 18th National Congress of the Communist Party of China from Data*. Available at: <http://www.moe.gov.cn/fbh/live/2022/54875/>, September 27, 2022.

Ministry of Education of the People’s Republic of China (2022b) *Vocational Education Law of the People’s Republic of China*. Available at: [http://www.moe.gov.cn/jyb\\_sjzl/sjzl\\_zcfg/zcfg\\_jyfl/202204/t20220421\\_620064.html](http://www.moe.gov.cn/jyb_sjzl/sjzl_zcfg/zcfg_jyfl/202204/t20220421_620064.html)

Ministry of Education of the People's Republic of China (2023) *The internet access rate in primary and secondary schools nationwide has reached 100%, and the digital literacy and skills of teachers and students have been significantly improved*. Available at: [https://hudong.moe.gov.cn/jyb\\_xwfb/xw\\_fbh/moe\\_2606/2023/cfh\\_0209/baodao/202302/t20230209\\_1043211.html](https://hudong.moe.gov.cn/jyb_xwfb/xw_fbh/moe_2606/2023/cfh_0209/baodao/202302/t20230209_1043211.html) (Accessed: 12 December 2024)

Ministry of Education of the People’s Republic of China (2024) *Statistical Bulletin of the National Education Development in 2023*. Available at: [http://www.moe.gov.cn/jyb\\_sjzl/sjzl\\_fztjgb/202410/t20241024\\_1159002.html](http://www.moe.gov.cn/jyb_sjzl/sjzl_fztjgb/202410/t20241024_1159002.html) (Accessed: 19 November, 2024).

Miles, M. B., and A. M. Huberman (1994) *Qualitative Data Analysis: An Expanded Sourcebook*, 2nd edition. Thousand Oaks, CA: Sage.

Mizani, H., Cahyadi, A., Hendryadi, H., Salamah, S., & Retno Sari, S. (2022) ‘Loneliness, student engagement, and academic achievement during emergency remote teaching during COVID-19: The role of the God locus of control’, *Humanities and Social Sciences Communications*, 9(1), pp.1-9.

Muhammad, M., Wallerstein, N., Sussman, A. L., Avila, M., Belone, L., & Duran, B. (2015) ‘Reflections on Researcher Identity and Power: The Impact of Positionality on Community Based Participatory Research (CBPR) Processes and Outcomes’, *Critical Sociology*, 41(7-8), 1045-1063. <https://doi.org/10.1177/0896920513516025>.

Mulder, M. (2017) *Competence-based vocational and professional education: Bridging the worlds*

of work and education. Cham: Springer.

Nachalwe, C. M. and Bwanga, O. (2021) 'Impact and challenges of consultancy role regarding the delivery of breast imaging services in the united kingdom: consultant breast radiographers' perspective', *Medical Journal of Zambia*, 48(1), pp. 46-53. <https://doi.org/10.55320/mjz.48.1.54>

National Institute for Health and Care Excellence (NICE) (2012) *Methods for the development of NICE public health guidance*. National Institute for Health and Care Excellence.

Naug, H., Colson, N. J., & Donner, D. (2011) 'Promoting metacognition in first year anatomy laboratories using plasticine modeling and drawing activities: a pilot study of the "blank page" technique', *Anatomical Sciences Education*, 4(4), pp. 231-234. <https://doi.org/10.1002/ase.228>

Nazari-Shirkouhi, S., Mousakhani, S., Tavakoli, M., Dalvand, M. R., Šaparauskas, J., & Antuchevičienė, J. (2020) 'Importance-performance analysis based balanced scorecard for performance evaluation in higher education institutions: an integrated fuzzy approach', *Journal of Business Economics and Management*, 21(3), pp. 647-678.

Nicol, D. J., & 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), pp. 199-218.

Nilholm, C., Sundberg, D., Forsberg, E., et al. (2021) 'The Aims and Meaning of Teaching as Reflected in High - Impact Reviews of Teaching Research', *Teaching and Teacher Education*, Volume 107. Available at: <https://doi.org/10.1016/j.tate.2021.103488>

Ning, B. T., Fan, X. W. (2022) 'The Generative Logic and Practical Path of Integration Education', *Journal of Hangzhou Normal*, 44 (05), pp. 62 - 69

Niu, R. X. (2018) 'Review, Reflection and Prospect of Teaching Evaluation Research in the Past 40 Years', *Curriculum, Teaching Material and Method*, 38 (11), pp. 60 - 66.

Noels, K. A., Pelletier, L. G., Clément, R., and Vallerand, R. J. (2000) 'Why are you learning a second language? Motivational orientations and self-determination theory', *Language Learning*, 50(1), pp. 57-85.

Nosratinia, M., Saveiy, M., & Zaker, A. (2014) 'Efl learners' self-efficacy, metacognitive awareness, and use of language learning strategies: how are they associated?', *Theory and Practice in Language Studies*, 4(5). <https://doi.org/10.4304/tpls.4.5.1080-1092>

O'Cathain, A., Murphy, E. and Nicholl, J. (2007) 'Integration and publications as indicators of "yield" from mixed methods studies', *Journal of mixed methods research*, 1(2), pp.147-163.

O'Malley, J. M. (1990) *Learning strategies in second language acquisition*. Cambridge UP.

O'Neill, M. A., & Palmer, A. (2004) 'Importance-performance analysis: a useful tool for directing continuous quality improvement in higher education', *Quality assurance in education*, 12(1), 39-52.

Oppermann, M. (2000) 'Triangulation—a methodological discussion', *International Journal of Tourism Research*, 2(2), pp.141-145.

Overton, T., & Potter, N. (2008) 'The role of goal orientation in self-regulated learning', *Chemistry Education Research and Practice*, 9, pp. 65-69.

Owan, V.J., Abang, K.B., Idika, D.O., Etta, E.O. and Bassey, B.A. (2023) 'Exploring the potential of artificial intelligence tools in educational measurement and assessment', *Eurasia Journal of Mathematics, Science and Technology Education*, 19(8), p. em2307.

Ozuru, Y., Dempsey, K., & McNamara, D. S. (2009) 'Prior Knowledge, Reading Skill, and Text Cohesion in the Comprehension of Science Texts', *Learning and Instruction*, 19(3), pp.228-242.

Paris, S. G. and Winograd, P. (1990) 'How Metacognition Can Promote Academic Learning and Instruction', *Dimensions of Thinking and Cognitive Instruction*, pp. 15-51.

Pan, H. J. (2021) '110 Years of Development of Chinese Primary School Science Curriculum (1912 - 2021)', *Education and Teaching Research*, 2021(35)(07), pp. 45 - 61.

Panadero, E. (2017) 'A review of self-regulated learning: Six models and four directions for

research', *Frontiers in Psychology*, 8, p. 422.

Panadero, E. and Järvelä, S. (2019) 'Socially Shared Regulation of Learning: A Review', *European Psychologist*, 24(1), pp. 117–128.

Patel, V. L., Yoskowitz, N. A., & Arocha, J. F. (2008) 'Towards effective evaluation and reform in medical education: a cognitive and learning sciences perspective', *Advances in Health Sciences Education*, 14(5), pp. 791-812.

Pearson, N., Naylor, P.J., Ashe, M.C., Fernandez, M., Yoong, S.L. and Wolfenden, L. (2020) 'Guidance for conducting feasibility and pilot studies for implementation trials', *Pilot and feasibility studies*, 6, pp.1-12.

Peng, B., Wang, W. Q., Zhang, J. L., Yuan, J. L., & Yu, J. B. (2021) 'The Education Evaluation Reform from the Perspective of Artificial Intelligence', *Forum on Contemporary Education*, 2021(06), pp. 1 - 15.

Peng, J. (2022) 'The Enlightenment of Ancient Chinese Private Schools on Modern School Education', *PR World*, 2022(01), pp. 120 - 121.

Peng, N. & Xiao, J. (2023) 'Research on oral learning strategies of English major students based on the FiF oral training system', *Overseas English*, (21), pp.74-76.

People's Daily (2022) *China Has Built the World's Largest Vocational Education System*. Available at: [http://www.gov.cn/xinwen/2022-05/29/content\\_5692904.htm](http://www.gov.cn/xinwen/2022-05/29/content_5692904.htm) (Accessed: 29 May, 2022).

Petruk, V., Rudenko, Y., Yurchenko, A., Kharchenko, I., Kharchenko, S., & Semenikhina, O. (2023) 'Analysis of the Results of the Pedagogical Experiment on the Integrated Analysis of the Average and Dispersions', *International Journal of Modern Education and Computer Science*. 2022. Vol. 14, № 6, pp. 25-34.

Phuong, L. T., Tuấn, T. Đ., & Phuc, N. T. N. (2019) 'Transformative social learning for agricultural sustainability and climate change adaptation in the vietnam mekong delta', *Sustainability*, 11(23), p. 6775.

Piaget, J. (1970). *Piaget's Theory*. In P. H. Mussen (Ed.), *Carmichael's Manual of Child Psychology*, Vol. 1, pp. 703-732. Wiley.

Piaget, J. (1973) *To Understand Is to Invent: The Future of Education*. New York: Grossman Publishers.

Pintrich, P. R. (2000) 'The Role of Goal Orientation in Self-Regulated Learning', *Handbook of Self-Regulation*, pp. 451–502.

Poehner, M. E. (2008) *Dynamic assessment: A Vygotskian approach to understanding and promoting second language development*. Berlin: Springer.

Poehner, M. E., & Lantolf, J. P. (2010) 'Vygotsky's teaching-assessment dialectic and L2 education: The case for dynamic assessment', *Mind, Culture, and Activity*, 17(4), pp.312-330.

Poth, C. and Munce, S.E. (2020) 'Commentary--Preparing today's researcher for a yet unknown tomorrow: Promising practices for a synergistic and sustainable mentoring approach to mixed methods research learning', *International Journal of Multiple Research Approaches*, 12(1).

Qiao, H., Dong, B. and Liu, C. (2012) 'Research on computer-assisted automatic scoring technology for PETS oral tests', *Foreign Language Testing and Teaching*, (03), pp. 47–52.

Qin, F. M., & Mo, K. (2022) 'A research on the CIPP-based quality evaluation of industry-education integration in vocational education', *Journal of Southwest University (Social Sciences Edition)*, 48 (03), pp. 194 - 203.

Qin, X. Q.(2021) 'On Factors Influencing Curriculum Development — Starting from Theoretical Basis of Curriculum', *Teaching of Forestry Region*, 12, pp. 48 - 52.

Rad, M. R. (2018) 'The effect of moodle-mediated instruction on the reading comprehension of iranian efl learners', *International Journal of Advanced Multidisciplinary Scientific Research*, 1(3), pp. 70-78. <https://doi.org/10.31426/ijamsr.2018.1.3.136>

- Råheim, M., Magnussen, L. H., Sekse, R. J. T., Lunde, Å., Jacobsen, T., & Blystad, A. (2016) 'Researcher-researched relationship in qualitative research: Shifts in positions and researcher vulnerability', *International Journal of Qualitative Studies on Health and Well-being*, 11, Article 30996.
- Rahman, M. S. (2016) 'The advantages and disadvantages of using qualitative and quantitative approaches and methods in language "testing and assessment" research: A literature review', *Journal of education and learning*, 6(1).
- Ramaswami, G., Susnjak, T., & Mathrani, A. (2023) 'Effectiveness of a learning analytics dashboard for increasing student engagement levels', *Journal of Learning Analytics*, 10(3), pp. 115-134.
- Ran, F. (2022) 'Research on the Concurrent Development of Information Technology and Vocational Education — Review of "Introduction to Research on Vocational Education Informatisation"', *China Sciencepaper*, 17 (10), pp. 1182.
- Raudenbush, S. W. & Bryk, A. S. (2002) *Hierarchical linear models: Applications and data analysis methods*. 2nd edn. Thousand Oaks, CA: Sage Publications.
- Reidenberg, J. R., & Schaub, F. (2018) 'Achieving big data privacy in education', *Theory and Research in Education*, 16(3), pp. 263-279.
- Renz, A., & Hilbig, R. (2020) 'Prerequisites for artificial intelligence in further education: Identification of drivers, barriers, and business models of educational technology companies', *International Journal of Educational Technology in Higher Education*, 17(1), p.14.
- Ridley, D. (2012) *The literature review: A step-by-step guide for students*. 2nd ed. London: Sage.
- Robertson, S.N., Humphrey, S.M. and Steele, J.P. (2019) 'Using Technology Tools for Formative Assessments', *Journal of Educators Online*, 16(2).
- Rogti, M. (2020) 'Contribution of meta-cognitive competence in preserving student self-regulation in algerian higher education: a case from ens of laghouat', *Arab World English Journal*, 11(4), pp. 445-460. <https://doi.org/10.24093/awej/vol11no4.28>
- Rutter, M. (1988) 'Longitudinal data in the study of causal processes: Some uses and some pitfalls', *Child Development*, 59(1), pp. 1-23.
- Ryan, R.M. & Deci, E.L. (2000) 'Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being', *American Psychologist*, 55(1), pp. 68-78.
- Ryan, R.M. and Deci, E.L. (2020) 'Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions', *Contemporary educational psychology*, 61, p.101860.
- Sadler, D.R. (1989) *Formative assessment and the design of instructional systems*. Instructional Science, 18(2), pp.119-144.
- Salaberry, M.R. (2001) 'The use of technology for second language learning and teaching: A retrospective', *The Modern Language Journal*, 85(1), pp. 39-56.
- Salsbury, T., Crossley, S. A., & McNamara, D. S. (2011) 'Psycholinguistic Word Information in Second Language Oral Discourse', *International Journal of Corpus Linguistics*, 16(1), pp.37-62.
- Samuel, R. S. (2022) *A Comprehensive Study of Education* (2nd edition), pp. 41 - 42.
- Samuel, R. S. (2022) *A Comprehensive Study of Education* (2nd edition), pp. 51 - 72.
- Samuel, R. S. (2022) *A Comprehensive Study of Education* (2nd edition). PHI Learning Private Limited, pp. 114.
- Sánchez-Verdejo, J., Anguita, J., and Fernandes, F. (2019) 'The teaching of foreign languages: proposals for improvement', *Chakíñan, Revista De Ciencias Sociales Y Humanidades*, 8, pp. 133-147. doi: 10.37135/chk.002.08.09
- Satyawan, I.M., Wahjoedi, W. and Swadesi, I.K.I. (2021) 'The effectiveness of online learning through Undiksha e-learning during the covid-19 pandemic', *Journal of Education Technology*, 5(2), pp.191-199.

- Savin-Baden, M. & Major, C. H. (2007) 'Using interpretative meta-ethnography to explore the relationship between innovative approaches to learning and their influence on faculty understanding of teaching', *Higher Education*, 54(6), pp. 833–852.
- Scalise, K. and Gifford, B. (2006) 'Computer-based assessment in e-learning: A framework for constructing "intermediate constraint" questions and tasks for technology platforms', *The Journal of Technology, Learning and Assessment*, 4(6).
- Schlosser, L., Hood, C. E., Hogan, E., Baca, B., & Gentile-Mathew, A. (2022) 'Choosing the right educational technology tool for your teaching: A data-privacy review and pedagogical perspective into teaching with technology', *Journal of Educational Technology Systems*, 51(2), pp. 236-251.
- Schmidt, T., & Strasser, T. (2022) 'Artificial intelligence in foreign language learning and teaching: a CALL for intelligent practice', *Anglistik: International Journal of English Studies*, 33(1), pp. 165-184.
- Schoonenboom, J., & Johnson, R. B. (2017) 'How to construct a mixed methods research design', *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 69(2), pp. 107-131.  
<https://doi.org/10.1007/s11577-017-0454-1>
- Schraw, G. and Moshman, D. (1995) 'Metacognitive Theories', *Educational Psychology Review*, 7(4), pp. 351–371.
- Scriven, M. (1967) 'The methodology of evaluation', in: R.L. Tyler, R.M. Gagne & M. Scriven, eds. *Perspectives of curriculum evaluation*. Chicago: Rand McNally, pp.39-83.
- Selwyn, N. (2016) *Education and technology: Key issues and debates*. Bloomsbury Publishing.
- Shadiev, R. and Yang, M. (2020) 'Review of studies on technology-enhanced language learning and teaching', *Sustainability*, 12(2), p.524.
- Shadiev, R., Wang, X., & Huang, Y. M. (2021) 'Cross-cultural learning in virtual reality environment: facilitating cross-cultural understanding, trait emotional intelligence, and sense of presence', *Educational Technology Research and Development*, 69(5), pp. 2917-2936.
- Shaheen, N., Shaheen, A., Ramadan, A., Hefnawy, M.T., Ramadan, A., Ibrahim, I.A., Hassanein, M.E., Ashour, M.E. & Flouty, O. (2023) 'Appraising systematic reviews: a comprehensive guide to ensuring validity and reliability', *Frontiers in Research Metrics and Analytics*, 8, p.1268045.  
<https://doi.org/10.3389/frma.2023.1268045>
- Shan, B. L. (2021) 'The Enlightenment of Ancient Chinese Academy Education on Middle School History Teaching', *Foundamental Education Forum*, 2021(25), pp. 37 - 38.
- Shao, X. F. (2016) 'A Historical Study of the Changes of the Chinese Teacher-student Relationships', *Journal of Hebei Normal University (Educational Science)*, 2016(18)(03), pp. 10 - 16.
- Sharma, S., Tyagi, V. and Vaidya, A. (2023) 'Technology Enabled Self-directed Learning: A Review and Framework', in *International Conference on Advances in Computing and Data Sciences*, Cham: Springer Nature Switzerland, pp. 228–240.
- Shen, L.T. (1996) 'A Historical Reflection on the Education System in Ancient China', *Journal of Nanjing University (Philosophy, Humanities and Social Sciences)*, 1996, (01), pp. 173 - 179.
- Shi, K. C. (2021) 'The Evolution and Characteristics of School Textbooks in Ancient China', *People's Tribune*, 2021(23), pp. 110 - 112.
- Shi, L. (2022) 'Exploring Multidimensional Teaching Evaluation and Promoting the Development of Students' Core Competencies — An Exploration of Multidimensional Teaching Evaluation in High School History Based on the Development of Subject Core Competencies', *History Research and Teaching*, 4, pp. 149 - 155.
- Shi, O., & Zhang, W. (2021) 'Basic Experience and Contemporary Challenges of Textbook Construction by the Communist Party of China in the Past Century', *Educational Science*, 4, pp. 1 - 9.
- Shi, Z.Y. (2018) 'A Philosophic Discussion on Developing Key Competencies of Chinese Students', *Curriculum, Teaching Material and Method*, 38 (09), pp. 36 - 41.

- Shi, Z.Y. (2022) 'The Policy Connotation and Practical Path of High-quality Educational Development', *People's Education*, 7, pp. 24 - 28.
- Shonfeld, M., Cotnam-Kappel, M., Judge, M., Ng, C. Y., Ntebutse, J. G., Williamson-Leadley, S., & Yildiz, M. N. (2021) 'Learning in digital environments: a model for cross-cultural alignment', *Educational Technology Research and Development*, pp. 1-20.
- Shorten, A. and Smith, J. (2017) 'Mixed methods research: expanding the evidence base', *Evidence-based nursing*, 20(3), pp.74-75.
- Shrestha, P., & Coffin, C. (2012) 'Dynamic assessment, tutor mediation and academic writing development', *Assessing Writing*, 17(1), pp.55-70.
- Shute, V.J. (2008) 'Focus on formative feedback', *Review of Educational Research*, 78(1), pp. 153–189.
- Siegel, H. (2022) 'Philosophy of Education', *Encyclopedia Britannica*, 27 Sep. 2022. Available at: <https://www.britannica.com/topic/philosophy-of-education> (Accessed: 19 October 2022).
- Silverman, D. (2006) *Interpreting Qualitative Data: Methods for Analyzing Talk, Text and Interaction*, 3rd edition, London: Sage.
- Şirin, Y.E., Aydın, Ö. & Bilir, F.P. (2018) 'Transformational-transactional leadership and organisational cynicism perception: physical education and sport teachers sample', *Universal Journal of Educational Research*, 6(9), 2008-2018.
- Snow, E. L., Allen, L. K., Jackson, G. T., & McNamara, D. S. (2015) 'Text Readability and Intuitive Simplifications: A Coh-Metrix Analysis of Texts for English Learners', *System*, 50, pp.1-12.
- Soares, L. T. (2020) 'A microanálise da aprendizagem musical autorregulada como ferramenta para o desenvolvimento da pedagogia da performance', *Percepta-Revista de Cognição Musical*, 7(2), 29-50.
- Song, C. X., Zhang, P., Zi, C. Y., Li, C. X. (2021) 'Application of Virtual Simulation Technology Combined with Intelligent Vocational Education Classroom in the Teaching of Childbirth Nursing Care', *Journal of Nursing Science*, 2021(36)(23), pp. 65 - 68.
- Song, L. L. (2022) 'The Rise of the Field of Vocational Education in Modern China — Taking Huang Yanpei and Tan Kah Kee's Vocational Education Ideas as Examples', *Vocational and Technical Education*, 43 (15), pp. 75 - 80.
- Spector, J. M. (2014) 'Conceptualizing the emerging field of smart learning environments', *Smart Learning Environments*, 1(2). <https://doi.org/10.1186/s40561-014-0002-7>.
- Stahl, G., Koschmann, T. & Suthers, D. (2006) 'Computer-supported collaborative learning: An historical perspective', in R. K. Sawyer (ed.) *Cambridge Handbook of the Learning Sciences*. Cambridge University Press, pp. 409-426.
- Stockwell, G. (2022) *Mobile-assisted language learning: Concepts, contexts and challenges*. Cambridge: Cambridge University Press.
- Su, F. and Zou, D. (2022) 'Technology-enhanced collaborative language learning: theoretical foundations, technologies, and implications', *Computer Assisted Language Learning*, 35(8), pp.1754-1788.
- Sun, H. (2011) 'Automated scoring of speech: An overview of research at home and abroad', *Foreign Language Education in China*, 4(2), pp. 28–36, 89–90.
- Sun, J. (2021) 'On the Dissolution and Reconstruction of the World of Educational Significance in Ancient China—An Investigation Centered on Zhu Xi's Criticism of Imperial Examinations in the Song Dynasty', *Academic Exploration*, 2021, (11), pp. 127 - 134.
- Sun, J. M. (2022) *From Confucius to Tao Xingzhi: The Essence of Chinese Normal Education Thought*. Beijing Institute of Technology Press.
- 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), pp. 1183-1202.

- Sun, P. Q. (2019) *History of China's Education* (4th edition). East China Normal University Press, pp: 11 - 13.
- Suraworachet, W., Zhou, Q. & Cukurova, M. (2023) 'Impact of combining human and analytics feedback on students' engagement with, and performance in, reflective writing tasks', *International Journal of Educational Technology in Higher Education*, 20(1), pp. 1-24.
- Taherdoost, H., (2022) 'What are different research approaches? Comprehensive Review of Qualitative, quantitative, and mixed method research, their applications, types, and limitations', *Journal of Management Science & Engineering Research*, 5(1), pp.53-63.
- Tan, W. F. (2022) 'A Brief Discussion on the Construction of New Forms of Textbooks in Colleges and Universities', *Science-Technology & Publication*, 9, pp. 86 - 92.
- Tan, W. J., Tan, X. L. (2022) 'The Historical Evolution and Characteristics of the Purpose of School Education in China'. *Teaching and Management*, 2022(30), pp. 18 - 13.
- Tan, Y. L., & Chai, Y. B. (2022) 'Research on the Training Mode of Outstanding Craftsmen Based on Huang Yanpei's Vocational Education Thoughts', *Education and Vocation*, 7, pp. 103 - 106.
- Tang, L. W. (2021) 'In-depth Inquiry and Preliminary Suggestions on Parent Participation in School Governance', *Primary and Secondary School Management*, 4, pp. 37 - 38.
- Tashakkori, A., Johnson, R. B., & Teddlie, C. (2020) *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Sage publications.
- Teng, F. (2020) 'The role of metacognitive knowledge and regulation in developing university EFL learners' writing performance', *Innovation in Language Learning and Teaching*, 14(5), pp. 436–450.
- Terry, G., & Braun, V. (2017) 'Short but often sweet: The surprising potential of qualitative survey methods', in V. Braun, V. Clarke, & D. Gray (Eds.), *Collecting qualitative data: A practical guide to textual, media and virtual techniques* (pp. 15–44). Cambridge University Press.
- Thabane, L., Ma, J., Chu, R., Cheng, J., Ismaila, A., Rios, L.P., Robson, R., Thabane, M., Giangregorio, L. and Goldsmith, C.H. (2010) 'A tutorial on pilot studies: the what, why and how', *BMC medical research methodology*, 10, pp.1-10.
- The Central Committee of the Communist Party of China, The State Council (1999) *Decision on Deepening Educational Reform and Comprehensively Promoting Quality-oriented Education*. Available at: [http://www.moe.gov.cn/srcsite/A05/s7052/199907/t19990719\\_162906.html](http://www.moe.gov.cn/srcsite/A05/s7052/199907/t19990719_162906.html)
- Theelen, H., Van den Beemt, A. & den Brok, P. (2019) 'Classroom simulations in teacher education to support preservice teachers' interpersonal competence: A systematic literature review', *Computers & Education*, 129, pp. 14–26.
- Thorndike, E. L. (1922) *The psychology of arithmetic*. New York: Macmillan.
- Thurmond, V.A. (2001) 'The point of triangulation', *Journal of nursing scholarship*, 33(3), pp.253-258.
- Tian, Z.P. (2016) 'Some Thoughts on the Education in the Republic of China', *Journal of Educational Studies*, 12 (04), pp. 102 - 111.
- Tondeur, J., Van Braak, J., Ertmer, P. A., & Ottenbreit-Leftwich, A. (2017) 'Understanding the relationship between teachers' pedagogical beliefs and technology use in education: A systematic review of qualitative evidence', *Educational technology research and development*, 65, pp. 555-575.
- Tong, Y. (2012) 'An empirical study of the cultivation of autonomy through metacognitive strategy training', *International Journal of Education and Management Engineering*, 2(9), pp. 70-75.  
<https://doi.org/10.5815/ijeme.2012.09.12>
- Treharne, G.J. and Riggs, D.W. (2015) 'Ensuring quality in qualitative research', *Qualitative research in clinical and health psychology*, 2014, pp.57-73.
- Tynjälä, P. (2013) 'Toward a 3-P model of workplace learning: A literature review', *Vocations and Learning*, 6(1), pp. 11-36.

Ushioda, E. (2003) 'Motivation as a socially mediated process', in D. Little, J. Ridley & E. Ushioda (eds.) *Learner autonomy in the foreign language classroom: Learner, teacher, curriculum and assessment*. Dublin: Authentik.

Van der Kleij, F. M., Adie, L. E. & Cumming, J. J. (2019) 'A meta-review of the student role in feedback', *International Journal of Educational Research*, 98, pp. 303-323.

Van Teijlingen, E. and Hundley, V. (2002) 'The importance of pilot studies', *Nursing Standard (through 2013)*, 16(40), p.33.

Vaughan, N. D., Cleveland-Innes, M., & Garrison, D. R. (2013) *Teaching in blended learning environments: Creating and sustaining communities of inquiry*. Athabasca University Press.

Veenman, M. V., Van Hout-Wolters, B., and Afflerbach, P. (2006) 'Metacognition and learning: conceptual and methodological considerations', *Metacognition and Learning*, 1(1), pp. 3–14.

Venkatesh, V., & Davis, F. D. (2000) 'A theoretical extension of the technology acceptance model: Four longitudinal field studies', *Management Science*, 46(2), pp.186-204.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003) 'User acceptance of information technology: Toward a unified view', *MIS Quarterly*, 27(3), pp. 425-478.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wisdom, J., Creswell, J. W. (2013) 'Mixed Methods: Integrating Quantitative and Qualitative Data Collection and Analysis While Studying Patient-Centered Medical Home Models', *AHRQ and PCHM*. JPSE 2 (2): 25-36.

Vygotsky, L.S. (1978) *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.

Wahdain, E. A., & Ahmad, M. N. (2014) 'User acceptance of Information Technology: Factors, theories and applications', *Journal of Information Systems Research and Innovation*, 6 (1), pp. 17-25.

Wan, R. (2020) 'Constructing a Home-school Collaborative Educational Community', *Educational Science Research*, 2020(10), pp: 1.

Wang, C., Wang, H. X. (2022) 'School Curriculum Integration from the Perspective of the Integration of Five Domains of Education: Realistic Investigation and Rational Reflection', *Curriculum, Teaching Material and Method*, 42(09), pp. 79 - 86.

Wang, J. (2021) 'Research on college oral English teaching based on the FiF oral training system', *English Square*, (21), pp.98-100. doi:10.16723/j.cnki.yygc.2021.21.032.

Wang, H. (2022) 'Analysis of the Policy Changes of Industry-education Integration in Chinese Vocational Education', *Chinese Vocational and Technical Education*, 27, pp. 5 - 12.

Wang, J., Zou, J. & Lin, J. (2018) 'Effect research of smartphone-based mobile learning on autonomous learning ability of higher vocational students', *Nursing Research*, 32(9), pp. 1451–1453.

Wang, J., Katz, I., & Yue, J. L. (2021) 'Quality Education Policy, Neoliberalism and the Rise of Shadow Education in China'. *China Youth Study*, 7, pp. 110 - 119.

Wang, J. Y. (2019) 'Educational Thoughts in The Analects of Confucius', *Teaching Reference of Middle School Politics*, 2019(12), pp. 75 - 77.

Wang, L. (2024) 'Examining the effectiveness of automated writing evaluation feedback in EFL/ESL writing: A meta-analysis', *Foreign Languages in China*, 21(4), pp. 69–77.

Wang, M., & Zhu, Z. H. (2021) 'Parent Participation from the Perspective of School Change: Connotation, Dilemmas and Breakthroughs', *Contemporary Education Sciences*, 12, pp. 60 - 65.

Wang, W. (2018) 'An empirical study on cultivating autonomous learning ability of vocational college students under the blended teaching model', *China Vocational and Technical Education*, (35), pp. 10–15.

- Wang, W. L. (2022) 'A Casual Talk on Educational Evaluation', *Modern University Education*, No. 38(06), pp. 60 - 64.
- Wang, X. H. (2022) 'Loose-leaf Type Textbooks Development in Vocational Colleges Based on the Philosophy of Competence-based Education', *Journal of Vocational Education*, 38 (09), pp. 55 - 61.
- Wang X. L. (2013) 'Research on the Purpose of Education in Ancient China', *Theory Research*, 2013(36), pp. 185 - 186.
- Wang, Y. N. (2022) 'Mode Construction and Realisation Path of Specialty Group Integration in Higher Vocational Colleges under the Background of the "Double High Plan"', *Chinese Vocational and Technical Education*, 2022(07), pp. 53 - 60.
- Warren, C. A. (2002) 'Qualitative interviewing', *Handbook of interview research: Context and method*, 839101, 103-116.
- Washington, S., Karlaftis, M.G., Mannering, F. and Anastasopoulos, P. (2020), *Statistical and econometric methods for transportation data analysis*. Chapman and Hall/CRC.
- Wei, J. (2015) 'Investigation of the autonomous learning ability of vocational college students', *Education and Vocation*, (04), pp. 78–79.
- White, B. Y. and Frederiksen, J. R. (1998) 'Inquiry, Modeling, and Metacognition: Making Science Accessible to All Students', *Cognition and Instruction*, 16(1), pp. 3–118.
- Wijayanti, E., Nugroho, A. P., Mursyid, A. M. M., & Rahmawansyah, R. (2023) 'Teaching efl students in post-pandemic era: ict tools and obstacles', *AL-ISHLAH: Jurnal Pendidikan*, 15(2), pp. 2591-2604. <https://doi.org/10.35445/alishlah.v15i2.2943>
- Wijaya, K. (2021) 'Graduate students' perceptions on their metacognitive awareness in academic writing', *Elt Echo the Journal of English Language Teaching in Foreign Language Context*, 6(1), p. 50. <https://doi.org/10.24235/eltecho.v6i1.7957>
- Wiliam, D. (2011) *Embedded formative assessment*. Bloomington, IN: Solution Tree Press.
- Winne, P. H. and Hadwin, A. F. (1998) 'Studying as Self-Regulated Learning', *Metacognition in Educational Theory and Practice*, pp. 277–304.
- Wu, H.C., Fan, K. (2011) 'A Brief Discussion on the Change of Educational Guidelines in the Early Years of the Republic of China—From Five Educational Proposals to Four Educational Types', *Journal of Hebei Normal University (Educational Science)*, 2011, 13(09), pp. 23 - 29.
- Wu, T. H. (2022) 'Research on the Ecological Classroom of Higher Vocational English Based on Factor Analysis', *Journal of Wuhan Institute of Shipbuilding Technology*, 21 (02), pp. 68 - 73.
- Wu, W. B., Weng, W. B., Zhang, L. (2022) 'On the Construction of Learning Field Based on Vocational Education informatisation "Five in One, Integration of Virtual and Real"', *Vocational and Technical Education*, 43 (05), pp. 66 - 70.
- Wu, Y. (2022) 'Chinese Modernisation and the Reform, Innovation and Development of Higher Education', *China Higher Education Research*, No. 11, pp. 21 - 29.
- Wu, Z. J. (2022) 'The Application of Inquiry Teaching in High School Mathematics Teaching', *People's Education*, 6, p. 79.
- Xia, X. M. (2022) 'Project-based Learning Evaluation Pointing to Core Competencies', *Journal of the Chinese Society of Education*, 9, pp. 50 - 57.
- Xiang, P. (2020) 'Research on the impact of online learning in MOOC environments on the autonomous learning of vocational college students', *Modern Vocational Education*, 4, pp. 43–45.
- Xiao, H.Y. & Hu, X. (2022) 'Vocational Education Industry-Education Integration Policy: Characteristics, Deficiencies and Optimisation Suggestions', *Chinese Vocational and Technical Education*, 10, pp. 61 - 66.
- Xiao, L. L., Zhang, Y. Z., & Sun, J. F. (2020) 'Wisdom of Education in Ancient China'. Southeast University.

- Xu, S. F., Li, Y. W., & Ma, D. X. (2022) 'The Reform of Polymer Chemistry Teaching Method Aiming at the Cultivation of Applied Talents', *Polymer Bulletin*, 9, pp. 119 - 125.
- Xu, J., He, W. & Xu, P. (2022) 'The theoretical thinking of the high-quality development of China's higher vocational education in the new era and its practical approach', *Journal of Higher Education*, 43(4), pp.78-86.
- Xue, X. and Su, X. (2024) 'Exploration of the learning evaluation system for higher vocational college students based on the cultivation of independent learning ability', *Journal of Educational Studies of Renmin University of China*, (02), pp. 100–112.
- Xue, Y.M., Gao, X.L. (2023) 'Thinking of the Integrity of the Five Educations of Morality, Intelligence, Physical Education, Art and Labor from the Perspective of System Theory', *Chinese Journal of Systems Science*, pp. 99 - 103.
- Yan, S. J. (2021) 'Textbook Construction in the New Era: Cultivating the Soul, Inspiring Intelligence and Increasing Wisdom', *The Teaching of Thought and Political Study*, 2021(04), pp. 4 - 6.
- Yan, W., Lowell, V. L. & Yang, L. (2024) 'Developing English language learners' speaking skills through applying a situated learning approach in VR-enhanced learning experiences', *Virtual Reality*, 28, p. 167. <https://doi.org/10.1007/s10055-024-01061-5>
- Yang, P. & Liu, Y. (2016) 'Provincial Coordination and Balanced Development for Library Vocational Education in China', *Peking University Education Review*, 14(3), pp.59-83.
- Yang, S.S., Lou, S.Z. (2020) 'The Realistic Contradictions, Attributions and Solutions in the Coordinated Development of Urban and Rural Education', *Journal of the Chinese Society of Education*, No. 5, pp. 44 - 48.
- Yang, Z.F. (2024) *Over 97,000 MOOCs Have Been Launched in China*. Available at: <https://www.news.cn/politics/20241212/5c203a2099994920a9ce25c634954779/c.html> (Accessed: 12 December 2024).
- Yang, Z.S., Li, S.N. (2023) 'How Is Chinese-style Educational Modernisation Possible?', *Social Science Front*, No. 7, pp. 234 - 243.
- Yao, Y. J. (2021) 'Spring Silkworm, Gardener, Candle — My Beloved Teacher in My Heart', *Huangzhong (Journal of Wuhan Conservatory of Music)*, 2, pp. 13 - 15 + 164.
- Yesilyurt, Y.E. (2023) 'AI-enabled assessment and feedback mechanisms for language learning: transforming pedagogy and learner experience', in *Transforming the Language Teaching Experience in the Age of AI*. Hershey, PA: IGI Global, pp. 19-38.
- Ye, F. G. (2019) 'The Vanity and Poverty of Educational Evaluation', *Tsinghua Journal of Education*, 40 (01), pp. 18 - 21.
- You, Y. (2022) 'Huang Yanpei's Vocational Education School-running Thoughts and Their Enlightenment', *Education and Vocation*, 18, pp. 99 - 103.
- Yu, J. J. (2019) 'Discussion on the Difficulties and Pain Points Encountered by Teachers in Higher Vocational Colleges in Class', *Think Tank Era*, 21, pp. 87 + 100.
- Yu, J., Liao, Y.Y. & Lin, Y.X. (2022) 'A Study on the Discourse Rhythm Characteristics of Chinese English Learners under Different Oral Task Types', *Foreign Languages and Their Teaching*, 2, pp. 79 - 90 + 147 - 148.
- Yu, T. Y. (2020) 'Teachers: Imparting Knowledge, Answering Questions', *Foreign Economics & Management*, 41 (01), pp. 149 - 150.
- Zeng, R. (2003) 'A Survey and Analysis of College Students' Interest and Motivation in English Learning', *Journal of Xinjiang University (Philosophy and Social Science Edition)*, S1, pp. 123-125.
- Zeng, Y. (2019) 'Fostering integrity through education: An interpretation from the perspective of the evolution of education thoughts in ancient China', *Contemporary Education and Culture*, 11 (01), pp. 7 - 11.
- Zhan, B., & Zhang, J. (2021) 'Survey on Learning Motivation of English Majors in Higher Vocational Colleges', *Journal of Tonghua Normal University*, 42 (01), pp. 124 - 129.

- Zhang, C. & Ya, G. (2024) 'Mechanisms and strategies for enhancing autonomous learning behavior among higher vocational college students', *Vocational and Technical Education*, 45(14), pp. 60–66.
- Zhang, H. (2019) 'On John Dewey and China's Educational Reform', *Journal of East China Normal University (Educational Sciences)*, Vol. 37(02), pp. 18 - 28.
- Zhang, H. W., Wang, H. X., Guo, J. F., Wang, X. H. (2020) 'Online Teaching of Embedded System Based on Tencent Classroom and Virtual Simulation Technology', *Experimental Technology and Management*, 37 (12), pp. 170 - 174.
- Zhang, J. J., & Cui, F. Z. (2022) 'The Essential Characteristics of Huang Yanpei's Vocational Education Thoughts and the Construction of the Modern Vocational Education System', *Education and Vocation*, 13, pp. 5 - 11.
- Zhang, N., Franklin, M., & Hodges, J. (2023) 'Use of Online Formative Quizzes to Enhance Long-Term Retention of Knowledge in Chiropractic Students', *World*, 10(1).
- Zhang, P. F., Xu, J. C., Gao, P. W., & Wang, T. (2021) 'How Teachers' Exam-oriented Education Concepts are Formed: A Case Study from the Perspective of Life History'. *Education Research Monthly*, 9, pp. 67 - 75 + 84.
- Zhang, W., & Liu, Y. L. (2022) 'On the High-quality Development of Textbooks in the New Era', *Journal of Educational Science of Hunan Normal University*, Vol. 21(05), pp. 32 - 39.
- Zhang, Y. (2022) 'A Preliminary Discussion on the Whole-process Democracy in School Management', *Primary and Secondary School Management*, 3, pp. 30 - 33.
- Zhang, Z.W., Wang, W.G. (2022) 'The Philosophical Wisdom of the Integration of Moral, Intellectual, Physical, Aesthetic and Labor Education in Universities in the New Era', *Journal of Graduate School of Chinese Academy of Social Sciences*, 42 (02), pp. 20 - 40 + 144 - 145.
- Zhao, X. L., & Cai, G. C. (2020) 'The Reasons for the Decline of "Teacher's Dignity" and the Creative Transformation', *Jiangsu Higher Education*, 10, pp. 69 - 76.
- Zhao, Y. T., Zhong, S. C., Tang, Yewei (2022) 'Research on the Ecology of e-Learning Space from the Perspective of Technology Empowerment', *China Educational Technology*, 10.
- Zhao, Z. Q., & Gao, F. (2022) 'Professional Competence of Students in Chinese Vocational Colleges and Its Influencing Factors—An Analysis on COMET Data of Five Major Areas of Manufacturing', *China Educational Technology*, 6, pp. 47 - 55.
- Zheng, Q. Q., & Yang, K. K. (2021) 'The Centennial Course, Achievements and Prospects of the Development of Higher Education by the Communist Party of China', *University Education Science*, 2,
- Zheng, R. L., & Xu, D. B. (2021) 'The New Form of Examination-oriented Education in Senior High School under the Background of New College Entrance Examination', *Journal of Educational Science of Hunan Normal University*, 20 (04), pp. 31 - 38.
- Zhi Yan Consulting (2022) *Analysis of the Development Status of China's Educational informatisation Industry in 2021: The Industry Continues to Develop in Depth*. Available at: [http://t.10jqka.com.cn/pid\\_240800381.shtml](http://t.10jqka.com.cn/pid_240800381.shtml)
- Zhong, D. H. (2022) 'Strengthening the Confidence of Higher Education with Chinese Characteristics and Accelerating the High-Quality Development of Higher Education', *China Higher Education Research*, 7, pp. 1 - 2 + 14.
- Zhou, H. X. (2012) 'The Development of Chinese Curriculum Thought and the Current Situation of Curriculum Reform', *Journal of Teaching and Management*, 6, pp. 101 - 103.
- Zhou, Y., & Xu, H. H. (2022) 'The Responsibilities and Actions of Rural Teachers towards Rural Revitalisation', *Journal of Soochow University (Educational Science Edition)*, 10 (01), pp. 10 - 19.
- Zhu, D. Q., & Shi, X. J. (2021) 'From Hierarchy to Type: The Centennial Development of Vocational Education in China', *Journal of Southwest University (Social Sciences Edition)*, 47 (02), pp. 103 - 117 + 228.

- Zhu, Z.W. (2020) 'Deeply Studying General Secretary Xi Jinping's Important Expositions on Education to Facilitate the High-Quality Development of Basic Education in the New Era', *Journal of the Chinese Society of Education*, 9, pp. 1 - 3.
- Zhu, Y., Zuo, H. and Chen, Y. (2023) 'Digital transformation in vocational education: Challenges, strategies, and an experimental proposal', In *2023 4th International Conference on Artificial Intelligence and Education (ICAIE 2023)*, pp. 643-650. Atlantis Press.
- Zhou, T. (2014) 'An empirical examination of initial trust in mobile payment', *Wireless Personal Communications*, 77, pp. 1519-1531.
- Zhou, Y. (2018) 'Teaching strategies for cultivating oral communication skills of college students based on the FiF oral training ecosystem', *Journal of Leshan Teachers College*, (6), pp.111-116.
- Zimmerman, B. J. (2002) 'Becoming a Self-Regulated Learner: An Overview', *Theory Into Practice*, 41(2), pp. 64-70.
- Zimmerman, B. J., & Schunk, D. H. (Eds.) (2013) *Self-regulated learning and academic achievement: Theoretical perspectives*. Routledge.
- Zuo, L.L. (2020) 'The Enlightenment of the Ecological Ethics Concepts Contained in Confucius' Confucianism Thought on Contemporary College Students', *Sinogram Culture*, 6, pp. 168 - 169.

## Appendixes

### Appendix 1 Scales of Questionnaire and Data Analysis in the Pilot Study

A preliminary survey on students' using WBLT/WBFAT at JTC

| Questions                                                | Answers                           | The number of participants<br>(N=30) |
|----------------------------------------------------------|-----------------------------------|--------------------------------------|
| 1. Have you used a WLT ?                                 | Y                                 | 29                                   |
|                                                          | N                                 | 1                                    |
| 2. Have you used a WFAT ?                                | Y                                 | 16                                   |
|                                                          | N                                 | 14                                   |
| 3. Which kind of WLT/WFAT have you used ?                | FIF OELAS                         | 2                                    |
|                                                          | Liulishuo English                 | 15                                   |
|                                                          | Italki                            | 5                                    |
|                                                          | Kouyuyi                           | 4                                    |
|                                                          | Others                            | 3                                    |
| 4. How often do you use WLT/WFAT ?                       | 1-30 minutes a week               | 18                                   |
|                                                          | 3-60 minutes a week               | 7                                    |
|                                                          | 61-90 minutes a week              | 4                                    |
|                                                          | More than 90 minutes              | 0                                    |
| 5. What do you think about the effectiveness of WLT/WFAT | Very good                         | 2                                    |
|                                                          | Good                              | 3                                    |
|                                                          | Neutral                           | 6                                    |
|                                                          | Bad                               | 18                                   |
| 6. What is the most significant limitation of WLT/WFAT   | Limitation of learning resources  | 4                                    |
|                                                          | Limitation of evaluation function | 10                                   |
|                                                          | Limitation of feedback            | 6                                    |
|                                                          | Limitation of interaction         | 8                                    |

Data analysis of the scales of questionnaire in the pilot study

| Questions                                                     | Mode | Median | Variance | SD    |
|---------------------------------------------------------------|------|--------|----------|-------|
| 1. Have you used a web-based learning tool (WLT)?             | 1    | 15.0   | 392.00   | 19.80 |
| 2. Have you used a web-based FA tool(WFAT)?                   | 14   | 15.0   | 2.00     | 1.41  |
| 3. Which kind of WLT/WFAT have you used?                      | 2    | 4.0    | 27.70    | 5.26  |
| 4. How often did you use the WLT/WFAT?                        | 0    | 5.5    | 59.58    | 7.72  |
| 5. What do you think about the effectiveness of the WLT/WFAT? | 2    | 4.5    | 54.25    | 7.37  |
| 6. What was the most significant limitation of the WLT/WFAT?  | 4    | 7.0    | 6.67     | 2.58  |

## Appendix 2 Item-Total Statistics of the Scales of Questionnaire and Data Analysis in the Survey of

### Students' Attitudes Towards Using WFAT

|                                                                          | Scale Mean<br>if Item<br>Deleted | Scale<br>Variance if<br>Item Deleted | Corrected<br>Item-Total<br>Correlation | Cronbach's<br>Alpha if Item<br>Deleted |
|--------------------------------------------------------------------------|----------------------------------|--------------------------------------|----------------------------------------|----------------------------------------|
| S1.I agree to use the WFAT unconditionally                               | 70.76                            | 29.978                               | .236                                   | .792                                   |
| S2.I agree to use the WFAT if required                                   | 70.17                            | 30.090                               | .245                                   | .791                                   |
| S3.I agree to use the WFAT if it was proved<br>useful to learning        | 70.25                            | 30.122                               | .309                                   | .787                                   |
| S4.I agree to use the WFAT that suits my needs                           | 70.38                            | 30.813                               | .171                                   | .794                                   |
| S5.I agree to use of WFAT in oral English learning                       | 70.17                            | 28.412                               | .591                                   | .772                                   |
| S6.I agree to use WFAT in other courses                                  | 70.82                            | 30.452                               | .255                                   | .790                                   |
| S7.I agree to start using WFAT now                                       | 70.69                            | 28.843                               | .433                                   | .780                                   |
| S8.I want to use WFAT to increase engagement in<br>autonomous learning   | 70.32                            | 27.795                               | .552                                   | .772                                   |
| S9.I want to use WFAT to increase interest in<br>learning                | 70.47                            | 29.980                               | .335                                   | .786                                   |
| S10.I want to use WFAT to increase joys in<br>learning                   | 70.41                            | 29.363                               | .381                                   | .783                                   |
| S11.I want to use WFAT to improve confidence in<br>learning              | 70.63                            | 29.337                               | .441                                   | .780                                   |
| S13.I want to use WFAT to increase interaction<br>with peers remotely    | 70.65                            | 27.840                               | .494                                   | .775                                   |
| S14.I want to use WFAT to increase interaction<br>with teachers remotely | 70.60                            | 28.344                               | .409                                   | .781                                   |
| S15.I want to use WFAT to enhance the<br>personalisation of learning     | 70.73                            | 29.147                               | .451                                   | .779                                   |
| S16.I want to use WFAT to enhance the autonomy<br>of learning            | 70.47                            | 30.099                               | .281                                   | .789                                   |
| S17.I want to use WFAT to enhance the<br>convenience of learning         | 70.15                            | 30.045                               | .303                                   | .787                                   |
| S18.I wish I could easily get permission to use<br>WFAT                  | 70.28                            | 29.117                               | .409                                   | .781                                   |
| S19.I wish to use WFAT in class                                          | 70.59                            | 29.515                               | .341                                   | .785                                   |
| S20.I wish to use WFAT in autonomous learning<br>after class             | 70.32                            | 30.050                               | .243                                   | .791                                   |
| S21.I wish to choose a WFAT that is suitable for<br>my needs             | 70.35                            | 30.061                               | .277                                   | .789                                   |
| S22.I wish to use WFAT for free                                          | 70.29                            | 30.816                               | .175                                   | .794                                   |

**Appendix 3 Raw Data of Students' Engagement of the 14 Variables in OEAL & FA**

| Participants | Group | Data of the 14 variables of students' OEAL & FA |     |       |    |    |    |    |    |     |     |     |     |     |     |
|--------------|-------|-------------------------------------------------|-----|-------|----|----|----|----|----|-----|-----|-----|-----|-----|-----|
|              |       | V1                                              | V2  | V3    | V4 | V5 | V6 | V7 | V8 | V9  | V10 | V11 | V12 | V13 | V14 |
| ST1          | EG    | 1471                                            | 712 | 21.00 | 30 | 36 | 27 | 18 | 17 | 149 | 76  | 76  | 0   | 1   | 46  |
| ST2          | EG    | 1617                                            | 728 | 22.00 | 36 | 26 | 27 | 18 | 21 | 150 | 113 | 113 | 2   | 2   | 35  |
| ST3          | EG    | 1376                                            | 742 | 21.00 | 29 | 36 | 17 | 17 | 16 | 136 | 76  | 76  | 0   | 0   | 46  |
| ST4          | EG    | 1570                                            | 724 | 40.00 | 44 | 16 | 15 | 12 | 11 | 138 | 100 | 111 | 2   | 2   | 44  |
| ST5          | EG    | 1435                                            | 622 | 36.00 | 44 | 37 | 25 | 9  | 11 | 162 | 98  | 98  | 3   | 0   | 51  |
| ST6          | EG    | 1411                                            | 663 | 22.00 | 41 | 38 | 36 | 22 | 6  | 165 | 112 | 112 | 2   | 1   | 34  |
| ST7          | EG    | 1354                                            | 639 | 22.00 | 37 | 30 | 23 | 19 | 18 | 149 | 107 | 107 | 0   | 0   | 32  |
| ST8          | EG    | 1387                                            | 599 | 22.00 | 43 | 22 | 19 | 21 | 19 | 146 | 107 | 107 | 0   | 3   | 40  |
| ST9          | EG    | 1731                                            | 756 | 22.00 | 41 | 42 | 22 | 21 | 21 | 169 | 110 | 110 | 0   | 3   | 49  |
| ST10         | EG    | 1394                                            | 615 | 44.00 | 44 | 22 | 22 | 0  | 0  | 132 | 93  | 93  | 2   | 0   | 30  |
| ST11         | EG    | 1342                                            | 636 | 22.00 | 37 | 22 | 22 | 8  | 13 | 124 | 70  | 70  | 2   | 4   | 34  |
| ST12         | EG    | 1619                                            | 718 | 22.00 | 40 | 23 | 39 | 20 | 17 | 161 | 94  | 94  | 3   | 0   | 92  |
| ST13         | EG    | 1520                                            | 701 | 34.00 | 40 | 22 | 34 | 11 | 15 | 156 | 93  | 93  | 0   | 0   | 62  |
| ST14         | EG    | 1378                                            | 645 | 22.00 | 22 | 22 | 39 | 10 | 12 | 127 | 75  | 75  | 3   | 3   | 66  |
| ST15         | EG    | 1368                                            | 680 | 40.00 | 46 | 41 | 18 | 9  | 8  | 162 | 111 | 111 | 0   | 0   | 40  |
| ST16         | EG    | 1447                                            | 653 | 22.00 | 32 | 32 | 27 | 19 | 15 | 147 | 93  | 93  | 2   | 3   | 75  |
| ST17         | EG    | 1764                                            | 733 | 22.00 | 58 | 36 | 45 | 29 | 23 | 213 | 149 | 149 | 0   | 0   | 32  |
| ST18         | EG    | 1640                                            | 744 | 22.00 | 37 | 22 | 41 | 19 | 19 | 160 | 108 | 108 | 2   | 0   | 43  |
| ST19         | EG    | 1717                                            | 784 | 22.00 | 39 | 22 | 41 | 22 | 23 | 169 | 121 | 121 | 0   | 2   | 51  |
| ST20         | EG    | 1689                                            | 821 | 22.00 | 44 | 27 | 42 | 27 | 23 | 185 | 116 | 116 | 2   | 0   | 49  |
| ST21         | EG    | 1394                                            | 629 | 22.00 | 32 | 22 | 25 | 11 | 9  | 121 | 76  | 76  | 0   | 2   | 103 |
| ST22         | EG    | 1472                                            | 689 | 22.00 | 37 | 21 | 36 | 15 | 11 | 142 | 97  | 97  | 0   | 0   | 53  |
| ST23         | CG    | 1160                                            | 540 | 29    | 29 | 0  | 29 | 0  | 0  | 22  | 87  | 22  | 0   | 2   | 0   |
| ST24         | CG    | 1280                                            | 663 | 41    | 36 | 0  | 32 | 0  | 0  | 22  | 109 | 22  | 2   | 1   | 6   |

**Appendix 3 (continued)**

| Participants | Group | Data of the 14 variables of students' OEAL & FA |     |    |    |    |    |    |    |    |     |     |     |     |     |
|--------------|-------|-------------------------------------------------|-----|----|----|----|----|----|----|----|-----|-----|-----|-----|-----|
|              |       | V1                                              | V2  | V3 | V4 | V5 | V6 | V7 | V8 | V9 | V10 | V11 | V12 | V13 | V14 |
| ST25         | CG    | 1360                                            | 690 | 43 | 43 | 0  | 45 | 3  | 8  | 22 | 131 | 22  | 2   | 2   | 0   |
| ST26         | CG    | 1000                                            | 505 | 28 | 37 | 1  | 27 | 0  | 3  | 22 | 92  | 22  | 3   | 3   | 0   |
| ST27         | CG    | 1040                                            | 589 | 27 | 47 | 0  | 40 | 0  | 0  | 22 | 114 | 22  | 0   | 0   | 5   |
| ST28         | CG    | 1120                                            | 609 | 28 | 46 | 0  | 36 | 2  | 0  | 22 | 110 | 22  | 0   | 0   | 9   |
| ST29         | CG    | 1080                                            | 577 | 27 | 40 | 0  | 32 | 0  | 3  | 22 | 99  | 22  | 0   | 3   | 0   |
| ST30         | CG    | 1440                                            | 804 | 46 | 64 | 0  | 47 | 7  | 0  | 22 | 157 | 22  | 0   | 0   | 5   |
| ST31         | CG    | 1360                                            | 766 | 41 | 58 | 1  | 51 | 0  | 0  | 22 | 150 | 22  | 3   | 3   | 0   |
| ST32         | CG    | 1280                                            | 697 | 32 | 54 | 0  | 40 | 0  | 7  | 22 | 126 | 22  | 3   | 0   | 8   |
| ST33         | CG    | 840                                             | 384 | 26 | 21 | 15 | 21 | 10 | 9  | 21 | 68  | 21  | 0   | 2   | 23  |
| ST34         | CG    | 800                                             | 410 | 29 | 36 | 0  | 20 | 0  | 0  | 20 | 85  | 20  | 0   | 0   | 0   |
| ST35         | CG    | 840                                             | 455 | 29 | 42 | 2  | 21 | 13 | 0  | 21 | 92  | 21  | 0   | 2   | 34  |
| ST36         | CG    | 840                                             | 481 | 29 | 46 | 0  | 21 | 0  | 0  | 21 | 96  | 21  | 2   | 3   | 15  |
| ST37         | CG    | 1320                                            | 754 | 46 | 55 | 1  | 51 | 2  | 0  | 22 | 152 | 22  | 3   | 0   | 0   |
| ST38         | CG    | 1240                                            | 714 | 44 | 54 | 0  | 43 | 13 | 3  | 22 | 141 | 22  | 0   | 2   | 13  |
| ST39         | CG    | 880                                             | 465 | 33 | 44 | 1  | 22 | 0  | 0  | 22 | 99  | 22  | 3   | 0   | 0   |
| ST40         | CG    | 1000                                            | 571 | 25 | 60 | 6  | 28 | 1  | 3  | 22 | 113 | 22  | 0   | 1   | 24  |
| ST41         | CG    | 960                                             | 579 | 37 | 50 | 2  | 26 | 0  | 0  | 22 | 113 | 22  | 2   | 0   | 0   |
| ST42         | CG    | 880                                             | 447 | 31 | 49 | 0  | 22 | 8  | 0  | 22 | 102 | 22  | 0   | 0   | 17  |
| ST43         | CG    | 1360                                            | 769 | 49 | 57 | 0  | 51 | 2  | 3  | 22 | 157 | 22  | 2   | 1   | 12  |
| ST44         | CG    | 840                                             | 489 | 34 | 46 | 1  | 21 | 0  | 0  | 21 | 101 | 21  | 0   | 0   | 0   |

*Variable 1: Time spending; Variable 2: Time of submitted recording; Variable 3: Number of comprehensive tasks completed; Variable 4: The number of single tasks completed with pronunciation accuracy; Variable 5: Number of fluency tasks completed; Variable 6: The number of lexical and grammar tasks completed; Variable 7: Number of tasks completed in practising coherence of oral English; Variable 8: Number of tasks completed in practising; Variable 9: The number of times you get instant feedback; Variable 10: Number of repeated tasks; Variable 11: Number of policy adjustments; Variable 12: The number of times you rate other students; Variable 13: The number of times you were graded by other students; Variable 14: number of peer interactions.*

**Appendix 4 Raw Data of Students' Scores in the Five Oral English Tests**

| Parti-<br>pants | Group | Monthly Test1   |              |                        |      | Monthly Test2   |              |                        |     | Monthly Test3   |              |                        |     | Monthly Test4   |              |                        |     | Monthly Test5   |              |                        |      |
|-----------------|-------|-----------------|--------------|------------------------|------|-----------------|--------------|------------------------|-----|-----------------|--------------|------------------------|-----|-----------------|--------------|------------------------|-----|-----------------|--------------|------------------------|------|
|                 |       | FiF's<br>rating | My<br>rating | Martin<br>'s<br>rating | Avg. | FiF's<br>rating | My<br>rating | Martin<br>'s<br>rating | Avg | FiF's<br>rating | My<br>rating | Martin<br>'s<br>rating | Avg | FiF's<br>rating | My<br>rating | Martin<br>'s<br>rating | Avg | FiF's<br>rating | My<br>rating | Martin<br>'s<br>rating | Avg. |
| ST1             | EG    | 86              | 80           | 83                     | 83   | 80              | 76           | 78                     | 78  | 88              | 84           | 82                     | 85  | 89              | 86           | 88                     | 88  | 92              | 89           | 90                     | 90   |
| ST2             | EG    | 77              | 71           | 74                     | 74   | 75              | 73           | 75                     | 74  | 81              | 78           | 80                     | 80  | 84              | 81           | 83                     | 83  | 83              | 79           | 80                     | 81   |
| ST3             | EG    | 92              | 87           | 89                     | 89   | 85              | 81           | 83                     | 83  | 95              | 89           | 87                     | 90  | 93              | 85           | 85                     | 88  | 95              | 93           | 94                     | 94   |
| ST4             | EG    | 79              | 74           | 76                     | 76   | 77              | 73           | 75                     | 75  | 87              | 80           | 79                     | 82  | 92              | 89           | 91                     | 91  | 88              | 84           | 88                     | 87   |
| ST5             | EG    | 91              | 85           | 88                     | 88   | 90              | 86           | 88                     | 88  | 93              | 89           | 87                     | 90  | 95              | 86           | 88                     | 90  | 93              | 89           | 90                     | 91   |
| ST6             | EG    | 94              | 89           | 91                     | 91   | 86              | 80           | 82                     | 83  | 96              | 92           | 90                     | 93  | 95              | 92           | 94                     | 94  | 96              | 92           | 93                     | 94   |
| ST7             | EG    | 94              | 87           | 91                     | 91   | 88              | 84           | 86                     | 86  | 93              | 89           | 87                     | 90  | 96              | 93           | 95                     | 95  | 95              | 92           | 85                     | 91   |
| ST8             | EG    | 89              | 84           | 86                     | 86   | 85              | 80           | 82                     | 82  | 95              | 91           | 89                     | 92  | 95              | 88           | 90                     | 91  | 97              | 93           | 92                     | 94   |
| ST9             | EG    | 86              | 83           | 83                     | 84   | 86              | 82           | 84                     | 84  | 92              | 88           | 88                     | 89  | 93              | 90           | 92                     | 92  | 95              | 91           | 92                     | 93   |
| ST10            | EG    | 77              | 72           | 74                     | 74   | 79              | 77           | 79                     | 78  | 76              | 72           | 70                     | 73  | 75              | 72           | 75                     | 74  | 79              | 78           | 79                     | 79   |
| ST11            | EG    | 83              | 79           | 80                     | 81   | 85              | 81           | 83                     | 83  | 90              | 86           | 85                     | 87  | 89              | 84           | 86                     | 86  | 90              | 86           | 89                     | 88   |
| ST12            | EG    | 91              | 86           | 88                     | 88   | 90              | 86           | 88                     | 88  | 89              | 85           | 83                     | 86  | 91              | 88           | 93                     | 91  | 92              | 88           | 89                     | 90   |
| ST13            | EG    | 88              | 80           | 85                     | 84   | 85              | 79           | 81                     | 82  | 96              | 92           | 91                     | 93  | 95              | 80           | 82                     | 86  | 95              | 90           | 92                     | 92   |
| ST14            | EG    | 93              | 88           | 90                     | 90   | 88              | 84           | 86                     | 86  | 93              | 89           | 87                     | 90  | 92              | 88           | 90                     | 90  | 96              | 93           | 94                     | 94   |
| ST15            | EG    | 89              | 84           | 86                     | 86   | 92              | 88           | 90                     | 90  | 95              | 91           | 89                     | 92  | 95              | 92           | 90                     | 92  | 96              | 92           | 93                     | 94   |
| ST16            | EG    | 86              | 81           | 83                     | 83   | 83              | 79           | 81                     | 81  | 88              | 84           | 84                     | 85  | 90              | 87           | 89                     | 89  | 89              | 90           | 91                     | 90   |
| ST17            | EG    | 85              | 80           | 82                     | 82   | 82              | 80           | 82                     | 81  | 89              | 85           | 83                     | 86  | 90              | 83           | 85                     | 86  | 92              | 88           | 89                     | 90   |
| ST18            | EG    | 82              | 77           | 79                     | 79   | 85              | 81           | 83                     | 83  | 88              | 84           | 82                     | 85  | 90              | 87           | 89                     | 89  | 87              | 83           | 86                     | 85   |
| ST19            | EG    | 86              | 81           | 83                     | 83   | 85              | 82           | 84                     | 84  | 92              | 88           | 88                     | 89  | 93              | 90           | 92                     | 92  | 91              | 89           | 90                     | 90   |
| ST20            | EG    | 76              | 73           | 73                     | 74   | 75              | 71           | 73                     | 73  | 75              | 71           | 69                     | 72  | 81              | 80           | 83                     | 81  | 79              | 75           | 76                     | 77   |
| ST21            | EG    | 90              | 85           | 87                     | 87   | 86              | 83           | 85                     | 85  | 94              | 90           | 88                     | 91  | 94              | 91           | 93                     | 93  | 93              | 89           | 90                     | 91   |
| ST22            | EG    | 83              | 80           | 80                     | 81   | 84              | 80           | 82                     | 82  | 89              | 85           | 80                     | 85  | 87              | 84           | 86                     | 86  | 91              | 87           | 88                     | 89   |
| ST23            | CG    | 85              | 80           | 82                     | 82   | 82              | 78           | 80                     | 80  | 86              | 82           | 80                     | 83  | 88              | 82           | 86                     | 85  | 90              | 88           | 89                     | 89   |

**Appendix 4 (continued)**

| Parti-<br>pants | Group | Monthly Test1   |              |                        |      | Monthly Test2   |              |                        |     | Monthly Test3   |              |                        |     | Monthly Test4   |              |                        |     | Monthly Test5   |              |                        |      |
|-----------------|-------|-----------------|--------------|------------------------|------|-----------------|--------------|------------------------|-----|-----------------|--------------|------------------------|-----|-----------------|--------------|------------------------|-----|-----------------|--------------|------------------------|------|
|                 |       | FiF's<br>rating | My<br>rating | Martin<br>'s<br>rating | Avg. | FiF's<br>rating | My<br>rating | Martin<br>'s<br>rating | Avg | FiF's<br>rating | My<br>rating | Martin<br>'s<br>rating | Avg | FiF's<br>rating | My<br>rating | Martin<br>'s<br>rating | Avg | FiF's<br>rating | My<br>rating | Martin<br>'s<br>rating | Avg. |
| ST24            | CG    | 87              | 82           | 84                     | 84   | 86              | 82           | 84                     | 84  | 88              | 84           | 83                     | 85  | 84              | 81           | 83                     | 83  | 90              | 86           | 87                     | 88   |
| ST25            | CG    | 88              | 83           | 85                     | 85   | 85              | 80           | 82                     | 82  | 90              | 86           | 84                     | 87  | 92              | 89           | 91                     | 91  | 90              | 86           | 88                     | 88   |
| ST26            | CG    | 87              | 80           | 84                     | 84   | 85              | 81           | 83                     | 83  | 86              | 82           | 80                     | 83  | 90              | 87           | 89                     | 89  | 88              | 84           | 85                     | 86   |
| ST27            | CG    | 86              | 81           | 83                     | 83   | 85              | 81           | 83                     | 83  | 86              | 82           | 81                     | 83  | 85              | 82           | 84                     | 84  | 90              | 86           | 87                     | 88   |
| ST28            | CG    | 90              | 85           | 87                     | 87   | 92              | 88           | 90                     | 90  | 90              | 86           | 84                     | 87  | 95              | 92           | 94                     | 94  | 91              | 87           | 88                     | 89   |
| ST29            | CG    | 78              | 73           | 75                     | 75   | 75              | 71           | 73                     | 73  | 80              | 76           | 74                     | 77  | 84              | 81           | 83                     | 83  | 85              | 81           | 85                     | 84   |
| ST30            | CG    | 84              | 79           | 81                     | 81   | 83              | 80           | 82                     | 82  | 86              | 82           | 80                     | 83  | 85              | 82           | 84                     | 84  | 88              | 84           | 85                     | 86   |
| ST31            | CG    | 84              | 79           | 81                     | 81   | 87              | 83           | 85                     | 85  | 93              | 89           | 87                     | 90  | 90              | 85           | 87                     | 87  | 88              | 84           | 85                     | 86   |
| ST32            | CG    | 82              | 77           | 79                     | 79   | 84              | 80           | 82                     | 82  | 87              | 83           | 81                     | 84  | 85              | 82           | 85                     | 84  | 87              | 83           | 84                     | 85   |
| ST33            | CG    | 85              | 80           | 82                     | 82   | 83              | 80           | 82                     | 82  | 90              | 86           | 84                     | 87  | 89              | 86           | 88                     | 88  | 92              | 88           | 89                     | 90   |
| ST34            | CG    | 83              | 78           | 80                     | 80   | 82              | 78           | 80                     | 80  | 85              | 81           | 80                     | 82  | 86              | 83           | 85                     | 85  | 87              | 83           | 84                     | 85   |
| ST35            | CG    | 86              | 81           | 83                     | 83   | 85              | 81           | 83                     | 83  | 92              | 88           | 86                     | 89  | 90              | 85           | 87                     | 87  | 94              | 90           | 91                     | 92   |
| ST36            | CG    | 91              | 87           | 88                     | 89   | 90              | 86           | 88                     | 88  | 88              | 84           | 82                     | 85  | 92              | 87           | 89                     | 89  | 90              | 86           | 87                     | 88   |
| ST37            | CG    | 85              | 80           | 82                     | 82   | 80              | 76           | 78                     | 78  | 87              | 83           | 81                     | 84  | 91              | 88           | 90                     | 90  | 90              | 86           | 86                     | 87   |
| ST38            | CG    | 87              | 82           | 84                     | 84   | 86              | 82           | 84                     | 84  | 85              | 81           | 79                     | 82  | 89              | 82           | 84                     | 85  | 87              | 83           | 84                     | 85   |
| ST39            | CG    | 84              | 79           | 81                     | 81   | 85              | 81           | 83                     | 83  | 90              | 86           | 79                     | 85  | 86              | 81           | 80                     | 82  | 90              | 85           | 86                     | 87   |
| ST40            | CG    | 87              | 82           | 84                     | 84   | 85              | 80           | 82                     | 82  | 87              | 83           | 81                     | 84  | 85              | 80           | 82                     | 82  | 87              | 83           | 85                     | 85   |
| ST41            | CG    | 92              | 88           | 89                     | 90   | 90              | 86           | 88                     | 88  | 95              | 91           | 88                     | 91  | 88              | 85           | 87                     | 87  | 95              | 87           | 88                     | 90   |
| ST42            | CG    | 86              | 81           | 83                     | 83   | 88              | 84           | 86                     | 86  | 87              | 83           | 81                     | 84  | 88              | 85           | 87                     | 87  | 89              | 84           | 85                     | 86   |
| ST43            | CG    | 90              | 86           | 87                     | 88   | 88              | 85           | 87                     | 87  | 85              | 81           | 80                     | 82  | 89              | 87           | 85                     | 87  | 89              | 85           | 86                     | 87   |
| ST44            | CG    | 77              | 72           | 74                     | 74   | 75              | 71           | 73                     | 73  | 78              | 74           | 75                     | 76  | 79              | 73           | 73                     | 75  | 85              | 81           | 82                     | 83   |

## Appendix 5 Records of Class Observation on Students Oral English Performance

### 课堂观察记录 (Chinese)

#### 第一次观察记录(实验组) 28 Dec 2022

##### 1.学生在课堂上的参与度和表现

整体面貌描述: 大部分学生在课堂上状态还不错, 积极参与到课堂中的任务中, 并和老师进行互动。  
存在的问题: 课上有大约 5-6 个学生不太积极不主动, 他们不参与到课堂活动中, 也不直视教师, 不和教师互动。有 2-3 个学生心不在焉, 他们经常低下头沉默, 或往窗外望去, 或玩手机。

##### 2.口语表达的准确性

发音的准确性:

大部分学生在英语口语表达中的大部分发音是正确的。但有些学生在部分元音的发音上错误, 比如将 /i:/ 被误发成短音元音 /ɪ/, 部分学生在 /ɔɪ//əʊ//əʊ/ 等双元音发音方面存在错误。有些学生在摩擦音 (如 /f/ /v/ ) 发音上出错。

词汇运用的准确性:

大部分学生能使用大纲中的词汇来实现英语口语表达。但很多学生词汇量有限, 存在过度使用熟悉的词汇的问题, 在英语口语表达中都只能使用比较简单的词汇, 且反复使用某些熟悉的词汇, 导致表达单一、重复。例如, 学生频繁使用 good 来表达各种积极的意思, 而忽视了 excellent、fantastic 等更恰当的替代词。

语法正确性:

学生在口语表达时, 经常在时态、语态、主谓一致、介词等方面出现表述错误的次数较多。

##### 3.口语表达的流利性

断断续续: 学生在表达时, 每句话之间会有较长的停顿, 例如说一句 “I went to the store...”, 然后停顿好几秒, 再继续说 “and... I... bought... some... milk”, 造成表达不连贯。

磕磕巴巴: 学生常常因为想不起词汇或语法规则, 导致中途停顿、重复某个词或打破语句流畅性。例如, 学生说 “I want... I want... to go... to... to... the park”。

句子之间衔接欠佳: 学生在说不同句子时, 无法使用适当的衔接词, 导致句子显得生硬。比如 “I like pizza. It is good. I eat it every day”, 句子间没有合理的衔接词或关联词, 表达显得生硬。

长时间停顿: 在表达复杂句子或观点时, 学生往往会有较长时间的停顿, 思考接下来要说什么, 造成听者的理解困难。例如: “Yesterday I... uh... I... went to the cinema with my... uh... with my friend”。

重复使用简单连接词: 学生在表达不同想法时, 频繁使用 “and” 来连接句子或短语, 缺乏多样性。例如: “I went to the park and I played football and I met my friend and we ate lunch”。

##### 4.口语表达的连贯性和完整性

学生的口语表达明显缺乏连贯性和完整性。句子之间没有合理的衔接, 表达内容显得支离破碎, 听者难以理解学生的意图。

例 1: 学生说 “I like pizza. I eat pizza. My friend likes pizza too”, 这些句子之间没有任何衔接词或逻辑连接, 内容显得生硬。

例 2: 学生表达不完整: “Yesterday, I... um... went to... um... the store”, 因频繁停顿, 句子不连贯, 信息不完整。

#### 第一次观察记录(对照组) 29 Dec 2022

##### 1.学生在课堂上的参与度和表现

整体面貌描述: 大部分学生在课堂表现较好, 他们能集中注意力, 积极参与到课堂中的任务中, 并和老师进行互动。

存在的问题: 有部分学生需要老师点名才会回答问题或参与课堂讨论, 个别学生被点名后仍不愿意回答教师的提问。有 7-8 个学生在用手机在玩或在做其他与上课内容无关的事情 (如写其他科目的作业, 看故事书或睡觉)。

##### 2.口语表达的准确性

发音的准确性: 大部分学生在英语口语表达中的大部分发音是正确的。但有些学生在部分元音的发音

上错误，比如将/i:/ 被误发成短音元音/ɪ/，部分学生在/ɔɪ//aʊ//əʊ/等双元音发音方面存在错误。有些学生在摩擦音（如/f/ /v/ ）发音上出错。

词汇运用的准确性：很多学生词汇量有限，存在过度使用熟悉的词汇的问题，在英语口语表达中都只能使用比较简单的词汇，且反复使用某些熟悉的词汇，导致表达单一、重复。有些学生在表达中存在词不达意或单词使用错误（例如，将 fun 和 funny ， lend 和 borrow 混用）。

语法正确性：学生在口语表达时，经常在时态、语态、主谓一致、介词等等方面出现表述错误。

### 3.口语表达的流利性

频繁停顿，思路不清：学生说“I... I went to... um... the market... uh... to buy... uh... food”，在表达过程中多次停顿，明显感到学生在寻找合适的词汇或句子结构，导致思路不连贯。

断断续续，表达内容不完整：学生想表达某个经历，但话语中断，“Last weekend... I with my friend... we... go to... but... didn't go”，听者无法理解完整意思，表达内容缺失严重。

句子之间缺乏衔接：学生说“I like football. It is a sport. I play on the weekend. My team is good”，句子之间没有使用适当的衔接词，信息之间的关系不清晰，显得没有连贯性。

缺乏逻辑感：学生在描述一个活动时，顺序和逻辑混乱：“First we eat lunch. After that we... go to the cinema. I buy a ticket... then... we walk... and eat lunch”。其中描述顺序前后不一致，让人难以理解。

连贯性差：学生在回答问题时，频繁使用不连贯的短语，如“I want... I think... uh... we... should... because... but... uh... I don't know”，没有形成完整的句子，表达不流畅，缺乏逻辑连贯性。

### 4.口语表达的连贯性和完整性

在第一次观察中，对照组学生的口语表达缺乏连贯性和完整性。他们在表达过程中句子之间的衔接几乎不存在，内容显得支离破碎，听者很难捕捉到完整的信息。

例 1：学生说“I like chocolate. Chocolate is sweet”。在这两句中，学生虽然表达了自己的喜好，但缺乏任何连接词，内容重复且过于简单，未能展开具体的观点或提供更多信息。

例 2：学生表达不完整：“Yesterday, I went to... um... the school... for...”。由于频繁停顿，句子显得不连贯，听者无法理解学生想要表达的具体内容。

## 第二次观察记录（实验组）8 Feb 2023

### 1.学生在课堂上的参与度和表现

整体面貌描述：学生们在课堂上总体上表现较好，大部分学生能专注课堂并积极参与到课堂中的任务中。大部分学生表现得比较自信，他们和老师进行互动频繁。

存在的问题：少量学生（大概 5 个）不太积极，who 不主动回答教师提问或不主动参与课堂活动。有 4 个学生在上课期间一直在玩手机来看电影或打游戏。

### 2.口语表达的准确性

发音的准确性：有学生在部分元音发音上出错（如/æ//ɑ:/的发音时候开口不够大）；有些学生把 learn 中“ear”的发音/ɜ:/发成了/ə/；部分学生在双元音发音方面存在明显的错误。有学生在摩擦音和 Lateral 的发音上出错。

词汇运用的准确性：很多学生偏向于使用自己熟悉且简单的词汇来表达，导致表达单一、重复。有些学生在表达中存在词不达意或单词使用错误（例如，将 fun 和 funny ， lend 和 borrow 混用）。有些学生在表述时词汇搭配不当（如 do a mistake-make a mistake, big rain-heavy rain, do a meal-make a meal, pay attention on - pay attention to）。

语法正确性：学生在口语表达时，经常在时态、语态、主谓一致、介词等等方面出现表述错误仍然非常明显。

### 3.口语表达的流利性

顺序之间无连接词：学生说“I go to school. I study math. I eat lunch. I come home”，句子之间缺乏连接词，没有表达出句子之间的逻辑关系或时间顺序。

逻辑不清：学生在表达观点时说“I think... um... school is important... because... uh... we learn things... but... sometimes it's boring... um... I don't know”，表达内容前后不一致，逻辑混乱。

时态错误，混淆表达：学生说“Yesterday I go to the store and buy milk. Then I go home and I drink the milk”，

时态错误使得表达内容难以理解, 缺乏逻辑。

缺乏衔接词导致信息跳跃: 学生描述假期经历时说“I went to the beach. The sun was bright. We swim. The food was delicious”, 每个句子都没有衔接词, 内容显得跳跃, 没有连贯性。

词汇匮乏导致重复: 学生想描述一天的活动, “I... I do homework. Then I do... um... I do more homework. Then... I do other things... like... homework”, 词汇不足导致重复, 内容显得单调、混乱。

#### 4. 口语表达的连贯性和完整性

第二次观察显示, 学生的表达仍然缺乏连贯性和完整性。尽管能说出一些独立的句子, 但句子之间依然缺乏逻辑关系或适当的连接。

例 1: 学生说“I have a dog. My dog is big. We play outside”, 虽然每个句子都是独立的完整句子, 但缺乏逻辑衔接, 句子之间无连贯性。

例 2: 学生说“Last weekend, I went to my... um... friend's house”, 句子结构不完整, 听者无法得到足够的信息。

### 第二次观察记录 (对照组) 9 Feb 2023

#### 1. 学生在课堂上的参与度和表现

描述: 大部分学生在课堂表现较好, 他们能集中注意力, 积极参与到课堂中的任务中, 并和老师进行互动。

存在的问题: 有部分学生需要老师点名才会回答问题或参与课堂讨论, 个别学生被点名后仍不愿意回答教师的提问。有 7-8 个学生在用手机在玩或在做其他与上课内容无关的事情 (如写其他科目的作业, 看故事书或睡觉)。

#### 2. 口语表达的准确性

发音的准确性: 学生在部分单元音发音上出错, 特别是长元音和短元音经常被他们混淆; 部分学生在双元音发音方面存在明显的错误, 双元音发音不饱满。学生在摩擦音和 Lateral 的发音上经常出错。

词汇运用的准确性: 学生喜欢反复使用自己熟悉且简单的词汇来表达, 导致表述单一、重复。学生在表达中存在词不达意或单词使用错误、词汇搭配不当等情况经常发生。

语法正确性: 学生在口语表达时, 经常在时态、语态、主谓一致、介词等等方面出现表述错误仍然非常明显。

#### 3. 口语表达的流利性

无序重复: 学生说“Yesterday I... yesterday I went to... I went to the park... um... yesterday”, 同一个词或短语被重复多次, 表达混乱, 显得毫无逻辑。

表达不连贯: 学生想描述周末活动, “On Saturday... I... with my friend... um... we... we go... park... and... um... football”, 句子缺失必要的成分, 无法形成连贯的表达。

句子顺序混乱: 学生描述自己的日常作息时说“I wake up... I go to school... after school I eat breakfast... then I sleep”, 活动顺序颠倒, 缺乏逻辑。

依赖简单词汇和重复: 学生说“I like movies. And... um... I like movies... and... um... I watch movies... and... um... movies are good”, 过度依赖“and”和重复的表达, 逻辑不清晰。

长时间停顿: 学生在试图解释一件事情时说“I... uh... last week I... um... go to... uh... library”, 每个句子之间有很长的停顿, 使得听者难以理解。

#### 4. 口语表达的连贯性和完整性

在第二次观察中, 学生的口语表达仍然缺乏连贯性和完整性, 尽管他们尝试使用简单句子, 但句子之间的逻辑关系依然薄弱, 整体表达显得不流畅。

例 1: 学生说“I have a cat. The cat is black”。虽然这两句都是独立的完整句子, 但缺乏进一步的描述或连接, 整体表达显得单调, 无法给听者提供更多的背景信息。

例 2: 学生说“Last weekend, I went... um... out with my... um... friends”。在这句话中, 虽然提到与朋友的活动, 但因停顿导致信息不完整, 听者无法理解具体的活动内容。

### 第三次观察记录 (实验组) 2 Mar 2023

### 1.学生在课堂上的参与度和表现

描述：在课堂上，大多数学生展现出高度的参与意识，积极完成课堂任务，并且能够很好地遵守课堂纪律。他们专心听讲，表现出对课程内容的浓厚兴趣，课堂互动也比较活跃。

存在的问题：少部分学生在课堂上显得有些被动，他们不愿与教师或同学进行深入讨论。有些学生会时不时低头看手机或在课桌下玩游戏，甚至有学生因为疲惫打瞌睡，影响了学习效率。

### 2.口语表达的准确性

发音的准确性：部分学生在部分元音和部分辅音（如 /θ//ð//ʒ//l//v/）的发音上出错。有些学生在发 /m/ 时嘴巴未关闭。学生在长单词发音上存在较大困难，词重音把握不准，有学生在句子的语调上出错。

词汇运用的准确性：学生在不同语境中重复使用大纲中的基础词汇来表达，缺乏词汇使用的灵活性和多样性；有些学生在表达中存在词不达意，很多学生在表达中频繁使用中式英语来表达。

语法正确性：学生在口语表达时，经常在时态、语态、主谓一致、介词、冠词等方面出现表述错误。有部分学生刻意地使用复杂句来取代简单句但最后出错了。

### 3.口语表达的流利性

频繁停顿。学生说“I... um... went to... the park... and... uh... played football”，每个句子和短语之间有明显的停顿。学生表达“Last week... I... uh... studied for my test”，因为思路不清而频繁停顿。

表达经常重复。学生说“I like... I like... I like playing basketball”，无法顺利表达完整的句子。

在表达观点时，重复词汇：“The movie is... is... very interesting. It's... it's fun”。

句子衔接差，学生说“I like soccer. It's fun. I play every weekend”，句子之间缺乏连贯的逻辑或衔接词。

描述一个事件时，缺少衔接词：“I went to school. I did homework. I watched TV”。

语速很慢。学生缓慢地说“I... went... to... the store... and... bought... some... food”，语速非常慢，听者难以保持注意力。学生表达时速度极慢：“I... uh... think... the answer... is... right”。

### 4.口语表达的连贯性和完整性

在第三次观察中，学生的表达稍有改善，但连贯性和完整性问题依然存在。学生尝试使用连接词，但经常使用不当或停顿过多。

例 1：学生说“I like reading... and... um... football. It is fun”，学生试图使用“and”进行连接，但停顿打断了表达的连贯性。

例 2：学生说“I go to school... then... um... I study math”，虽然句子结构完整，但由于不必要的停顿，影响了表达的连贯性。

## 第三次观察记录（对照组）2 Mar 2023

### 1.学生在课堂上的参与度和表现

描述：在课堂上，大多数学生积极参与教学活动，能够集中注意力并有效配合老师的教学节奏，课堂气氛较为活跃。

存在的问题：部分学生显得缺乏参与动力，不仅不主动回答问题，甚至在小组讨论时也保持沉默。有些学生会在上课时频繁查看手机，甚至偷偷玩游戏，还有些学生在课堂上时常分神，心不在焉地望向窗外，完全游离于课堂之外。此外，有学生时不时借口上厕所离开教室，长时间不归，打断了课堂节奏。

### 2.口语表达的准确性

发音的准确性：学生在部分元音和部分辅音的发音上发生错误频繁。学生在单词重音上经常犯错，在句子的语调上也经常出错。

词汇运用的准确性：学生单一、重复地使用简单的词汇来表达不同的语境内容，缺乏词汇使用的灵活性和多样性；学生在词汇搭配和上经常犯错，很多学生在表达中频繁使用中式英语来表达。

语法正确性：学生在口语表达时，经常在时态、语态、主谓一致、介词、冠词等方面出现表述错误。

### 3.口语表达的流利性

表达中频繁停顿。在表达复杂句子时，学生说“I want... to go to... uh... the beach... because... uh... it's sunny”。

学生在回答问题时停顿，“I think... uh... the answer is... maybe... two”。

无意义重复表达。学生回答时说“I want... I want... to go to... um... to the mall”，重复不必要的词汇。在描述经历时，反复使用某个词：“We went... we went... to the restaurant... and then we went... to the cinema”。句子衔接差。学生表达观点时说“The weather is nice. I went for a walk. The flowers are pretty”，没有使用任何衔接词。在描述活动时，没有连贯性：“We met at the café. We talked. We left”。词汇匮乏。描述物体时，学生说“It’s big... um... it’s like big”，缺乏更多的描述性词汇。学生在讨论时说“The movie was... uh... it was good”，未能进一步具体描述电影。语速过慢。学生试图描述复杂的概念：“The... uh... weather... is... nice”，每个词之间的间隔很大。在讲述故事时，学生缓慢说出每个句子：“We... walked... to the park... and... then... we played”。

#### 4.口语表达的连贯性和完整性

在第三次观察中，对照组学生的表达依然存在连贯性和完整性问题。尽管他们尝试连接不同的句子，但常常由于停顿而使表达显得不自然，思路时常中断。

例 1：学生说“I enjoy swimming... and... um... running”。在这里，学生尝试使用“and”来连接两个兴趣，但由于停顿影响了流畅性，导致听者需要花时间理解句子的意思。

例 2：学生说“I go to the library... then... I read books”。虽然表达了两个相关的活动，但由于多次停顿，整体表达缺乏连贯性，无法有效传达思路。

### 第四次观察记录（实验组）22 Mar 2023

#### 1.学生在课堂上的参与度和表现

描述：多数学生在课堂上表现得较为专注，能够认真听讲并参与课堂讨论，课堂上问答互动频繁，讨论氛围较为热烈。

存在的问题：少部分学生似乎对课程内容兴趣不大，他们在课堂上只做表面上的参与，并未积极思考问题。有些学生会在课堂中做与学习无关的事情，比如和同学私下聊天、玩手机，甚至有个别学生在桌下吃零食。此外，有学生上课时因为没有准备好教材或学习资料，导致无法有效参与课程。

#### 2.口语表达的准确性

发音的准确性：学生在发音的准确性方面还存在很多不足。和前一阶段相比，学生在单元音和双元音上的发音得到了改善，错误在逐步减少，但部分学生在摩擦音/v/ /θ//ð//ʒ/发音上仍然出错。有些学生不能区分 /m/ /n//ŋ/ 的发音。有部分学生在词重音方面发生错误的频率较大。有学生在句子的语调上出错。

词汇运用的准确性：学生在英语口语表达中的词汇重复使用的问题仍然很明显，很多学生在表达中受到母语影响，使用大量的中式英语来表达，这给语言的理解带来了一定的困扰（如学生常使用

Open/Close the light

-Turn on/Turn off the light; I very like it-I like it very much.You are too great-You are amazing or You are fantastic.

I suggest you to do this-I suggest you do this;Let’s go out to play -Let’s go out and have some fun or Let’s hang out.I’m so boring-I’m so bored 或 This is boring.)

语法正确性：学生在口语表达时，经常在时态、语态、主谓一致、介词等方面出现表述错误。有部分学生刻意地使用复杂句来取代简单句但最后出错了。

#### 3.口语表达的流利性

停顿还是很频繁。学生在试图描述计划时，“Next weekend I... um... plan to... uh... visit my... uh... grandmother”。学生描述一天的活动时，“I... went to... um... school... then... uh... after that... I... came home”，不断地停顿使得句子断裂。

学生在表述中经常重复某个词多遍。回答问题时，重复某个短语：“I think... I think... the answer is correct”。描述某件事时，学生说“I went to the park... and... and I played football... and... and it was fun”，重复“and”。句子衔接较差。学生在回答问题时缺少细节衔接：“I like reading books. It’s good. I do it often”。在描述一次旅行时，学生说“I went to the beach. I ate ice cream. It was fun”，句子之间没有过渡词汇。

#### 4.口语表达的连贯性和完整性

在第四次观察中，学生的连贯性和完整性问题依然明显，特别是在较长的句子或复杂表达中，句子往

往断裂或缺乏逻辑联系。

例 1：学生说“I went to the park. I play... um... football”，动词时态错误加上停顿，使句子既不连贯也不完整。

例 2：描述一个活动时，学生说“We went to the zoo... saw animals... and it was... fun”，句子缺少必要的衔接词，整体表达不完整。

#### **第四次观察记录（对照组） 23 Mar 2023**

##### **1.学生在课堂上的参与度和表现**

描述：课堂整体氛围较好，学生们大多数都能集中注意力，积极投入到课堂讨论中，与老师和同学的互动频繁。

存在的问题：少数学生在课中教师组织的小组活动和讨论环节中没有按照教师要求分组并参与到活动或者讨论中，而是呆在原位不动。有的学生在课堂上私自使用耳机听音乐，还有些学生利用笔记本电脑或平板电脑进行与课程无关的操作，如观看视频或购物。个别学生上课迟到并且没有道歉，影响了课堂秩序。个别学生发现有外来的听课人员之后也不会改变这些行为。

##### **2.口语表达的准确性**

发音的准确性：学生在部分单元音、部分双元音和摩擦音、舌边音、舌侧近音的发音上仍然容易出错。在鼻音 /m/ /n/ /ŋ/ 的发音时经常混淆。在词重音、句子的语调上经常出错。

词汇运用的准确性：学生在英语口语表达中受母语影响很大，经常使用中国式英语或者生硬地直译来表达。

语法正确性：学生在口语表达时，在时态、语态、主谓一致、介词等方面出现表述错误仍比较常见。

##### **3.口语表达的流利性**

表达中频繁停顿。学生回答问题时说“I think... um... it's... uh... maybe... right”，中途多次停顿使得表达断断续续。试图讲述过去的经历：“Last year... I... uh... traveled to... uh... Paris”，因缺乏自信或词汇导致频繁停顿。

无意义重复表达。学生描述一件事时说“I think... I think... I need to... go... I think... to the library”，重复表达“我觉得”。

在表达某个想法时，学生不断重复“like”：“It was... like... like... really good... like... the best movie”。

句子衔接差。在描述早晨活动时，学生说“I wake up. I brush teeth. I eat breakfast”，没有任何连接词，显得内容断裂。学生表达计划时说“I will study. I will play games. It is nice”，缺乏衔接词使句子显得生硬。

##### **4.口语表达的连贯性和完整性**

在第四次观察中，学生的连贯性和完整性问题依然明显，尤其是在较长的句子中，表达往往断裂，缺乏足够的信息，听者的理解受到限制。

例 1：学生说“I visited my aunt. She lives in the city”。虽然每句话都是完整的，但缺乏进一步的细节描述，如关于她的生活或与她的活动，这使得内容显得薄弱，无法引发听者的兴趣。

例 2：学生描述活动时说“We played games. It was fun”。这句话虽然完整，但缺乏更多信息，如游戏的种类或玩的过程，整体表达显得简单且无趣。

#### **第五次观察记录（实验组） 12 Apr 2023**

##### **1.学生在课堂上的参与度和表现**

描述：大多数学生在课堂上展现了较强的学习意愿，他们积极参与讨论，能够跟随教师的思路，学习状态较为积极。整体上，课堂的互动效果较好。

存在的问题：少部分学生上课期间显得心不在焉，甚至有时候直接无视老师的提问，课堂参与度很低。有的学生上课期间不做笔记，也不看课本，只是漫不经心地坐着，还有些学生则一直低头玩手机或与同学在纸条上来回传递信息。此外，有学生因准备不足，在课堂上找借口回避参与讨论。

##### **2.口语表达的准确性**

发音的准确性：学生在发音的准确性方面得到了较大的改善，学生在词重音、句子语调方面的错误的

有了明显改善。少部分学生摩擦音、鼻音等方面还存在问题，仍有学生不能正确地发/l/和/v/的发音。

词汇运用的准确性：学生在词汇运用的准确性方面有了一些提升，但并不明显。首先，学生逐渐学会了根据不同的语境来选用不同的词汇来表达，词汇的多样性得到了增强，这是非常棒的。但仍有不少学生在表述中频繁使用中国式英语，或者经常出现词义错误或者搭配错误等现象。

语法正确性：学生在语法方面的错误逐渐减少，他们能克服一些基本的错误，比如在前一阶段在时态、语态上的错误很大程度上得到了消除。学生在简单句表达上准确率较高，但是学生在复杂句的表达上仍存在较大问题。

### 3.口语表达的流利性

学生在表达中的停顿减少了。学生在描述假期活动时，虽然有少量停顿，但整体句子流畅：“During the summer, I went to the beach and had a great time”，只在稍难的单词处有短暂停顿。

学生在表达中的重复现象减少。学生在讨论一个问题时，虽然有少量重复，但能快速纠正：“I think... I mean, I believe that the answer is correct”，不再像之前一样多次重复同样的词汇。

学生表达时，句子衔接性得到改进。学生在描述一天的行程时能正确使用衔接词：“First, we went to the museum, and then we had lunch at a café”，句子之间逻辑顺畅。

语速比以前快了一些。学生在表达想法时，语速适中且流畅：“I think the best part of the trip was visiting the historical landmarks”，没有明显的拖延或过慢的表达。

通过这些观察可以看出，经过一段时间的训练，学生在英语口语表达流利性方面有了提高，停顿减少、重复减少、句子衔接变得更自然、语法更加准确、词汇量增加，语速也更加自然流畅。这些进步表明学生在训练中的努力和效果。

### 4.口语表达的连贯性和完整性

在第五次观察中，学生在连贯性和完整性方面有了明显的改善。能够使用简单的衔接词，句子之间的逻辑关系逐渐清晰，表达更流畅。

例 1：学生说“I went to the park with my friends, and we played football for two hours”，正确使用了连接词“and”，句子之间的连贯性有所提高。

例 2：学生描述一件事时说“I like reading because it helps me relax”，使用“because”使句子连贯，表达更加完整。

## 第五次观察记录（对照组） 13 Apr 2023

### 1.学生在课堂上的参与度和表现

描述：多数学生在课堂上表现出了较高的参与热情，能够积极配合教师的讲解与任务安排，课堂互动良好，整体气氛较为融洽

存在的问题：部分学生的课堂表现仍然存在问题，他们不主动参与互动，对课程内容的反应迟钝。有些学生会在上课时使用手机拍摄老师讲义或做其他事情，完全忽略了课堂内容。还有的学生上课时在教室角落窃窃私语，甚至传纸条，分散了自己和其他同学的注意力。另外，有学生上课期间无故迟到或早退，影响了课堂秩序。

### 2.口语表达的准确性

发音的准确性：学生在发音的准确性方面未得到明显的改善，学生在前阶段容易出现发音错误的元音和辅音错误仍然频繁出现，特别是三个鼻音，舌边音/l/和、舌侧近音/v/的发音错误明显。很多学生对于多音节单词的重音搞不清楚，对长单词的发音经常出错。

词汇运用的准确性：学生在词汇运用的准确性方面有了一些改善。学生在表单中的词汇多样性使用较前一阶段有了增强。但中国式英语仍然非常明显，词汇搭配和词义错误情况经常发生。

语法正确性：学生在语法方面的错误逐渐减少，一些基本的语法错误不再发生。学生在简单句表达方面越来越准确，但是在复杂句的表达上仍存在较大问题。

### 3.口语表达的流利性

频繁停顿方面没有改进。在描述假期活动时，学生说“During... um... the summer... I... uh... went to... um... the beach”，频繁停顿使表达断断续续。学生说“I... um... I want to... uh... go to the store... but I... um... forgot”，停顿过多，使表达显得混乱。试图描述日常活动：“Every morning... I... uh... wake up... and... um...

eat breakfast”，频繁的停顿使整个表达显得非常不流畅。

无意义重复仍然明显。描述一个人时，学生说“He is... he is... very smart... and... and... he is good at math”，重复词汇和短语，显得不流畅。表达观点时，学生不断重复“think”：“I think... I think... it's a good idea... I think”。在描述事件时，学生说“We went to... went to... the park... and... and we played football”，重复动词使句子显得冗长。

句子衔接仍然较差。学生说“I watched a movie. It was good. My friend liked it”，句子之间缺少衔接词，显得不连贯。在描述一天的活动时，学生说“I went to school. I did my homework. I watched TV”，缺少衔接词，信息显得割裂。描述一个经历时，学生说“We went to the restaurant. I ate a pizza. The pizza was delicious”，没有使用适当的连接词，句子显得分散。

语法错误频繁。学生说“He don't like pizza”，动词形式错误，影响了表达的准确性。学生说“I meeted him yesterday in the store”，错误地使用了“go”的过去时，影响了语法的准确性。学生在描述活动时说“She playing basketball yesterday”，缺少动词“was”，导致语法错误。

#### 4. 口语表达的连贯性和完整性

在第五次观察中，对照组学生的口语表达开始出现一些小的改进，尽管整体连贯性和完整性仍然不足，但他们在句子结构和信息表达上稍微有所提高。

例 1：学生说“I went to the store and bought some snacks”。尽管这个句子相对简单，且没有细节描述，但使用了“and”连接了两个相关的信息，使得表达稍显连贯。

例 2：在描述兴趣时，学生说“I like to cook. Cooking is fun”。学生尝试使用不同的句子来表达相同的想法，虽然重复了“I like to cook”，但第二句增加了对“烹饪”这一兴趣的补充说明，体现出思维的连贯性。

### 第六次观察记录（实验组） 10 May 2023

#### 1. 学生在课堂上的参与度和表现

描述：课堂上大部分学生表现出积极的学习状态，能专心听讲并认真做笔记，积极回应教师提出的问题，整体表现较为理想。

存在的问题：少部分学生的参与度较低，他们在课堂上不太活跃，对教师的问题没有回应。有些学生上课时随意翻阅课外书籍或在教室后排与同学私语，不仅影响了自己，也分散了周围同学的注意力。还有学生未经允许擅自使用电子设备，打断了课堂的正常进行。

#### 2. 口语表达的准确性

发音的准确性：学生的发音准确性得到了明显的改善和提高。相比较前阶段，学生英语口语表达中的发音方面的错误大部分得到了消除，除少部分学生在少量辅音（如 /θ//ð/）和个别近音(Approximants)上偶尔出错，大部分学生的发音都是准确的。学生在词重音、句子的语调上方面几乎不会再出错。

词汇运用的准确性：学生在词汇准确性方面得到了明显的改善。很多学生从一开始只能使用大纲中的简单词汇来呈现单一和重复的表述，到现在能根据语境灵活使用近义词或者替代性表述来表达。学生在表述时词汇搭配不当、词义错误大大降低了。尽管仍然有部分学生偶尔会使用中国式英语，但他们克服中国式英语的意识得到了增强。

语法正确性：学生在语法方面的错误逐渐减少，学生很少在时态、语态等基本语法上犯错误。学生在简单句表达上准确率较高，在复杂句的表达上虽然还存在较大的困难和问题。

#### 3. 口语表达的流利性

口语表达频繁停顿减少。在表达一个想法时，学生更自信：“Last weekend, I visited my grandmother and we talked for hours”，基本没有停顿，语速自然。

重复现象减少。在描述活动时，学生说：“We went to the park, played football, and had ice cream afterward”，句子流畅，没有不必要的重复。

衔接性得到改进。在表达意见时，学生使用了合适的连接词：“I enjoy reading books because they help me relax after a long day”，句子连贯，过渡自然。

语法错误减少。在讨论某个话题时，学生说：“She doesn't like spicy food, so we ordered something mild”，正确使用了否定形式，语法大幅改进。

语速加快了。在描述日常活动时，学生语速自然：“Every morning, I wake up, have breakfast, and then go

for a jog”，表达流畅，语速恰当。

#### 4.口语表达的连贯性和完整性

在第六次观察中，学生的连贯性和完整性有了显著的改善。学生能清楚地表达完整的想法，句子之间的逻辑关系明显增强，表达较前几次观察更加自然。

例 1：学生说“Yesterday, I went to the mall, bought some clothes, and then had dinner with my family”，句子流畅，信息完整，且逻辑连贯。

例 2：在讨论爱好时，学生说“I enjoy playing basketball because it's fun and keeps me healthy”，句子衔接自然，表达完整且具有逻辑性。

### 第六次观察记录（对照组） 11 May 2023

#### 1.学生在课堂上的参与度和表现

描述：大部分学生课堂表现较为积极，能够跟上教师的教学节奏，认真参与到课堂互动中，整体表现不错，教学效果较好。

存在的问题：少部分学生的注意力并不集中，他们表现出对课堂活动的参与热情不足，有些学生上课时目光游离，经常看向窗外或低头看手机，表现出明显的疲惫感。有学生打着哈欠，显得缺乏精神投入。有学生在教师提出问题神情茫然，领会不到教师提出的问题。有些学生上课时无故离开座位，有的学生借口参加小组活动而在教室随意走动，有几个学生在下课之前提前离开了课堂。

#### 2.口语表达的准确性

发音的准确性：学生的发音准确性未得到明显改善和提高。学生在发音上的错误率仍然较高，甚至部分学生在一些简单的元音和辅音的发音上仍然经常犯错误。

词汇运用的准确性：学生在词汇准确性方面得到一定的改善，但成效并不明显。中国式英语仍然非常明显，词汇搭配和词义错误情况仍然经常发生。

语法正确性：学生在语法方面的错误逐渐减少，在简单句表述和时态、语态等基本语法上的错误很少发生，但他们在复杂句的表达上问题仍然很明显。

#### 3.口语表达的流利性

频繁停顿有所改进。学生在日常表达中停顿减少，尤其是在简单的句子中，流利性有所提升。学生说“Yesterday, I went to the park and played football with my friends”，几乎没有停顿，表达比较顺畅。描述早晨活动时，学生说“I wake up, have breakfast, and then go to school”，表达自然，停顿减少。

学生在简单句中重复现象减少，能表达自己的想法。学生说“I think this is the right answer”，没有像以前那样重复“I think”。在描述经历时，学生说“We went to the museum and learned a lot”，没有出现不必要的重复。

在表达简单想法时，语速有所提升，整体流畅性较好。学生说“Last weekend, I visited my grandparents and had a great time”，语速均匀且自然。在描述假期计划时，学生说“I plan to go hiking with my family next month”，语速适中，表达流利。

表达的衔接性仍未得到改进。学生在复杂句子中的句子衔接仍显不足，句子之间的逻辑关系不够清晰。学生说“I like reading books. I also like football. It is good”，缺乏衔接词，句子之间的关系不明确。在描述一个活动时，学生说“We went to the park. Played football. It was fun”，句子之间没有明显的逻辑衔接。语法错误，尤其是时态和主谓一致问题，仍然较为常见。学生说“She don't like the food”，在动词形式上仍然出现错误。描述过去的活动时，学生说“Yesterday, I go to the store”，时态错误没有明显改善。

#### 4.口语表达的连贯性和完整性

在第六次观察中，对照组学生的口语表达依然存在连贯性和完整性的问题，虽然有轻微的进步，但整体表达仍然显得不够清晰。

例 1：学生说“Yesterday, I went to the park, and we had a picnic”。虽然这句话使用了连接词“and”，使得信息更连贯，但仍缺乏具体细节，比如食物的种类或活动的时间，整体内容仍显得简单。

例 2：描述天气时，学生说“It was sunny, and I enjoyed it”。虽然句子之间使用了连接词“and”，表达了对阳光天气的喜爱，但缺乏更多的描述和具体感受，使得整体表达仍显得有些连贯和不完整。

## **Translation of Classroom Observation Records**

### **First Observation Record (Experimental Group) - 28 December 2022**

#### **1. Student Participation and Performance in Class**

Overall Description: The majority of students were in a good state during the class. They actively engaged in the tasks and interacted with the teacher.

Existing Problems: Approximately 5 - 6 students were not very active or proactive. They did not participate in class activities, did not make eye contact with the teacher, and did not interact with the teacher. There were 2 - 3 students who were absent - minded. They often lowered their heads, remained silent, looked out of the window, or played with their mobile phones.

#### **2. Accuracy of Oral Expression**

Pronunciation Accuracy: Most students pronounced most words correctly in their English oral expressions. However, some students made mistakes in the pronunciation of certain vowels. For example, /i:/ was mispronounced as the short vowel /ɪ/. Some students also had errors in the pronunciation of diphthongs such as /ɔɪ/, /aʊ/, and /əʊ/. Additionally, some students made mistakes in the pronunciation of fricative sounds (such as /f/ and /v/).

Vocabulary Usage Accuracy: Most students could use the vocabulary in the syllabus to express themselves in spoken English. However, many students had a limited vocabulary and had the problem of over - using familiar words. They could only use relatively simple words in their spoken English expressions and repeatedly used certain familiar words, resulting in monotonous and repetitive expressions. For example, students frequently used "good" to express various positive meanings, while ignoring more appropriate alternatives such as "excellent" and "fantastic".

Grammar Correctness: When students expressed themselves orally, they often made many mistakes in aspects such as tenses, voices, subject - verb agreement, and prepositions.

#### **3. Fluency of Oral Expression**

Stuttering: When students expressed themselves, there were long pauses between sentences. For example, when saying "I went to the store...", they would pause for several seconds and then continue with "and... I... bought... some... milk", making the expression unsmooth.

Hesitating: Students often paused in the middle of their speech, repeated a word, or disrupted the fluency of the sentence because they couldn't remember a word or a grammar rule. For example, a student said "I want... I want... to go... to... to... the park".

Poor Sentence Cohesion: When students spoke different sentences, they were unable to use appropriate cohesive words, making the sentences seem rigid. For example, "I like pizza. It is good. I eat it every day". There were no reasonable cohesive or correlative words between the sentences, and the expression seemed stiff.

Long Pauses: When expressing complex sentences or ideas, students often had long pauses to think about what to say next, causing difficulties for the listener to understand. For example: "Yesterday I... uh... I... went to the cinema with my... uh... with my friend".

Overusing Simple Conjunctions: When expressing different ideas, students frequently used "and" to connect sentences or phrases, lacking diversity. For example: "I went to the park and I played football and I met my friend and we ate lunch".

#### **4. Coherence and Completeness of Oral Expression**

Students' oral expressions were clearly lacking in coherence and completeness. There were no reasonable connections between sentences, and the expressed content seemed fragmented, making it difficult for the listener to understand the students' intentions.

Example 1: A student said "I like pizza. I eat pizza. My friend likes pizza too". There were no connecting words or logical links between these sentences, and the content seemed stiff.

Example 2: The student's expression was incomplete: "Yesterday, I... um... went to... um... the store". Due to frequent pauses, the sentence was not coherent and the information was incomplete.

### **First Observation Record (Control Group) - 29 December 2022**

#### **1. Student Participation and Performance in Class**

Overall Description: Most students performed well in class. They could concentrate and actively participated in the tasks and interacted with the teacher.

Existing Problems: Some students needed to be called on by the teacher to answer questions or participate in class discussions. Individually, a few students still refused to answer the teacher's questions even after being called. There were 7 - 8 students who were playing with their mobile phones or doing other things unrelated to the class (such as doing homework for other subjects, reading storybooks, or sleeping).

#### **2. Accuracy of Oral Expression**

Pronunciation Accuracy: Most students pronounced most words correctly in their English oral expressions. However, some students made mistakes in the pronunciation of certain vowels. For example, /i:/ was mispronounced as the short vowel /ɪ/. Some students also had errors in the pronunciation of diphthongs such as /ɔɪ/, /aʊ/, and /əʊ/. Additionally, some students made mistakes in the pronunciation of fricative sounds (such as /f/ and /v/).

Vocabulary Usage Accuracy: Many students had a limited vocabulary and had the problem of over - using familiar words. They could only use relatively simple words in their spoken English expressions and repeatedly used certain familiar words, resulting in monotonous and repetitive expressions. Some students also had problems with expressing the wrong meaning or using the wrong words in their expressions (for example, confusing "fun" and "funny", "lend" and "borrow").

Grammar Correctness: When students expressed themselves orally, they often made mistakes in aspects such as tenses, voices, subject - verb agreement, and prepositions.

#### **3. Fluency of Oral Expression**

Frequent Pauses and Unclear Thoughts: A student said "I... I went to... um... the market... uh... to buy... uh... food". There were many pauses during the expression, and it was obvious that the student was looking for the right words or sentence structures, resulting in an incoherent train of thought.

Stuttering and Incomplete Expressions: A student wanted to express an experience, but the words were interrupted: "Last weekend... I with my friend... we... go to... but... didn't go". The listener could not understand the complete meaning, and the expressed content was seriously lacking.

Lack of Sentence Cohesion: A student said "I like football. It is a sport. I play on the weekend. My team is good". There were no appropriate cohesive words between the sentences, and the relationship between the information was not clear, making it seem incoherent.

Lack of Logic: When a student described an activity, the order and logic were confused: "First we eat lunch. After that we... go to the cinema. I buy a ticket... then... we walk... and eat lunch". The described order was inconsistent, making it difficult to understand.

Poor Coherence: When answering questions, students frequently used incoherent phrases, such as "I want... I think... uh... we... should... because... but... uh... I don't know", without forming complete sentences. The expression was not smooth and lacked logical coherence.

#### **4. Coherence and Completeness of Oral Expression**

In the first observation, the oral expressions of the students in the control group lacked coherence and completeness. There were almost no connections between sentences during the expression process, and the content seemed fragmented. It was difficult for the listener to capture the complete information.

Example 1: A student said "I like chocolate. Chocolate is sweet". Although the student expressed their preference, there were no connecting words. The content was repetitive and too simple, and no specific views or more information were provided.

Example 2: The student's expression was incomplete: "Yesterday, I went to... um... the school... for...". Due to frequent pauses, the sentence seemed incoherent, and the listener could not understand the specific content the student wanted to express.

## **Second Observation Record (Experimental Group) - 8 February 2023**

### **1. Student Participation and Performance in Class**

Overall Description: Overall, the students performed well in class. Most students could focus on the class and actively participated in the tasks. Most students showed more confidence and interacted frequently with the teacher.

Existing Problems: A small number of students (about 5) were not very active. They did not actively answer the teacher's questions or participate in class activities. Four students were watching movies or playing games on their mobile phones during the class.

### **2. Accuracy of Oral Expression**

Pronunciation Accuracy: Some students made mistakes in the pronunciation of certain vowels (for example, when pronouncing /æ/ and /ɑ:/, the mouth opening was not large enough). Some students mispronounced the "ear" in "learn" as /ə/ instead of /ɜ:/ . Some students had obvious errors in the pronunciation of diphthongs. Some students also made mistakes in the pronunciation of fricative sounds and lateral sounds.

Vocabulary Usage Accuracy: Many students preferred to use familiar and simple words to express themselves, resulting in monotonous and repetitive expressions. Some students had problems with expressing the wrong meaning or using the wrong words in their expressions (for example, confusing "fun" and "funny", "lend" and "borrow"). Some students also had inappropriate vocabulary collocations (such as "do a mistake - make a mistake", "big rain - heavy rain", "do a meal - make a meal", "pay attention on - pay attention to").

Grammar Correctness: When students expressed themselves orally, mistakes in aspects such as tenses, voices, subject - verb agreement, and prepositions were still very obvious.

### **3. Fluency of Oral Expression**

Lack of Connecting Words between Sequences: A student said "I go to school. I study math. I eat lunch. I come home". There were no connecting words between the sentences, and the logical relationship or time sequence between the sentences was not expressed.

Unclear Logic: When expressing an opinion, a student said "I think... um... school is important... because... uh... we learn things... but... sometimes it's boring... um... I don't know". The expressed content was inconsistent and the logic was confused.

Tense Errors and Confused Expressions: A student said "Yesterday I go to the store and buy milk. Then I go home and I drink the milk". The tense errors made the expressed content difficult to understand and lack logic.

Lack of Connecting Words Leading to Information Jumps: When describing a vacation experience, a student said "I went to the beach. The sun was bright. We swim. The food was delicious". There were no connecting words for each sentence, and the content seemed to jump without coherence.

Repetition Due to Lack of Vocabulary: When a student wanted to describe a day's activities, they said "I... I do homework. Then I do... um... I do more homework. Then... I do other things... like... homework". The lack of vocabulary led to repetition, and the content seemed monotonous and chaotic.

### **4. Coherence and Completeness of Oral Expression**

The second observation showed that the students' expressions still lacked coherence and completeness. Although they could say some independent sentences, there was still a lack of logical relationships or appropriate connections between the sentences.

Example 1: A student said "I have a dog. My dog is big. We play outside". Although each sentence was an independent and complete sentence, there was a lack of logical connection, and there was no coherence between the sentences.

Example 2: A student said "Last weekend, I went to my... um... friend's house". The sentence structure was incomplete, and the listener could not get enough information.

## **Second Observation Record (Control Group) - 9 February 2023**

### **1. Student Participation and Performance in Class**

Description: Most students performed well in class. They could concentrate and actively participated in the tasks and interacted with the teacher.

Existing Problems: Some students needed to be called on by the teacher to answer questions or participate in class discussions. Individually, a few students still refused to answer the teacher's questions even after being called. There were 7 - 8 students who were playing with their mobile phones or doing other things unrelated to the class (such as doing homework for other subjects, reading storybooks, or sleeping).

### **2. Accuracy of Oral Expression**

Pronunciation Accuracy: Students made mistakes in the pronunciation of some monophthongs, especially often confusing long and short vowels. Some students had obvious errors in the pronunciation of diphthongs, and the diphthongs were not pronounced fully. Students often made mistakes in the pronunciation of fricative sounds and lateral sounds.

Vocabulary Usage Accuracy: Students liked to repeatedly use familiar and simple words to express themselves, resulting in monotonous and repetitive expressions. Problems such as expressing the wrong meaning, using the wrong words, and inappropriate vocabulary collocations often occurred in their expressions.

Grammar Correctness: When students expressed themselves orally, mistakes in aspects such as tenses, voices, subject - verb agreement, and prepositions were still very obvious.

### **3. Fluency of Oral Expression**

Disordered Repetition: A student said "Yesterday I... yesterday I went to... I went to the park... um... yesterday". The same word or phrase was repeated many times, and the expression was chaotic and seemed illogical.

Incoherent Expression: When a student wanted to describe weekend activities, they said "On Saturday... I... with my friend... um... we... we go... park... and... um... football". The sentence lacked necessary components and could not form a coherent expression.

Confused Sentence Order: When a student described their daily routine, they said "I wake up... I go to school... after school I eat breakfast... then I sleep". The activity order was reversed, lacking logic.

Relying on Simple Words and Repetition: A student said "I like movies. And... um... I like movies... and... um... I watch movies... and... um... movies are good". There was excessive reliance on "and" and repetitive expressions, and the logic was not clear.

Long Pauses: When a student was trying to explain something, they said "I... uh... last week I... um... go to... uh... library". There were long pauses between each sentence, making it difficult for the listener to understand.

### **4. Coherence and Completeness of Oral Expression**

In the second observation, the students' oral expressions still lacked coherence and completeness. Although they tried to use simple sentences, the logical relationship between the sentences was still weak, and the overall expression seemed unsmooth.

Example 1: A student said "I have a cat. The cat is black". Although these two sentences were independent and complete sentences, there was a lack of further description or connection. The overall expression seemed monotonous and could not provide more background information to the listener.

Example 2: A student said "Last weekend, I went... um... out with my... um... friends". In this sentence, although an activity with friends was mentioned, due to the pauses, the information was incomplete, and the listener could not understand the specific activity content.

### **Third Observation Record (Experimental Group) - 2 March 2023**

#### **1. Student Participation and Performance in Class**

Description: In the class, most students showed a high sense of participation. They actively completed the class tasks, could abide by the class discipline well, listened attentively, showed a strong interest in the course content, and the class interaction was relatively active.

Existing Problems: A small number of students seemed a bit passive in class. They were reluctant to have in - depth discussions with the teacher or classmates. Some students would look at their mobile phones from time to time or play games under the desk. Even some students dozed off due to fatigue, affecting their learning efficiency.

#### **2. Accuracy of Oral Expression**

Pronunciation Accuracy: Some students made mistakes in the pronunciation of certain vowels and some consonants (such as /θ/, /ð/, /ʒ/, /l/, /t/). Some students did not close their mouths when pronouncing /m/. Students had great difficulties in pronouncing long words, and they could not grasp the word stress accurately. Some students also made mistakes in the intonation of sentences.

Vocabulary Usage Accuracy: Students repeatedly used the basic vocabulary in the syllabus to express themselves in different contexts, lacking flexibility and diversity in vocabulary usage. Some students had problems with expressing the wrong meaning, and many students frequently used Chinglish in their expressions.

Grammar Correctness: When students expressed themselves orally, they often made mistakes in aspects such as tenses, voices, subject - verb agreement, prepositions, and articles. Some students deliberately used complex sentences instead of simple sentences but made mistakes in the end.

#### **3. Fluency of Oral Expression**

Frequent Pauses: A student said "I... um... went to... the park... and... uh... played football". There were obvious pauses between each sentence and phrase. When expressing "Last week... I... uh... studied for my test", the student paused frequently because of unclear thoughts.

Frequent Repetition: A student said "I like... I like... I like playing basketball" and could not express a complete sentence smoothly.

Repeating Words When Expressing Opinions: "The movie is... is... very interesting. It's... it's fun".

Poor Sentence Cohesion: A student said "I like soccer. It's fun. I play every weekend". There was a lack of coherent logic or connecting words between the sentences. When describing an event, there were no connecting words: "I went to school. I did homework. I watched TV".

Slow Speech: A student slowly said "I... went... to... the store... and... bought... some... food". The speech speed was very slow, and it was difficult for the listener to maintain attention. When a student expressed their opinion, the speed was extremely slow: "I... uh... think... the answer... is... right".

#### **4. Coherence and Completeness of Oral Expression**

In the third observation, the students' expressions showed a slight improvement, but the problems of coherence and completeness still existed. Students tried to use connecting words but often used them inappropriately or paused too much.

Example 1: A student said "I like reading... and... um... football. It is fun". The student tried to use "and" for connection, but the pause interrupted the coherence of the expression.

Example 2: A student said "I go to school... then... um... I study math". Although the sentence structure was complete, the unnecessary pause affected the coherence of the expression.

### **Third Observation Record (Control Group) - 2 March 2023**

#### **1. Student Participation and Performance in Class**

Description: In the class, most students actively participated in the teaching activities. They could concentrate and effectively follow the teacher's teaching rhythm, and the classroom atmosphere was quite lively.

Existing Problems: Some students seemed to lack the motivation to participate. They not only did not actively answer questions but also remained silent during group discussions. Some students frequently checked their mobile phones during the class and even secretly played games. There were also some students who were often distracted in class, looking out of the window absent-mindedly and completely disengaged from the class. In addition, some students would occasionally leave the classroom on the pretext of going to the toilet and not return for a long time, disrupting the class rhythm.

## **2. Accuracy of Oral Expression**

Pronunciation Accuracy: Students frequently made mistakes in the pronunciation of some vowels and consonants. They often got the word stress wrong and also made errors in the intonation of sentences.

Vocabulary Usage Accuracy: Students used simple words in a single-minded and repetitive way to express different contexts, lacking flexibility and diversity in vocabulary usage. They often made mistakes in vocabulary collocations, and many students frequently used Chinglish in their expressions.

Grammar Correctness: When students expressed themselves orally, they often made mistakes in aspects such as tenses, voices, subject-verb agreement, prepositions, and articles.

## **3. Fluency of Oral Expression**

Frequent Pauses: When expressing complex sentences, a student said, "I want... to go to... uh... the beach... because... uh... it's sunny".

When answering questions, students paused, for example, "I think... uh... the answer is... maybe... two".

Meaningless Repetition: When answering, a student said, "I want... I want... to go to... um... to the mall", repeating unnecessary words. When describing an experience, they repeatedly used a certain word: "We went... we went... to the restaurant... and then we went... to the cinema".

Poor Sentence Connection: When expressing opinions, a student said, "The weather is nice. I went for a walk. The flowers are pretty", without using any connecting words. When describing an activity, there was no coherence: "We met at the café. We talked. We left".

Lack of Vocabulary: When describing an object, a student said, "It's big... um... it's like big", lacking more descriptive words. When discussing, a student said, "The movie was... uh... it was good", without further elaborating on the movie.

Slow Speech: When trying to describe a complex concept, a student said, "The... uh... weather... is... nice", with a large interval between each word. When telling a story, a student slowly uttered each sentence: "We... walked... to the park... and... then... we played".

## **4. Coherence and Completeness of Oral Expression**

In the third observation, the expressions of the students in the control group still had problems with coherence and completeness. Although they tried to connect different sentences, the pauses often made the expressions seem unnatural, and their train of thought was often interrupted.

Example 1: A student said, "I enjoy swimming... and... um... running". Here, the student tried to use "and" to connect two hobbies, but the pause affected the fluency, making it take time for the listener to understand the meaning of the sentence.

Example 2: A student said, "I go to the library... then... I read books". Although two related activities were expressed, due to multiple pauses, the overall expression lacked coherence and could not effectively convey the train of thought.

## **Fourth Observation Record (Experimental Group) - 22 March 2023**

### **1. Student Participation and Performance in Class**

Description: Most students were quite focused in class. They listened carefully and participated actively in class discussions. There were frequent Q&A interactions, and the discussion atmosphere was rather

enthusiastic.

Existing Problems: A small number of students seemed to have little interest in the course content. They only made a superficial show of participation in class and did not think actively. Some students did things unrelated to learning during class, such as chatting privately with classmates, playing with mobile phones, and there were even individual students eating snacks under the desk. In addition, some students could not participate effectively in the class because they did not prepare teaching materials or learning resources in advance.

## **2. Accuracy of Oral Expression**

Pronunciation Accuracy: Students still had many deficiencies in pronunciation accuracy. Compared with the previous stage, the pronunciation of monophthongs and diphthongs had improved, and the number of errors was gradually decreasing. However, some students still made mistakes in the pronunciation of fricative sounds such as /v/, /θ/, /ð/, and /ʒ/. Some students could not distinguish the pronunciation of /m/, /n/, and /ŋ/. Some students had a relatively high frequency of errors in word stress, and some students made mistakes in sentence intonation.

Vocabulary Usage Accuracy: The problem of repeated vocabulary use in students' spoken English expressions was still obvious. Many students were influenced by their mother tongue and used a lot of Chinglish, which caused some difficulties in language understanding. (For example, students often used "Open/Close the light" instead of "Turn on/Turn off the light"; "I very like it" instead of "I like it very much"; "You are too great" instead of "You are amazing" or "You are fantastic"; "I suggest you to do this" instead of "I suggest you do this"; "Let's go out to play" instead of "Let's go out and have some fun" or "Let's hang out"; "I'm so boring" instead of "I'm so bored" or "This is boring".)

Grammar Correctness: When students expressed themselves orally, they often made mistakes in aspects such as tenses, voices, subject - verb agreement, and prepositions. Some students deliberately used complex sentences instead of simple sentences but made mistakes in the end.

## **3. Fluency of Oral Expression**

Frequent Pauses: When trying to describe a plan, a student said, "Next weekend I... um... plan to... uh... visit my... uh... grandmother". When describing a day's activities, a student said, "I... went to... um... school... then... uh... after that... I... came home". The continuous pauses made the sentences disjointed.

Frequent Repetition of Words: When answering questions, students repeated a certain phrase, for example, "I think... I think... the answer is correct". When describing something, a student said, "I went to the park... and... and I played football... and... and it was fun", repeating "and".

Poor Sentence Cohesion: When answering questions, students lacked detailed cohesion. For example, "I like reading books. It's good. I do it often". When describing a trip, a student said, "I went to the beach. I ate ice cream. It was fun", without transitional words between sentences.

## **4. Coherence and Completeness of Oral Expression**

In the fourth observation, the problems of coherence and completeness in students' expressions were still obvious, especially in longer sentences or complex expressions. Sentences often broke off or lacked logical connections.

Example 1: A student said, "I went to the park. I play... um... football". The verb tense error combined with the pause made the sentence neither coherent nor complete.

Example 2: When describing an activity, a student said, "We went to the zoo... saw animals... and it was... fun". The sentence lacked necessary connecting words, and the overall expression was incomplete.

## **Fourth Observation Record (Control Group) - 23 March 2023**

### **1. Student Participation and Performance in Class**

Description: The overall classroom atmosphere was good. Most students could concentrate and actively engage in class discussions, having frequent interactions with teachers and classmates.

Existing Problems: A small number of students did not group and participate in the group activities and discussion sessions organized by the teacher as required. Instead, they stayed in their original positions. Some students privately used headphones to listen to music during class, and some students used laptops or tablets to do things unrelated to the course, such as watching videos or shopping. Individual students were late for class without apologizing, affecting the classroom order. Some students did not change their behavior even when they noticed external observers.

## **2. Accuracy of Oral Expression**

Pronunciation Accuracy: Students still made mistakes easily in the pronunciation of some monophthongs, diphthongs, fricative sounds, lateral sounds, and lateral approximants. They often confused the pronunciation of nasal sounds /m/, /n/, and /ŋ/, and frequently made mistakes in word stress and sentence intonation.

Vocabulary Usage Accuracy: Students were greatly influenced by their mother tongue in spoken English expressions and often used Chinglish or literal translations.

Grammar Correctness: When students expressed themselves orally, mistakes in tenses, voices, subject - verb agreement, prepositions, etc., were still common.

## **3. Fluency of Oral Expression**

Frequent Pauses: When answering questions, a student said, "I think... um... it's... uh... maybe... right", with multiple pauses making the expression disjointed. When trying to tell about a past experience, "Last year... I... uh... traveled to... uh... Paris", frequent pauses occurred due to lack of confidence or vocabulary.

Meaningless Repetition: When describing something, a student said, "I think... I think... I need to... go... I think... to the library", repeating "I think". When expressing an idea, a student kept repeating "like": "It was... like... like... really good... like... the best movie".

Poor Sentence Cohesion: When describing morning activities, a student said, "I wake up. I brush teeth. I eat breakfast", without any connecting words, making the content seem fragmented. When expressing a plan, a student said, "I will study. I will play games. It is nice", lacking connecting words and making the sentences seem rigid.

## **4. Coherence and Completeness of Oral Expression**

In the fourth observation, the problems of coherence and completeness in students' expressions were still obvious, especially in longer sentences. The expressions often broke off, lacked sufficient information, and limited the listener's understanding.

Example 1: A student said, "I visited my aunt. She lives in the city". Although each sentence was complete, there was a lack of further detailed descriptions, such as her life or activities with her, making the content seem weak and unable to arouse the listener's interest.

Example 2: When describing an activity, a student said, "We played games. It was fun". Although this sentence was complete, it lacked more information, such as the types of games or the process of playing, making the overall expression seem simple and uninteresting.

## **Fifth Observation Record (Experimental Group) - 12 April 2023**

### **1. Student Participation and Performance in Class**

Description: Most students showed a strong willingness to learn in class. They actively participated in discussions, could follow the teacher's train of thought, and were in a relatively positive learning state.

Overall, the classroom interaction was quite effective.

Existing Problems: A small number of students seemed absent - minded during the class. Sometimes they even directly ignored the teacher's questions, with a very low level of classroom participation. Some students didn't take notes or look at the textbooks during the class. They just sat there casually. Some students kept looking down at their mobile phones or passing notes back and forth with classmates. In addition, some students made excuses to avoid participating in discussions due to lack of preparation.

### **2. Accuracy of Oral Expression**

**Pronunciation Accuracy:** The students' pronunciation accuracy has improved significantly. There have been obvious improvements in word stress and sentence intonation errors. However, a small number of students still have problems with fricative sounds and nasal sounds. Some students still can't pronounce /l/ and /r/ correctly.

**Vocabulary Usage Accuracy:** The students' accuracy in vocabulary usage has improved to some extent, but the improvement is not obvious. Firstly, students have gradually learned to choose different words according to different contexts, enhancing the diversity of vocabulary, which is great. However, many students still frequently use Chinglish in their expressions, and there are often phenomena such as incorrect word meanings or collocations.

**Grammar Correctness:** The students' grammar mistakes have gradually decreased. They can overcome some basic errors. For example, the errors in tenses and voices in the previous stage have been largely eliminated. The accuracy of simple sentence expressions is relatively high, but students still have major problems in expressing complex sentences.

### **3. Fluency of Oral Expression**

**Fewer Pauses:** When students described holiday activities, although there were a few pauses, the overall sentence was smooth: "During the summer, I went to the beach and had a great time", with only short pauses at slightly difficult words.

**Less Repetition:** When discussing a question, although there was a little repetition, students could correct it quickly: "I think... I mean, I believe that the answer is correct", instead of repeating the same words multiple times as before.

**Improved Sentence Cohesion:** When describing a day's itinerary, students could use connectives correctly: "First, we went to the museum, and then we had lunch at a café", with a smooth logic between sentences.

**Faster Speech:** When expressing ideas, students spoke at a moderate and smooth pace: "I think the best part of the trip was visiting the historical landmarks", without obvious procrastination or slow expression.

Through these observations, it can be seen that after a period of training, the students have improved in the fluency of spoken English expression. There are fewer pauses, less repetition, more natural sentence cohesion, more accurate grammar, an increased vocabulary, and a more natural and smooth speaking speed. These improvements demonstrate the students' efforts and the effectiveness of the training.

### **4. Coherence and Completeness of Oral Expression**

In the fifth observation, the students' coherence and completeness have improved significantly. They can use simple connectives, and the logical relationships between sentences have gradually become clear, making the expressions more fluent.

**Example 1:** A student said, "I went to the park with my friends, and we played football for two hours". By correctly using the connective "and", the coherence between sentences has been improved.

**Example 2:** When describing something, a student said, "I like reading because it helps me relax". Using "because" makes the sentence coherent and the expression more complete.

## **Fifth Observation Record (Control Group) - 13 April 2023**

### **1. Student Participation and Performance in Class**

**Description:** Most students showed a high enthusiasm for participation in class. They could actively cooperate with the teacher's explanations and task arrangements. The classroom interaction was good, and the overall atmosphere was rather harmonious.

**Existing Problems:** The classroom performance of some students still had problems. They did not actively participate in interactions and were slow to respond to the course content. Some students used their mobile phones to take pictures of the teacher's handouts or did other things during the class, completely ignoring the class content. Some students whispered in the corners of the classroom or even passed notes, distracting

themselves and other classmates. In addition, some students were late or left early without reason during the class, affecting the classroom order.

## **2. Accuracy of Oral Expression**

**Pronunciation Accuracy:** There has been no obvious improvement in the students' pronunciation accuracy. The vowel and consonant errors that were prone to occur in the previous stage still appeared frequently, especially the pronunciation errors of the three nasal sounds, the lateral sound /l/, and the lateral approximant /ɹ/. Many students couldn't figure out the stress of polysyllabic words and often made mistakes in the pronunciation of long words.

**Vocabulary Usage Accuracy:** The students' accuracy in vocabulary usage has improved to some extent. The diversity of vocabulary used in their expressions has increased compared with the previous stage. However, Chinglish is still very obvious, and problems such as incorrect vocabulary collocations and word meanings often occur.

**Grammar Correctness:** The students' grammar mistakes have gradually decreased, and some basic grammar mistakes no longer occur. The students are becoming more and more accurate in expressing simple sentences, but they still have major problems in expressing complex sentences.

## **3. Fluency of Oral Expression**

**No Improvement in Frequent Pauses:** When describing holiday activities, a student said, "During... um... the summer... I... uh... went to... um... the beach". The frequent pauses made the expression disjointed. A student said, "I... um... I want to... uh... go to the store... but I... um... forgot". Too many pauses made the expression seem chaotic. When trying to describe daily activities: "Every morning... I... uh... wake up... and... um... eat breakfast", the frequent pauses made the whole expression very unsmooth.

**Obvious Meaningless Repetition:** When describing a person, a student said, "He is... he is... very smart... and... and... he is good at math", repeating words and phrases, making the expression unsmooth. When expressing an opinion, a student kept repeating "think": "I think... I think... it's a good idea... I think". When describing an event, a student said, "We went to... went to... the park... and... and we played football", repeating the verb and making the sentence seem long - winded.

**Poor Sentence Cohesion:** A student said, "I watched a movie. It was good. My friend liked it". There were no connectives between the sentences, making the expression incoherent. When describing a day's activities, a student said, "I went to school. I did my homework. I watched TV". Without connectives, the information seemed fragmented. When describing an experience, a student said, "We went to the restaurant. I ate a pizza. The pizza was delicious", without using appropriate connectives, making the sentences seem scattered.

**Frequent Grammar Errors:** A student said, "He don't like pizza", with an incorrect verb form that affected the accuracy of the expression. A student said, "I meeted him yesterday in the store", wrongly using the past tense of "go", affecting the grammatical accuracy. When describing an activity, a student said, "She playing basketball yesterday", lacking the verb "was" and resulting in a grammar error.

## **4. Coherence and Completeness of Oral Expression**

In the fifth observation, the oral expressions of the students in the control group began to show some small improvements. Although the overall coherence and completeness were still insufficient, there was a slight improvement in sentence structure and information expression.

**Example 1:** A student said, "I went to the store and bought some snacks". Although this sentence was relatively simple and without detailed description, using "and" to connect two related pieces of information made the expression slightly more coherent.

**Example 2:** When describing an interest, a student said, "I like to cook. Cooking is fun". The student tried to use different sentences to express the same idea. Although "I like to cook" was repeated, the second sentence added a supplementary explanation of the "cooking" interest, reflecting the coherence of thinking.

## **Sixth Observation Record (Experimental Group) - 10 May 2023**

### **1. Student Participation and Performance in Class**

Description: Most students showed a positive learning state in class. They listened attentively, took notes carefully, and actively responded to the teacher's questions. Overall, their performance was quite satisfactory.

Existing Problems: A small number of students had a low level of participation. They were not very active in class and did not respond to the teacher's questions. Some students casually flipped through extracurricular books during class or whispered with classmates at the back of the classroom, which not only affected themselves but also distracted the attention of their surrounding classmates. There were also students who used electronic devices without permission, disrupting the normal progress of the class.

### **2. Accuracy of Oral Expression**

Pronunciation Accuracy: The students' pronunciation accuracy has been significantly improved. Compared with the previous stage, most of the pronunciation errors in students' spoken English expressions have been eliminated. Except that a few students occasionally make mistakes in the pronunciation of a small number of consonants (such as /θ/, /ð/) and individual approximants, most students' pronunciation is accurate. Students rarely make mistakes in word stress and sentence intonation.

Vocabulary Usage Accuracy: The students' vocabulary accuracy has also improved remarkably. Many students, who could only use simple vocabulary from the syllabus to present single and repetitive expressions at the beginning, can now flexibly use synonyms or alternative expressions according to the context. The occurrences of inappropriate vocabulary collocations and wrong word meanings in students' expressions have been greatly reduced. Although some students still occasionally use Chinglish, their awareness of overcoming it has been enhanced.

Grammar Correctness: The students' grammar mistakes have gradually decreased. They rarely make mistakes in basic grammar such as tenses and voices. The accuracy of simple sentence expressions is relatively high, but they still face great difficulties and problems in expressing complex sentences.

### **3. Fluency of Oral Expression**

Fewer Frequent Pauses: When expressing an idea, students are more confident. For example, "Last weekend, I visited my grandmother and we talked for hours", with almost no pauses and a natural speaking speed.

Less Repetition: When describing an activity, a student said, "We went to the park, played football, and had ice cream afterward". The sentence is smooth and there is no unnecessary repetition.

Improved Cohesion: When expressing an opinion, a student used an appropriate connective: "I enjoy reading books because they help me relax after a long day". The sentence is coherent with a natural transition.

Fewer Grammar Errors: When discussing a topic, a student said, "She doesn't like spicy food, so we ordered something mild", correctly using the negative form and showing a significant improvement in grammar.

Faster Speech: When describing daily activities, a student spoke at a natural speed: "Every morning, I wake up, have breakfast, and then go for a jog", with a fluent expression and an appropriate speed.

### **4. Coherence and Completeness of Oral Expression**

In the sixth observation, the students' coherence and completeness have improved significantly. Students can clearly express complete ideas, the logical relationships between sentences have been significantly strengthened, and the expressions are more natural compared to previous observations.

Example 1: A student said, "Yesterday, I went to the mall, bought some clothes, and then had dinner with my family". The sentence is smooth, the information is complete, and the logic is coherent.

Example 2: When discussing hobbies, a student said, "I enjoy playing basketball because it's fun and keeps me healthy". The sentence has a natural connection and the expression is complete and logical.

## **Sixth Observation Record (Control Group) - 11 May 2023**

### **1. Student Participation and Performance in Class**

Description: Most students performed relatively actively in class. They could keep up with the teacher's teaching rhythm and earnestly participate in classroom interactions. Overall, their performance was good and

the teaching effect was satisfactory.

**Existing Problems:** A small number of students were inattentive. They showed a lack of enthusiasm for classroom activities. Some students' eyes wandered during class, often looking out of the window or looking down at their mobile phones, showing obvious fatigue. Some students yawned, appearing to lack mental engagement. Some students looked blank when the teacher asked questions and couldn't understand the questions. Some students left their seats without reason during class, and some students used the excuse of participating in group activities to wander around the classroom. Several students left the classroom early before the end of the class.

## **2. Accuracy of Oral Expression**

**Pronunciation Accuracy:** There has been no obvious improvement in the students' pronunciation accuracy. The error rate in pronunciation is still high. Some students still often make mistakes in the pronunciation of some simple vowels and consonants.

**Vocabulary Usage Accuracy:** The students' vocabulary accuracy has improved to some extent, but the effect is not obvious. Chinglish is still very prominent, and problems such as inappropriate vocabulary collocations and wrong word meanings still occur frequently.

**Grammar Correctness:** The students' grammar mistakes have gradually decreased. They rarely make mistakes in simple sentence expressions and basic grammar such as tenses and voices. However, they still have obvious problems in expressing complex sentences.

## **3. Fluency of Oral Expression**

**Improvement in Frequent Pauses:** Students made fewer pauses in their daily expressions, especially in simple sentences, and their fluency has improved. For example, a student said, "Yesterday, I went to the park and played football with my friends", with almost no pauses and a smooth expression. When describing morning activities, a student said, "I wake up, have breakfast, and then go to school", with a natural expression and fewer pauses.

**Less Repetition in Simple Sentences:** Students made fewer repetitions in simple sentences and could express their thoughts. For example, a student said, "I think this is the right answer", without repeating "I think" as before. When describing an experience, a student said, "We went to the museum and learned a lot", without unnecessary repetitions.

**Improved Speech Speed in Expressing Simple Ideas:** When expressing simple ideas, students' speech speed has increased, and the overall fluency is better. For example, a student said, "Last weekend, I visited my grandparents and had a great time", with a uniform and natural speed. When describing holiday plans, a student said, "I plan to go hiking with my family next month", with a moderate speed and a fluent expression.

**Lack of Improvement in Sentence Cohesion:** The cohesion of sentences in complex sentences is still insufficient, and the logical relationships between sentences are not clear enough. For example, a student said, "I like reading books. I also like football. It is good", lacking connectives and making the relationships between sentences unclear. When describing an activity, a student said, "We went to the park. Played football. It was fun", without obvious logical connections between sentences.

**Common Grammar Errors:** Grammar errors, especially in tenses and subject - verb agreement, are still relatively common. For example, a student said, "She don't like the food", with an incorrect verb form. When describing a past activity, a student said, "Yesterday, I go to the store", and the tense error has not improved significantly.

## **4. Coherence and Completeness of Oral Expression**

In the sixth observation, the oral expressions of the students in the control group still have problems with coherence and completeness. Although there are slight improvements, the overall expression still seems unclear.

**Example 1:** A student said, "Yesterday, I went to the park, and we had a picnic". Although the connective

"and" was used to make the information more coherent, it still lacks specific details such as the types of food or the time of the activity, and the overall content still seems simple.

Example 2: When describing the weather, a student said, "It was sunny, and I enjoyed it". Although the connective "and" was used between the sentences to express the love for sunny weather, it lacks more descriptions and specific feelings, making the overall expression seem somewhat incoherent and incomplete.

## **Appendix 6 Summary of Four Interviews Between the Researcher and the Instructor**

### **First Interview Record**

Time: 9 Feb, 2023 (W8)

Venue: Meeting room of English Department

Interviewee: Martin

**The researcher:** Martin, thank you for accepting this interview. I'm conducting this interview for the needs of my PhD thesis research. The theme of my PhD thesis is about the impact of WFAT on the development of vocational college students' oral English abilities. The content of our interview will focus on the oral English performance of students in the experimental group and the control group in the classroom. The content of our interview will only be used for this research and will not be used for any other purposes. I will also keep the interview content and data absolutely confidential and protect the personal privacy of you and the students. In addition, in order to regularly understand the progress of students' classroom performance, I plan to interview you once a month in the following time. To record the interview content more comprehensively, I would like to record and audio - tape the interview. Do you agree?

**Martin:** Okay, I'm honored to participate in your research. I will also carefully observe the oral English performance of the two groups of students in the classroom. I agree to the audio - taping of the interview.

**The researcher:** Great, thank you very much. Let's get to the point. First of all, what have you observed about the enthusiasm and initiative of students in the experimental group and the control group in using oral English in the classroom over the past month?

**Martin:** During this period, there is no significant difference in the enthusiasm and initiative of students in the experimental group and the control group in using oral English in the classroom. Overall, the enthusiasm of students is generally not high. Specifically, most students are passive in classroom discussions. They rarely speak up or express their opinions in English on their own initiative. Instead, they mostly wait to be called on by the teacher before reluctantly participating. Even when called on to speak, many students' answers are usually short, lacking in - depth thinking or language expansion. They often just repeat or piece together sentences simply.

In addition, students are not very proactive in speaking English, especially in expression scenarios that are more complex or require higher language - organizing abilities, such as during group discussions or when answering open - ended questions. Many students seem to lack confidence and would rather remain silent or avoid expressing their opinions. In the classroom, we often observe that some students are distracted, such as looking down at their mobile phones or whispering with classmates, which further affects the interactive atmosphere of the classroom. These phenomena reflect the common problems of insufficient motivation and poor learning habits among vocational college students in oral English learning, especially in the classroom environment.

**The researcher:** What do you think causes the low enthusiasm of students in oral expression in the classroom?

**Martin:** On the one hand, most vocational college students have a relatively weak English foundation, which makes them feel a lot of pressure when using English for oral expression. Many students' English learning in middle school was exam - oriented, focusing on grammar and vocabulary memorization while lacking listening and speaking training. This learning experience makes them often feel at a loss when facing real - life communication and choose to remain silent for fear of making mistakes. On the other hand, I think it's due to their lack of self - confidence. Due to the long - term lack of oral English practice, they lack confidence in their English - expressing abilities. They are afraid of being laughed at for making mistakes in class, so they tend to avoid speaking. In addition, among vocational college students, many people have a low evaluation of their own language - learning abilities and are prone to develop a fixed mindset of "not being able to learn English well", which further weakens their willingness to actively use English.

**The researcher:** You just mentioned that students lack self - confidence. Is this related to the classroom

model or the classroom atmosphere?

**Martin:** Absolutely. The classroom atmosphere has a significant impact on students' self - confidence. A positive classroom atmosphere can further stimulate students' enthusiasm and self - confidence in classroom participation. Such an environment makes them feel safer and willing to try using newly - learned vocabulary and expressions, thus encouraging their oral English expressions. On the contrary, a tense and dull classroom atmosphere may further affect their self - confidence and participation willingness. Of course, I also need to reflect on my classroom teaching model. My classroom teaching model may be quite different from the traditional teacher - centered teaching model. My classroom mainly focuses on students answering questions, discussing problems, and expressing opinions. This sudden change in the classroom model may have caused them discomfort. In the later stage, I will adjust my teaching strategy and gradually promote student - centered classroom teaching.

**The researcher:** Do you think the students have made any progress in pronunciation during this stage? Is there a significant difference between the two groups of students?

**Martin:** I think the students' pronunciation is generally okay. Although some students make mistakes frequently in diphthongs, lateral sounds, plosives, etc. For example, in the pronunciation of diphthongs, some students often pronounce /aɪ/ as a more single /a/, or lack the gliding feature of the diphthong in the pronunciation of /eɪ/. In the pronunciation of lateral sounds, for words like "light" and "help", some students may replace the lateral sound /l/ with voiceless consonants, resulting in unclear pronunciation. In terms of plosives, especially /t/ and /p/, some students often neglect aspiration at the end of words. For example, they may pronounce "cat" similar to "cad".

Regarding the spelling and pronunciation of polysyllabic words, some students are prone to make mistakes in the stress position or omit certain syllables. For example, for a word like "interesting", some students may pronounce it as /'ɪntrestɪŋ/, ignoring the middle syllable /ə/. This problem may be related to the fact that they have not received systematic phonics training, making it difficult for them to infer the correct pronunciation according to spelling rules when encountering unfamiliar words.

However, it's worth affirming that the students perform well in intonation. When reading sentences or having daily conversations, most of them can correctly use rising and falling intonation to express different emotions and semantics. For example, the intonation naturally rises when asking questions, and it drops appropriately in declarative sentences. This mastery of intonation indicates that students have a strong overall perception ability in pronunciation and intonation.

As for the differences in pronunciation between the two groups of students, there doesn't seem to be a significant difference at present. Whether in the experimental group or the control group, students show similar types and frequencies of pronunciation errors. For example, students in both the experimental group and the control group may weaken the /t/ in "water" to the flap /ɾ/ in American English, or tend to substitute /s/ or /θ/ for "th" when pronouncing it. This similarity may indicate that there is no significant difference in the acquisition of pronunciation and the teaching models used by the two groups for the time being. However, with the in - depth of teaching, there may be subtle changes, which need further observation and analysis.

**The researcher:** What about the students' use of vocabulary and grammar? How did the two groups of students perform?

**Martin:** In terms of the use of vocabulary and grammar, students generally show a certain basic level, but there are still many problems that need to be improved. Specifically, in terms of vocabulary, students' commonly used vocabulary is relatively limited. Especially when describing more abstract or complex topics, they tend to repeatedly use basic vocabulary. For example, some students repeatedly use "like" or "good" when expressing preferences, lacking richer vocabulary alternatives such as "enjoy", "prefer", "fantastic", or "excellent". For some low - frequency or academic words, students generally show insufficient mastery. For example, when talking about "environmental issues", some students may only use "pollution" or "dirty", and are not familiar with words like "sustainability" or "conservation".

In addition, the accuracy of vocabulary collocation is also a common problem. For example, some students may say "solve a question" instead of the more idiomatic "solve a problem" when expressing "solve a problem". Such collocation errors are relatively common in both groups of students, indicating that students still have room for improvement in the contextual learning of vocabulary.

In terms of grammar, students' performance shows certain hierarchical differences. Some students can correctly use basic tenses and sentence patterns, such as the simple present tense, the simple past tense, and simple complex sentences. However, in more complex grammar structures, such as the passive voice, the subjunctive mood, or conditional sentences, students are prone to make mistakes. For example, some students may wrongly express "If I were you, I would..." as "If I was you, I will...". There are also students who are inconsistent in using the definite article "the". For example, they may say "go to school" but wrongly add "the" and make it "go to the school".

As for the performance of the two groups of students, there is no significant difference in the accuracy of vocabulary and grammar at present. Students in both the experimental group and the control group have similar common problems. For example, students in both groups may confuse tenses when describing past experiences, such as expressing past events in the simple present tense, like "Last year, I go to Beijing". Although students in the experimental group have more opportunities to practice vocabulary and grammar outside the classroom through the FiF tool, the impact of this practice on actual oral output may still need time to accumulate.

It's worth noting, however, that students in the experimental group seem to have a slight advantage in vocabulary diversity. For example, some students in the experimental group can use words like "exploring", "adventurous", "fascinating" when describing hobbies, while students in the control group tend to use basic vocabulary like "interesting" or "fun". This difference may be related to the online vocabulary resources and feedback accessed by the experimental group.

**The researcher:** What do you think of the fluency and completeness of the oral expressions of these two groups of students?

**Martin:** There are some notable characteristics and common problems in the fluency and completeness of students' oral expressions.

First, regarding fluency, most students can express themselves basically coherently when talking about daily topics, and their speaking speed is moderate. However, many students still experience frequent pauses, repetitions, and self - corrections in more complex expressions. For example, when describing weekend activities, a student may say, "On Saturday, I... I went to... um... to the park, and... and I... played with my friends." Students need more time to search for vocabulary or sentence patterns when organizing sentences, especially when it comes to unfamiliar vocabulary or longer sentences, and their fluency will decrease significantly. Regarding completeness, most students can express themselves basically completely when expressing simple opinions or events, but when it comes to complex topics or long paragraphs, they are prone to problems such as missing content or unclear logic. For example, when asked about future career plans, some students may only give simple answers like "I want to be a teacher because I like children.", lacking further elaboration on the reasons or goal details. In this regard, the performance of the two groups of students is basically the same, indicating that students' logical - organizing and extended - expressing abilities still need to be strengthened.

**The researcher:** From which aspects do you suggest students should improve their oral English abilities?

**Martin:** I think students should first build self - confidence and a positive language - learning attitude. Psychological factors have an important impact on the fluency and completeness of oral expressions. I suggest that students should try to speak more in class or during practice and not be afraid of making mistakes. Teachers and classmates can also help students gradually build confidence through positive feedback, making them realize that mistakes are part of the learning process, and only by speaking and practicing more can they truly improve their oral English abilities. Second, they should strengthen listening

input and imitate native - like intonation and expressions. Listening input is an important foundation for improving oral English. I suggest that students listen to a certain amount of English materials every day, such as news reports, podcasts, or movie lines, and try to imitate the intonation and speaking speed. For example, by watching English videos with subtitles, first understand the content, and then follow and imitate, constantly adjusting their intonation and pronunciation. Finally, they should improve their fluency through diverse practices. For example, regularly participate in group discussions or English corner activities to simulate real - life communication scenarios. Conduct 2 - 3 - minute impromptu speeches on topics every day (such as describing a day's activities or expressing opinions on a certain topic) to gradually improve the coherence of expression and speaking speed.

**The researcher:** Thank you very much for sharing your insights, Martin. Your feedback is very helpful for us to understand the students' performance. I'm looking forward to the next interview to further explore the development of students' oral and expressive abilities.

**Martin:** It's my pleasure.

### **Second Interview Record**

Time: 2 Mar, 2023 (W11)

Venue: Meeting room of English Department

Interviewee: Martin

**The researcher:** Martin, thank you for accepting my interview again. Today, let's continue to talk about the classroom English performance of the two groups of students at this stage. To record the interview content more comprehensively, I would like to record and audio - tape the interview. Do you agree?

**Martin:** Okay, I'm happy to be interviewed by you. I agree to the audio - taping of the interview.

**The researcher:** Great, thank you very much. Let's get to the point. First of all, what do you think of the enthusiasm and initiative of the two groups of students in using oral English in the classroom at this stage? Have the students improved in this regard compared with the previous stage? Is there a significant difference between the two groups of students in this regard?

**Martin:** Overall, the enthusiasm and initiative of students in using oral English in the classroom have indeed increased at this stage, but the degree of improvement varies. The students in the experimental group are more active than those in the control group, especially in classroom discussions and interactive activities. They are more willing to use English to express their opinions and ideas frequently, not just speaking when called on by the teacher. This initiative is manifested in group discussions and when answering open - ended questions. The students in the experimental group are more inclined to organize relatively complete expressions in English and try to expand the topic. The participation of students in the control group has also increased, but they are more likely to passively respond to the teacher's questions, and their willingness to participate actively is still insufficient.

**The researcher:** What about the self - confidence of the two groups of students in oral English expressions in the classroom? Have the students improved in this regard compared with the previous stage? Is there a significant difference between the two groups of students in this regard?

**Martin:** The self - confidence of students has generally increased, especially among students in the experimental group, who show more progress. When speaking in class, they seem more natural, their speaking speed is more stable than in the previous stage, and their intonation is more confident. The students in the experimental group are more willing to try to express complex opinions. Even if they may make mistakes, they are willing to complete their speeches. In contrast, students in the control group are still more cautious in similar situations. Although the self - confidence of students in the control group has also increased, some students tend to shorten their answers or avoid complex expressions due to uncertainty about the correctness of their language. This difference may be related to the more frequent practice opportunities obtained by the experimental group through using the WFAT tool.

**The researcher:** Do you think the students have made any progress in pronunciation during this stage? Is there a significant difference between the two groups of students?

**Martin:** Both groups of students have indeed made progress in pronunciation, especially in the accuracy of pronouncing common words and phrases. The students in the experimental group perform better in the stress position and syllable integrity. Especially for some polysyllabic words, their pronunciation is clearer and more natural. The pronunciation of students in the control group has also improved, but in terms of pronunciation details, such as the clarity of consonants or the glide of vowels, the experimental group is more stable. The progress of the experimental group may be related to their frequent use of the WFAT tool for pronunciation correction and imitation practice, while the control group mainly relies on classroom teaching for pronunciation correction, so the progress is slightly smaller.

**The researcher:** What about the students' use of vocabulary and grammar? Is there a difference in the performance of the two groups of students?

**Martin:** The students in the experimental group have made more obvious progress in vocabulary diversity and contextual use. They can flexibly choose vocabulary according to the specific topics of classroom discussions and try to expand their answers. The vocabulary use of students in the control group is relatively more limited. They usually tend to use basic vocabulary and common expressions and lack the ability to further expand. In terms of grammar, the students in the experimental group are more proficient in mastering basic tenses and using sentence structures. For example, they make fewer mistakes when describing events or expressing opinions, while students in the control group are still prone to make mistakes in subject - verb agreement and verb tenses, which affects the completeness and accuracy of their expressions.

**The researcher:** What do you think of the performance of the two groups of students in terms of the fluency, completeness, and appropriateness of oral expressions?

**Martin:** The oral expressions of both groups of students have improved at this stage, but the experimental group performs slightly better. In terms of fluency, the students in the experimental group can express their ideas more coherently, and the phenomena of pauses and repetitions have significantly decreased. In terms of completeness, the students in the experimental group can provide more comprehensive answers around a question, such as explaining the reasons or providing background information, while the expressions of students in the control group are often limited to simple sentences and lack content expansion. In terms of appropriateness, both groups of students still need to improve in choosing vocabulary and expressions. Sometimes, they may use words inaccurately or deviate from the context. However, the overall performance of the experimental group is closer to the actual language - use situation than that of the control group.

**The researcher:** Thank you very much for sharing your detailed insights, Martin. Your feedback is very valuable for our research and also provides important references for subsequent teaching practices. I'm looking forward to the next interview to discuss more progress of the students.

**Martin:** You are welcome.

### **Third Interview Record**

Time: 12 Apr, 2022 (W17)

Venue: Meeting room of English Department

Interviewee: Martin

**The researcher:** Martin, first of all, compared with the previous stage, do you think the enthusiasm and initiative of the two groups of students in oral English expressions in the classroom have improved? What's the overall situation like?

**Martin:** There has indeed been an improvement. The enthusiasm and initiative of both groups of students in oral English expressions in the classroom are better than in the previous stage, especially when participating in discussions and answering open - ended questions. However, there are still a small number of students who are not very active, especially those who have little interest in English or lack self - confidence. Usually,

these students remain silent during group activities, relying more on other classmates. Even when called on, their answers are often very short or they simply switch to Chinese.

Overall, the students in the experimental group are more active. For example, in one class, I asked the students to discuss a hypothetical situation: "If you could change one school rule, what would you choose to change?" The students in the experimental group were more willing to express their ideas in English and even took the initiative to start the discussion. In contrast, the students in the control group relied more on the teacher's prompts. Sometimes, even when called on, their answers were relatively simple, such as single - sentence responses like "Change the uniform", lacking further expansion and logical support.

**The researcher:** Do you think there have been any changes in the self - confidence of the two groups of students in oral English expressions in the classroom? Is there a significant difference between the experimental group and the control group?

**Martin:** In terms of self - confidence, the improvement in the experimental group is more obvious than that in the control group. This is mainly reflected in their performance when speaking. For example, their speaking speed is more stable, their intonation is more natural, and they are more willing to try complex sentence patterns. When answering questions, even if their grammar or vocabulary is not accurate sometimes, the students in the experimental group still strive to complete their expressions, and a stronger sense of self - confidence can be heard from their pronunciation and intonation. For example, in a class discussion about "future career plans", a student in the experimental group fully elaborated her interest in nursing in English. Even though there were some grammar mistakes in the middle, she still maintained strong logic and fluency. In contrast, students in the control group, in similar situations, mostly chose short answers or directly avoided complex expressions.

Although the self - confidence of students in the control group has also increased, when facing unfamiliar expressions or being asked to answer more complex questions, they often show hesitation and nervousness. For example, when discussing "an important childhood memory", a student tried to speak several times but finally chose to explain in the mother tongue. This difference may be related to the fact that the experimental group has more opportunities to get feedback and error - correction when using the WFAT tool, which helps them gradually build self - confidence.

**The researcher:** Have the two groups of students made any progress in pronunciation? Is there a significant difference between the experimental group and the control group?

**Martin:** Both groups of students have made progress in pronunciation. Especially in the pronunciation of some common words in daily communication, such as "people" and "important", they are more accurate. However, the experimental group performs better in the overall pronunciation quality. The students in the experimental group pay more attention to some details. For example, they handle the aspiration of consonant endings and the stress position better. They perform particularly well in polysyllabic words. For example, for words like "communication" and "responsibility", they can basically make the stress clear and the syllables complete.

Although the control group has also improved in basic pronunciation, they are still prone to problems in some details. For example, they may pronounce /θ/ as /s/, or omit the final sound /d/ in words like "world". The improvement of the students in the experimental group is more obvious, especially in phrase reading and free expression. The overall fluency and accuracy of their pronunciation are better, which may be related to their repeated pronunciation practice using the WFAT tool.

**The researcher:** What about the use of vocabulary? Is there a difference in the performance of the two groups of students?

**Martin:** The students in the experimental group indeed perform better in the accuracy, diversity, and specificity of vocabulary. They can not only choose words accurately but also use more advanced vocabulary and more specific descriptions in their answers. For example, in a class discussion on "how to maintain a healthy lifestyle", the students in the experimental group would mention "balanced diet", "regular exercise",

"mental health", and even try to use words like "stress management". In contrast, the students in the control group mostly use simple expressions like "eat vegetables" or "do sports", lacking further expansion. In terms of vocabulary collocations, the students in the experimental group are also more accurate. For example, they use "take action" instead of "do action", or "make progress" instead of "do progress". The students in the control group make relatively more mistakes in collocations, which also indicates that the experimental group has more abundant language input and practice opportunities, and their ability to use vocabulary in context has been better developed.

**The researcher:** What about the use of grammar? Is there a significant difference between the two groups of students?

**Martin:** In terms of grammar use, the students in the experimental group perform better. They make fewer mistakes in mastering basic tenses and sentence patterns. Especially when describing events, they can correctly use the past tense and the present perfect tense. For example, when describing past experiences, more students in the experimental group would use expressions with correct structures like "Last year, I travelled to..." or "It has been an amazing experience", while the students in the control group are more likely to confuse tenses, such as using incorrect sentences like "I travel last year".

In addition, the students in the experimental group begin to try using more complex sentence patterns, such as conditional sentences ("If I had more time, I would learn Spanish") and the passive voice ("This project was completed by our team"). The students in the control group tend to use simple sentences more, use complex sentence patterns less frequently, and are even prone to omit conjunctions or make grammar mistakes in longer sentences. For example, a student described an activity as "Yesterday I go park, meet friends", lacking coherence and correct tenses.

**The researcher:** What about fluency, completeness, and appropriateness?

**Martin:** In terms of fluency, the students in the experimental group perform more smoothly. They have fewer pauses when expressing themselves, and the connection between sentences is more natural. For example, when describing weekend activities, the students in the experimental group can clearly express the process and feelings of the events with almost no obvious interruptions or repetitions. In contrast, the students in the control group are still prone to frequent pauses and self-corrections due to uncertainties in vocabulary or grammar when expressing themselves.

In terms of completeness, the students in the experimental group can answer questions more comprehensively, especially in providing details and explanations. For example, when discussing "future plans", they can talk about specific goals, implementation methods, and possible difficulties, while the students in the control group mostly stay at simple descriptions, such as "I want to learn English well", lacking further expansion. In terms of appropriateness, there is not much difference in the performance of the two groups of students. Although the experimental group has made more obvious progress in other aspects, when choosing vocabulary and expressions, they sometimes still use words that are not precise enough or sentences that deviate from the theme. For example, a student used the expression "save trees" when discussing environmental protection, instead of the more appropriate "reduce deforestation".

**The researcher:** Thank you very much for your detailed feedback, Martin. Your observations are very important for our research. I'm looking forward to seeing further progress of the students in the next interview.

**Martin:** You are welcome.

#### **Fourth Interview Record**

Time: 11 May, 2023 (W20)

Venue: Meeting room of English Department

Interviewee: Martin

**The researcher:** Martin, thank you for accepting my interview again. Today, let's continue to talk about the

classroom English performance of the two groups of students at this stage. To record the interview content more comprehensively, I would like to record and audio - tape the interview. Do you agree?

**Martin:** Okay, I'm happy to be interviewed by you. I agree to the audio - taping of the interview.

**The researcher:** Great, thank you very much. Let's get to the point. First of all, what do you think of the enthusiasm and initiative of the two groups of students in using oral English in the classroom at this stage? Have the students improved in this regard compared with the previous stage? Is there a significant difference between the two groups of students in this regard?

**Martin:** Compared with the previous stage, the enthusiasm and initiative of both groups of students in oral English expressions in the classroom have increased, especially when participating in classroom discussions and group interactions. The students are more willing to try to express their opinions in English in class. In some cases, they even take the initiative to ask questions related to the classroom theme. The students in the experimental group are particularly outstanding in this regard. They are not only more willing to participate in classroom activities in English but also show more initiative during the activities. For example, in one class, I asked the students to discuss "environmental solutions for future cities". A student in the experimental group took the initiative to speak, not only expressing his own ideas in English but also asking for the opinions of other classmates, which drove the interactive atmosphere of the whole group.

In contrast, although the students in the control group have also made progress, their initiative is still relatively low. Most students in the classroom still rely more on being called on by the teacher to speak. Even when they participate, their speaking content is relatively short, and there are few cases of active expansion. For example, in group discussions, the students in the control group tend to follow the active members and lack the willingness to express themselves actively.

However, there are still students who are not active in both groups. These students usually have a weak English foundation or are not very interested in learning. They are relatively silent in class, trying to avoid interacting with others. Sometimes, even when called on, they try to end the communication with short answers. I observed that these students often choose to be silent because they are worried about making mistakes, which reflects that their self - confidence and language application ability still need further cultivation.

**The researcher:** What do you think are the reasons for the enhanced enthusiasm and initiative of students in oral English expressions in the classroom?

**Martin:** This is the result of the combined action of multiple factors. First of all, the students' internal motivation plays a big role. Many students have realized the importance of oral English ability for their future career development. For example, it can help them better participate in international exchanges, perform better in job interviews, and gain more opportunities in the workplace. Especially the students in the experimental group have clearer goals and a more positive learning attitude.

Secondly, the application of the WFAT tool in the experimental group is also an important reason. Through functions such as voice error - correction, instant feedback, and simulated conversations, this tool provides students with more opportunities for language input and output. This practice method helps students discover their own shortcomings and gradually improve. Especially, they can continue to strengthen their oral English ability outside the classroom, making them more confident to express themselves in English in class.

In addition, the transformation of the classroom teaching model has also provided students with more opportunities to communicate in English. For example, through student - centered group discussions and topic debates, students have more opportunities to practice and interact, and gradually get used to expressing themselves in English in the classroom. Especially the students in the experimental group are more open in class, which may be related to the continuous implementation of these activity forms.

Finally, the positive classroom atmosphere also plays a facilitating role. Students learn in a more supportive environment where mistakes are no longer regarded as failures but as part of the learning process. This greatly reduces their psychological pressure and makes them dare to try to express themselves in English.

**The researcher:** What about the self - confidence of the two groups of students in oral English expressions in the classroom? Have the students improved in this regard compared with the previous stage? Is there a significant difference between the two groups of students in this regard?

**Martin:** In terms of self - confidence, both groups of students have improved compared with the previous stage, but the improvement in the experimental group is more obvious. This improvement is reflected in many aspects. For example, in class, the students in the experimental group try to use more complex sentence patterns. Even if there are mistakes in their expressions, they can still complete their speeches fluently. Their speaking speed is more stable, and their intonation is more natural. For example, when discussing "how artificial intelligence changes our lives", a student was able to state his views completely without preparation and put forward some logical expansions, such as its impact on work efficiency and potential threats to privacy.

Although the self - confidence of students in the control group has also improved, it is more reflected in simple tasks, such as reading aloud or answering direct questions. They still seem less confident when facing complex questions and often pause or even give up halfway because they are not sure whether their vocabulary or grammar is correct. For example, in an open - ended discussion, a student tried to express his opinion but only completed half of it due to difficulties in language organization. This difference reflects the auxiliary effect of the WFAT tool on the students in the experimental group. The practice and feedback provided by the tool have enhanced their confidence in expression.

**The researcher:** Have the two groups of students made any progress in pronunciation? Is there a significant difference between the experimental group and the control group?

**Martin:** In terms of pronunciation, both groups of students have made obvious progress, which is mainly reflected in the accuracy of word pronunciation and the naturalness of sentence intonation. For example, there are many fewer mistakes in the pronunciation of some common diphthongs (such as /ai/) and final consonant sounds (such as /t/ and /k/). The progress of the students in the experimental group in pronunciation details is more remarkable. For example, when reading long sentences, they can use the correct stress and intonation better, especially when emphasizing certain key words, they seem more natural. Although the students in the control group have made progress, they still have problems with some pronunciation habits. For example, they often ignore the final sound or misplace the stress in polysyllabic words. For example, they may misplace the stress of "opportunity".

The pronunciation advantage of the students in the experimental group may be related to their repeated pronunciation practice and correction in the WFAT tool. This targeted training helps them discover and correct their problems more quickly.

**The researcher:** What about the use of vocabulary and grammar by the students? Is there a difference in the performance of the two groups of students?

**Martin:** In terms of vocabulary use, the accuracy and diversity of the students in the experimental group are more prominent. They can use more complex and specific vocabulary combined with the classroom theme. For example, when discussing "urban environment", they can use words such as "sustainable development", "green technology", "urban planning", while the students in the control group mostly stay at basic vocabulary such as "pollution" or "clean water", lacking deeper language expansion.

In terms of grammar, the basic grammar of the students in the experimental group is more solid, and they perform better in the use of tenses and sentence patterns. For example, they can use the perfect tense, conditional sentences, and the passive voice more accurately to express complex ideas. In contrast, the performance of the students in the control group in these aspects is relatively weak. Tense confusion and sentence - pattern mistakes are still common. For example, they often make mistakes when switching between the past tense and the simple present tense, and they are not proficient in using sentence - logic connectives.

**The researcher:** What do you think of the performance of the two groups of students in terms of the fluency,

completeness, and appropriateness of oral expressions?

**Martin:** In terms of fluency, the performance of the experimental group is more prominent. They have significantly fewer pauses in classroom discussions, the connection between sentences is more natural, and they can express complete views more smoothly. In contrast, the students in the control group still have more pauses and repetitions due to insufficient vocabulary reserves or difficulties in sentence - pattern organization. In terms of completeness, the students in the experimental group perform better. They are better at answering questions around the theme. For example, when describing future plans, they can expand from multiple angles such as goals, steps, and challenges, while the answers of the students in the control group are often relatively short and lack detailed support.

In terms of appropriateness, there is still room for improvement in the performance of both groups of students. For example, when discussing more abstract topics, such as "the significance of cultural diversity", many students would use simple words like "good" or "helpful" to describe, lacking more specific and precise expressions. This shows that they still need more guidance and training in language appropriateness.

The researcher: Do you think the students' oral English language ability has improved?

**Martin:** From the perspective of language communication functions, overall, the oral English expression abilities of both groups of students have indeed improved. Especially when completing basic daily communication tasks, such as answering teachers' questions, participating in simple conversations, reading aloud, and expressing personal opinions on a certain topic, these tasks are no longer difficult for them. The experimental group has made more significant progress than the control group overall. They perform more prominently in terms of fluency, completeness, and accuracy in the classroom. This difference may be related to the fact that the students in the experimental group have long - term used the WFAT tool for self - study and intensive practice. With the assistance of the tool, they have obtained more feedback and language practice opportunities. However, from the perspective of language ability itself, students' language ability is a complex multi - dimensional system, including pronunciation, vocabulary use, grammar use, the complexity of syntactic structures, and communicative competence, among other aspects. In these specific dimensions, although students have improved, there are still significant weaknesses. Especially when facing complex language tasks, such as giving a free speech, elaborating on a topic in depth, or expressing views with logic and hierarchy, their performance is often not ideal. Specifically, the common problems of students in complex language tasks include the following points. First, many students, including those in the experimental group, lack a clear logical structure when expressing complex views. For example, when asked to describe a hypothetical scenario or solve a specific problem, many students only stay at stating the phenomenon and cannot further expand on the reasons, impacts, or solutions.

Second, there is a lack of critical and in - depth thinking abilities. When answering open - ended questions or expressing themselves freely, many students can only express superficial opinions and lack the ability to conduct in - depth analysis or critical thinking. This problem may be related not only to language abilities but also to the cultivation of thinking habits. For example, when asked "What do you think of cultural diversity?", students usually can only simply answer that "culture is important" and have difficulty further elaborating on its specific significance or the challenges it brings.

**The researcher:** Thanks a lot.

**Martin:** You are welcome.

## Appendix 7 Summary of the Student Focus Group Discussion

### 第一部分：对 FiF 的使用感受

研究者：首先，我们来谈谈你们对 FiF 的初步印象。请每位同学分享一下你们使用这个工具的感受。

学生 A：我对 FiF 的第一印象是它非常直观，界面设计也很友好。刚开始使用时，我觉得这个工具很容易上手，特别是录音功能让我感到很有趣。通过录音，我能清晰地听到自己的发音，进而调整。

学生 B：我同样觉得界面很好用，颜色和布局都很吸引人。初次使用时我有点紧张，但看到界面如此友好后，我的紧张感减少了很多。整体上，我喜欢这个工具的灵活性。

学生 C：我对 FiF 的设计感到惊喜。它不仅功能齐全，还能让我在练习口语时感到放松。最让我高兴的是可以随时随地进行练习，真的很方便。

学生 D：我的初步印象是，这个工具确实提供了一个很好的平台，特别是能够让我记录和回放自己的声音。然而，我在使用过程中发现，FiF 平台的打分有时不准确。系统给出的评分让我有些困惑，尤其是在发音和流利度上。

学生 E：我也觉得反馈功能有时不够理想。有时候系统对错误的识别并不准确，比如我认为自己的发音很标准，但系统却显示我有错误，这让我感到沮丧。

学生 F：我觉得反馈有时显得程序化，缺乏针对性。比如，有些反馈让我觉得是固定的模板，没有考虑到我的具体情况。我希望能有更个性化的反馈，针对我出现的具体错误提供改进建议。

学生 G：我同意 F 的看法，反馈虽然及时，但缺乏个性化的策略指导功能。有时候，我需要的是关于如何改善我的表达的具体指导，而不是简单地指出我哪里错了。

学生 H：我在使用过程中也有些负面体验。FiF 的互动功能虽然好，但我还是觉得面对面的练习效果更好。FiF 的自动反馈有时候不能及时解决我在口语练习中遇到的问题。

学生 I：此外，有时我会觉得系统的智能化纠错功能缺失。很多时候，我希望能够获得更深入的分析，比如为什么我的某个句子结构有问题，或者具体的发音错误是在哪里。

学生 J：我在使用过程中发现，有时内容的相关性不够。我希望能够看到更多与我们课堂学习主题直接相关的练习，而不是那些看起来无关的内容。

学生 K：虽然我很享受使用 FiF 的过程，但也希望系统能提供更高质量的反馈。有时候，反馈的信息非常简单，我想要更深入的分析，这样我才能更好地理解自己的错误。

学生 L：我最初使用这个工具时，觉得它的功能非常强大，但有时候我会觉得有些复杂，尤其是在寻找特定功能的时候需要花时间去适应。

学生 M：我觉得 FiF 整体上是是个很好的学习工具，但我也有过不太愉快的体验。有时候系统提示不够明显，导致我在录制时不能及时发现问题。

学生 N：总体来说，我认为这个工具对我的学习帮助很大。虽然开始时有些负面体验，但随着使用的深入，我逐渐适应了这个平台，感到越来越自信。

学生 O：我对于 FiF 的使用感受总体来说是积极的，虽然在技术上有些小瑕疵，但它的实用性和功能性让我觉得非常值得使用。

学生 P：在使用过程中，我有时候会觉得自己一个人练习的效果不如在课堂上与同学互动时好。FiF 虽然有互动功能，但总觉得不如面对面的交流有效。

学生 Q：我同样在适应过程中有些挑战，但我发现通过反复练习，我对这个工具的使用变得更加熟悉，特别是在录音和播放回放的环节。

学生 R：我觉得整体上 FiF 给我带来了积极的体验，虽然有时候也会有些小问题，但它的优势明显。我相信，随着时间的推移，这些小问题都能得到改善。

学生 S：我对于 FiF 的功能表示认可，尽管遇到了一些小问题，但它的实用性让我感到这个工具对我的口语能力有很大的帮助。

学生 T：我觉得使用 FiF 的过程让我发现了自己的很多不足，尽管有些时候反馈不够细致，但整体体验让我受益良多。

研究者：非常感谢大家分享的使用感受！你们的反馈对我们改进 FiF 非常有价值。接下来，我们可以继续讨论对 FiF 的满意度以及功能的重要性。

### 第二部分：对 FiF 的满意度

研究者：接下来，我们讨论一下你们对 FiF 整体的满意度。你们可以分享一下自己满意或不满意的理由，以及对比一下市面上其他英语口语自动评价系统的优缺点。

学生 A：我对 FiF 的整体满意度很高。特别是它的反馈功能，让我能及时了解到自己的发音和语法错误。这种即时反馈的方式非常有效，特别是在我进行口语练习时，能够让我快速纠正错误。

学生 B：我同样很满意 FiF。相比市场上其他工具，我觉得它的界面更加友好，操作也更简单。大多数口语评价系统的界面都显得复杂，让人很难上手，而 FiF 则做得很好。

学生 C：不过，我有一些不满的地方。例如，FiF 的打分有时候不太准确。有时我会觉得自己发音很好，但系统的评分却很低。相比之下，我曾经使用过的某些口语评价工具，尽管也有打分问题，但在反馈的针对性上做得更好。

学生 D：我有类似的体验。FiF 有时候对错误的识别并不准确，这让我感到困惑。而且反馈有时显得程序化，缺乏个性化的指导。我觉得在这方面市场上某些流行的工具，例如某些知名的口语学习应用，就提供了更智能的纠错功能，可以更具具体地指出发音和语法错误的来源。

学生 E：我也觉得 FiF 的反馈缺乏个性化。每次练习后，虽然能得到反馈，但这些反馈往往是泛泛而谈，没有针对性，无法帮助我真正理解错误的根本原因。

学生 F：说到不满意，我还想提到的是，FiF 的互动功能虽然好，但有时候在实时讨论和交流方面的灵活性不足。相比市场上其他工具，一些应用程序提供了更强的社交功能，让我们能够更方便地与同学交流和讨论。

学生 G：但我依然认为 FiF 在流利性和发音方面的帮助是显而易见的。尽管有些不足，我在使用 FiF 的过程中，口语水平确实得到了提升。这种提升是其他自动评价系统所不能比拟的，特别是在自主练习的灵活性方面。

学生 H：我也很赞同 G 的观点。虽然 FiF 存在一些不足，但它依然是我见过的少数能真正帮助我提高口语表达的工具之一。其他工具虽然有各种优点，但往往在反馈的及时性和有效性上不如 FiF。

学生 I：我认为 FiF 的互动性是一个亮点，能够和同学一起练习是一种很好的学习方式。然而，正如前面提到的，其他工具在这方面可能做得更好。比如，有些应用程序提供了更强大的实时讨论功能，能够让学生之间更方便地交流和互相帮助。

学生 J：我还想说，FiF 的策略指导功能缺失是我比较不满意的一点。许多市面上的工具会提供一些学习策略和建议，帮助学生在遇到问题时能够更有效地解决，而 FiF 在这方面的支持相对较少。

学生 K：虽然存在这些不足，我觉得 FiF 在一定程度上能激励我学习。它的实时反馈和进度追踪让我对自己的学习有了更多的掌控感。相比之下，一些其他的自动评价系统则没有这样的激励机制。

学生 L：我觉得 FiF 整体上在设计上考虑了用户体验，虽然它也有一些技术问题和反馈不足，但我仍然认为它比许多市面上流行的工具更适合我们这种学习模式。

学生 M：我想补充的是，虽然市场上有很多自动评价系统，但 FiF 让我感觉更加个性化。我能在这里找到适合自己的练习方法，而不是仅仅跟着系统的步骤走。

学生 N：说到个性化，我希望 FiF 能够在这方面进一步改进。虽然它的反馈系统在一定程度上满足了我的需求，但如果能根据我的学习习惯和薄弱环节提供定制化的建议，那将更加理想。

研究者：大家的反馈非常重要。通过你们的讨论，我感受到 FiF 在某些方面的确有其优势，但同时也暴露出不少需要改进的地方。接下来，我们可以继续讨论 FiF 的功能重要性，以及具体的改进建议。

### **第三部分：FiF 的功能重要性**

研究者：接下来，我们讨论 FiF 的功能重要性。请每位同学分别分享你们对评价功能、反馈功能和互动功能在英语口语学习和评价中的看法，以及这些功能如何影响你们的英语口语发展。

学生 A：我认为 FiF 的评价功能非常重要。它让我能够清晰地了解自己的学习进展。每次练习结束后，系统会给出详细的评分，这让我知道自己在发音、流利度和语法方面的表现。这样的客观评价能够有效地指导我的学习，让我明确自己需要在哪些方面进行改进。

学生 B：我也非常重视评价功能。它不仅让我看到了自己的进步，还激励我设定新的学习目标。比如，我希望在下次练习中获得更高的分数，这种竞争意识让我更加努力。

学生 C：对于反馈功能，我觉得它的实用性和有效性在于能够提供即时的信息。每次练习后，我都能收到详细的反馈，特别是针对我的发音和语法错误。这种及时的反馈让我能快速调整自己的练习策略，

非常有效。

学生 D：反馈功能让我意识到自己的不足之处。有时候，我会在录音后看到一些我没有察觉的错误。及时的反馈让我能够在下一次练习中纠正这些问题，从而提高我的口语能力。

学生 E：我特别喜欢互动功能。能够和同学一起练习是一个很大的优势。我们可以互相纠正、分享学习经验，这让我在口语表达上更加自信。我发现，通过和同学们的讨论，我能更好地理解各种表达方式，这对我的学习非常有帮助。

学生 F：我同意 E 的看法。互动功能在促进学习方面起到了重要作用。通过与同学们的实时交流，我们能够互相激励，学习气氛也因此变得更加活跃。

学生 G：评价功能让我明确了自己的学习目标。每次看到自己的评分，我会更加努力去改进，尤其是在流利性和语法方面的表达。

学生 H：关于反馈，我认为它的有效性在于能够让我看到具体的错误和需要改进的地方。这种反馈的针对性让我在后续的练习中能有的放矢，特别是在发音和语法使用上。

学生 I：互动功能让我在学习中感到更轻松。通过和同学们的交流，我能获取不同的观点，帮助我更全面地理解课程内容。这种互动的学习方式不仅让我能更快提高口语能力，也让学习变得更加有趣。

学生 J：我觉得 FiF 的评价功能不仅仅是提供分数，它还为我设定了一个学习标准。看到自己的分数逐渐提高，我的自信心也随之增强。这种正向反馈对于我而言非常重要。

学生 K：反馈功能的实用性在于它能让我迅速了解自己的表现。每次练习后，能够立即看到反馈，这种及时性让我能在短时间内纠正错误，进而提高口语表达的准确性。

学生 L：互动功能则是我的一个学习动力。通过与同学们的合作练习，我能学到许多不同的表达方式，这让我在口语学习中更具信心。

学生 M：我认为评价功能对我的学习发展起到了至关重要的作用。通过清晰的评价标准，我能够不断调整自己的学习策略，朝着目标前进。

学生 N：反馈的有效性让我在每次练习后都有所期待。每次录音结束后，我都迫不及待想查看反馈，这种反馈让我感受到系统在支持我提升。

学生 O：我觉得互动功能让我们能够建立更好的学习关系。通过小组讨论，我们不仅可以学习口语，还能加深彼此的理解，这样的学习环境让我感到非常舒适。

学生 P：对于评价功能，我认为它能帮助我识别弱点。我经常查看自己的分数和评价，了解哪些方面需要加强，从而制定更有效的学习计划。

学生 Q：反馈功能的即时性让我能快速调整学习方向。我希望能看到更多具体的反馈，特别是在复杂句子的使用上，这样能更好地帮助我理解。

学生 R：互动功能的存在让我在学习过程中感受到支持。与同学的交流让我在口语表达上更加自然，也能在讨论中不断提升。

学生 S：我认为 FiF 的三大功能是相辅相成的。没有良好的评价，就无法明确自己的目标；没有及时的反馈，也无法发现问题；没有互动，学习的过程就会显得单调乏味。

学生 T：我希望在未来，FiF 能继续增强这些功能的智能化，让每个学生都能享受到最佳的学习体验，进一步促进我们的口语能力发展。

学生 U：总体来说，我觉得 FiF 在这三个方面的表现都很不错。它帮助我在口语学习中取得了显著的进步，同时也让我感受到语言学习的乐趣。

学生 V：我非常期待 FiF 能在未来的版本中改善评价、反馈和互动的功能，使我们能够在获得更多支持和帮助。

研究者：非常感谢大家的分享！你们对 FiF 的评价功能、反馈功能和互动功能的看法都很深入。通过你们的讨论，我感受到 FiF 在某些方面的确有其优势，但同时也暴露出不少需要改进的地方。接下来，我们可以继续讨论具体的改进建议，以便更好地满足你们的学习需求。

#### **第四部分：遇到的问题**

研究者：现在我们来谈谈在使用 FiF 过程中遇到的问题。请每位同学分享一下你们在使用中遇到的技术问题、内容方面的挑战，以及这些问题对你们的使用体验的影响。

学生 A：在使用 FiF 时，我遇到了一些技术问题，特别是网络连接不稳定。每当我在录音时，系统偶

尔会卡住，导致我不得不重新开始。这种情况让我感到很沮丧，因为我已经进入了练习状态。技术问题不仅影响了我的学习进度，还让我在完成练习时感到无比烦躁。

学生 B：我同样经历过网络延迟的问题。尤其在高峰时段，我的录音常常无法保存，这让我感到时间被浪费了。这样的技术问题让我在准备练习时变得焦虑，因为我不知道接下来会发生什么。

学生 C：我遇到的技术问题主要是系统崩溃。有一次，我在进行重要的练习时，系统突然关闭，我几乎完成了一段精彩的录音。这不仅让我失去了一次练习机会，还让我对这个工具的稳定性产生了怀疑。

学生 D：除了技术问题，内容方面的挑战也是一个大问题。有时候，我觉得系统提供的练习与我正在学习的内容不太相关。这种内容的错位让我在练习时感到困惑，导致我无法有效利用所学知识。

学生 E：我同意 D 的观点。FiF 的练习内容有时让我感到无趣，特别是在我想要针对特定主题进行深入练习时，系统却没有提供相应的内容。这种情况让我在学习过程中感到失落。

学生 F：在内容方面，我认为缺乏多样性也是一个挑战。有时候练习的形式单一，让我觉得学习变得枯燥。希望能够看到更多情景对话或小组讨论的练习，这样我才能在实践中更好地应用所学。

学生 G：技术问题确实影响了我的学习体验。有时我需要多次尝试才能成功录音，这让我感到挫败。这种挫败感让我在下次使用时会有些犹豫。

学生 H：我觉得技术问题直接影响了我的自信心。每当系统出现故障时，我都会担心自己是不是用错了工具，这种不确定感让我在练习时感到不安。

学生 I：关于内容的相关性，我认为 FiF 应该更好地匹配我们课堂上的学习内容。有时练习的主题与我所学的知识点无关，这让我在学习中感到迷茫。

学生 J：我也遇到过内容方面的挑战。希望 FiF 能根据我们的学习进度调整练习的难度，这样每个人都能在合适的水平上学习，避免无效的练习。

学生 K：关于技术问题，我觉得在使用过程中，系统的反应速度有时让人不满意。比如，有时候反馈加载很慢，这会影响我的学习进度和效率。

学生 L：内容的单一化让我在学习中感到乏味。有时我希望能看到一些具有挑战性的练习，这样我才能真正提高我的口语能力。

学生 M：我觉得技术问题影响了我的学习动力。每次遇到系统崩溃，我都需要重新开始，这让我很沮丧，常常想要放弃练习。

学生 N：在内容方面的挑战中，我发现某些练习的设置不够灵活。我希望能有更多选择，让我可以根据自己的需求选择适合的练习。

学生 O：关于技术问题，我希望 FiF 能在后续的更新中提高系统的稳定性。这样我就能更放心地进行录音练习，而不必担心技术故障。

学生 P：对于内容的相关性，我希望能看到更多与现实生活和交流相关的练习。这样可以让我更好地准备实际使用英语的情境。

学生 Q：我认为技术问题和内容的相关性直接影响了我的学习效果。有时我需要花费太多时间解决技术问题，这让我无法专注于提高口语能力。

学生 R：我发现互动功能在我的学习中也受到技术问题的影响。比如，在进行小组讨论时，如果有同学的网络不稳定，就会影响我们整个小组的交流效率。

学生 S：内容方面的挑战让我在某些练习中感到困惑。有时我并不知道哪些内容是我需要加强的，这让我在练习时缺乏方向。

学生 T：我觉得 FiF 的技术问题使我在进行练习时变得不那么自信。每次遇到系统故障，我都会担心自己是否能顺利完成练习。

学生 U：关于内容的相关性，我希望系统能够根据我们的学习进度推荐适合的练习，这样能让我更好地跟上学习节奏。

学生 V：总体而言，技术问题和内容方面的挑战都对我的学习体验造成了一定的影响。我希望 FiF 能不断改进，提供更好的学习支持。

研究者：感谢大家的分享！你们对使用过程中遇到的技术问题、内容方面的挑战以及这些问题对使用体验的影响的看法非常全面。接下来，我们可以继续讨论具体的改进建议，以便更好地满足你们的学习需求。

## 第五部分：FiF 需要改进的地方

研究者：现在我们来讨论 FiF 需要改进的地方。请每位同学分享一下你们在使用过程中发现的功能上的改进建议（重点谈评分的准确性、评价的即时性、反馈的准确性和个性化、策略指导），用户界面的友好性，以及内容的相关性和适应性。

学生 A：我认为 FiF 的评分准确性需要改进。有时候，我的发音很好，但系统给出的评分却很低。希望能利用更智能的算法来提高评分的准确性，让我们能够更准确地了解自己的水平。

学生 B：我同意 A 的观点。评分的即时性也是我希望改进的地方。有时我完成练习后需要等很长时间才能收到反馈，这让我无法在练习后及时调整。

学生 C：在反馈方面，我认为需要提高反馈的准确性。有时候系统只给出错误提示，但没有详细说明具体问题所在。如果能提供个性化的反馈，那我会更容易理解自己的不足。

学生 D：我也希望能看到策略指导。很多时候，我知道自己有错误，但不清楚如何去纠正。希望 FiF 能给出一些具体的练习建议，帮助我们改进。

学生 E：关于用户界面，我觉得 FiF 的界面虽然简洁，但有时不够直观。在寻找特定功能时，我常常需要花费时间。我希望能改进界面设计，让功能更加易于访问。

学生 F：我同意 E 的看法。用户界面应当更友好，尤其是对于新用户来说。更清晰的导航和指导将大大提升使用体验。

学生 G：我认为内容的相关性也是一个问题。有时候，练习内容与我们课堂学习的内容不太相关，这让我在练习时感到迷茫。希望能够加强内容的适应性，根据我们的学习进度推荐合适的练习。

学生 H：关于内容适应性，我希望 FiF 能根据我的学习水平和需要进行调整。有时我需要更具挑战性的内容，而不是基础的练习，这样才能真正提高我的口语能力。

学生 I：在技术方面，评分的准确性和反馈的及时性都对我的学习体验产生了直接影响。希望系统能利用更先进的技术来提高评分准确性和反馈速度。

学生 J：反馈的准确性让我觉得很重要。有时系统的反馈显得过于笼统，没有针对我的具体错误进行详细分析。如果能针对性地提出建议，将极大帮助我提升口语能力。

学生 K：我也认为反馈的个性化非常关键。希望 FiF 能够根据我以往的表现提供更适合我的反馈，这样我才能更有效地进行针对性练习。

学生 L：我觉得用户界面的友好性是使用 FiF 的一个重要方面。能够快速找到需要的功能，将让我们的学习变得更流畅，而不是因为操作问题而分心。

学生 M：在内容的相关性上，FiF 有时提供的练习与我们的学习目标不符。我希望能增加更多与现实生活相关的情景对话，这样的内容会更实用。

学生 N：我发现技术问题有时会影响用户体验，比如系统的反应速度慢，导致我在完成练习时感到挫败。希望能增强系统的稳定性，让我们可以顺利进行每一次练习。

学生 O：在内容方面，我也希望 FiF 能够根据我们的学习进度推荐更适合的内容。有时我在某个方面的进步比较慢，而系统却没有提供更多练习。

学生 P：对于评分的准确性，我希望 FiF 能在评分系统中引入更多的智能分析，以确保每个学生都能得到公平的评分。

学生 Q：关于反馈的即时性，我觉得系统应该在完成每次练习后，立即给出反馈，这样我才能立刻针对问题进行调整，而不必等待。

学生 R：我认为用户界面的友好性至关重要。简洁明了的界面会让我们在操作时减少困惑，更专注于口语练习。

学生 S：我希望 FiF 的内容能更具多样性。例如，增加与实际交流相关的练习，这样的内容将更具吸引力，并能让我在真实环境中应用所学。

学生 T：在功能方面，策略指导的缺失是最希望改进的地方。如果能在每次反馈中加入具体的练习建议，那将对我提高口语能力非常有帮助。

学生 U：对于内容的相关性和适应性，我希望能根据我的学习情况调整内容。如果能够根据我的弱点推荐适合的练习，那将大大提高我的学习效率。

学生 V：总结来说，我觉得在评分准确性、反馈及时性、用户界面的友好性和内容的适应性方面进行

改进，将会显著提升 FiF 的使用体验。

研究者：非常感谢大家的深入讨论和宝贵建议！你们对 FiF 在功能、用户界面和内容方面的改进建议非常全面。接下来，我们将继续探讨其他话题，以便更好地满足你们的学习需求。

#### 第六部分：用户期望

研究者：现在我们来讨论你们对 FiF 的期望。请每位同学分享一下对 FiF 未来发展的期望、新功能的期望，以及你们对学习支持和资源的需求。

学生 A：我对 FiF 未来发展的期望是能在评分系统上做出改进。我希望未来的版本能使用更先进的技术来提高评分的准确性，帮助我们更好地了解自己的口语能力。

学生 B：我期待 FiF 能增加更多个性化的功能，尤其是在反馈方面。希望系统能根据我的学习进度和表现来定制反馈，这样我能更有效地进行针对性练习。

学生 C：我希望 FiF 能够提供实时的策略指导。通过分析我的录音，系统可以给出具体的改进建议，让我在练习中更有方向感。

学生 D：对我而言，FiF 未来的版本希望能增加社交功能，让我们能与同学和教师更方便地交流。我觉得通过讨论和分享，能够更快地提高我们的口语能力。

学生 E：我期待 FiF 能够提供更多的学习资源，比如视频和音频材料，这些材料能与我们的学习内容直接相关，帮助我们在实际对话中更好地应用所学知识。

学生 F：我希望 FiF 能够引入更多文化背景的学习内容。了解语言背后的文化能帮我更好地理解和使用英语，也能提高我的跨文化交际能力。

学生 G：在新功能方面，我希望能增加一种社区功能，让学生能够互相支持，分享自己的学习经验。这种社区感会激励我们更积极地参与学习。

学生 H：我同样希望未来的 FiF 能够整合更多的学习资源，比如相关的在线课程和学习资料，帮助我们更全面地提高英语口语能力。

学生 I：我期待 FiF 能引入 AI 技术，提供更智能的练习方式。比如通过语音识别技术，能够实时分析我们的发音，并给出即时反馈。

学生 J：在学习支持方面，我希望 FiF 能提供更多的模拟对话练习，让我们在安全的环境中练习实际交流的技巧，这样我们在面对真实对话时会更加自信。

学生 K：我希望未来的 FiF 能增加更多的互动练习，比如小组讨论和角色扮演，这样的活动能让我在实际情境中练习口语。

学生 L：我希望能看到更多的个性化学习路径。根据我的学习情况，系统能推荐适合我的练习内容，这样的适应性会让我感到更加贴心。

学生 M：关于学习支持和资源的需求，我期待能有更多的练习材料，特别是与当前时事和热点话题相关的内容，这样我能在口语表达上更有效率。

学生 N：我对 FiF 未来发展的期望是能够加强与课堂内容的整合，让我们能在学习时更容易找到与课程相关的练习。

学生 O：我希望 FiF 能提供针对特定主题的课程包，让我们可以在需要提高的领域进行集中练习。

学生 P：在新功能方面，我希望能增加更多的反馈细化功能，让我能了解自己在不同方面的表现，例如语法、发音、流利性等，这样我可以更有针对性地进行改进。

学生 Q：我期待 FiF 能增强其移动端功能，让我们能够在手机上进行练习，这样能提高学习的灵活性。

学生 R：我希望未来的 FiF 能够支持更多的语言练习，比如提供多语种的支持，帮助不同背景的学生使用。

学生 S：我期待 FiF 能提供丰富的学习社区资源，比如学习小组和在线辅导，这样我们能在一起交流，互相学习。

学生 T：在学习支持方面，我希望能有专业的指导老师能够在平台上提供在线辅导，帮助我们解决学习中的问题。

学生 U：最后，我希望 FiF 能够不断更新，提供最新的学习资源和内容，这样我们能在不断变化的语言环境中保持学习的热度。

学生 V：总体来说，我对 FiF 的期望是能在功能、内容和学习支持上不断提升，以满足我们不断变化

的学习需求。

研究者：感谢大家的分享！你们对 FiF 未来发展的期望、新功能的期望以及学习支持和资源的需求非常全面。

### Translation of the Summary of Student Focus Group Discussion

#### Part 1: Initial Impressions of FiF

**The researcher:** First of all, let's talk about your initial impressions of FiF. Could each of you share your feelings about using this tool?

**Student A:** My first impression of FiF is that it's very intuitive, and the interface is quite user-friendly. When I first started using it, I found the tool easy to pick up, especially the recording function, which I found quite fun. By listening to my own recordings, I could clearly hear my pronunciation and make adjustments accordingly.

**Student B:** I also thought the interface was very easy to use, the colours and layout were quite appealing. I felt a bit nervous at first, but seeing how user-friendly the interface was really eased my nerves. Overall, I appreciate the flexibility of this tool.

**Student C:** I was pleasantly surprised by FiF's design. Not only is it fully functional, but it also helps me feel relaxed when practising speaking. What I'm most pleased with is that I can practise anytime, anywhere, which is really convenient.

**Student D:** My initial impression is that this tool does provide a good platform, especially since I can record and replay my voice. However, I found that the scoring on the FiF platform is sometimes inaccurate. The system's scores have left me somewhat confused, particularly in terms of pronunciation and fluency.

**Student E:** I also think the feedback function is not always ideal. Sometimes the system doesn't accurately identify errors; for example, I believe my pronunciation is quite standard, but the system indicates mistakes, which frustrates me.

**Student F:** I feel that the feedback sometimes seems generic and lacks specificity. For instance, some of the feedback feels like a fixed template and doesn't consider my particular situation. I'd prefer more personalised feedback that offers suggestions for improving specific errors.

**Student G:** I agree with F. While the feedback is timely, it lacks personalisation and strategic guidance. At times, I need specific advice on how to improve my expression, rather than just being told where I went wrong.

**Student H:** I've also had some negative experiences while using it. Although FiF's interaction function is good, I still feel that face-to-face practice is more effective. Sometimes the automated feedback on FiF doesn't address the problems and troubles I encounter during speaking practice in a timely manner.

**Student I:** Moreover, I sometimes feel that the system's intelligent error correction feature is lacking. Often, I wish for a deeper analysis, such as why a particular sentence structure is problematic or where exactly my pronunciation errors lie.

**Student J:** During my use of FiF, I've found that at times the content isn't entirely relevant. I would like to see more practice exercises directly related to our classroom learning themes rather than those that seem unrelated.

**Student K:** While I thoroughly enjoy using FiF, I also wish the system could provide higher-quality feedback. Sometimes the feedback is overly simplistic, and I want a more in-depth analysis so I can better understand my mistakes.

**Student L:** When I first used this tool, I thought it was very powerful, but at times I found it a bit complex, especially when trying to find specific functions, which required time to get used to.

**Student M:** Overall, I think FiF is a great learning tool, but I've also had some not-so-pleasant experiences. Occasionally, the system's prompts aren't clear enough, causing me to miss problems during recording.

**Student N:** In general, I feel this tool has greatly helped my learning. Although I had some negative

experiences at first, as I used it more, I became more comfortable with the platform and increasingly confident.

**Student O:** My overall experience with FiF has been positive. Although there are some minor technical problems and troubles, its practicality and functionality make it well worth using.

**Student P:** Sometimes, when using it, I feel that practising alone isn't as effective as interacting with classmates in class. While FiF has interaction features, they just don't seem as effective as face-to-face communication.

**Student Q:** I also faced some challenges during the adjustment phase, but through repeated practice, I became more familiar with the tool, particularly with recording and playback.

**Student R:** Overall, I think FiF has provided a positive experience for me. While there are occasional minor problems and troubles, its strengths are clear. I believe that with time, these small problems can be resolved.

**Student S:** I recognise FiF's functionality. Despite encountering some small problems and troubles, its practicality has made me feel that this tool is greatly enhancing my speaking abilities.

**Student T:** I think using FiF has made me more aware of my shortcomings. Although the feedback isn't always detailed enough, the overall experience has been very beneficial for me.

**The researcher:** Thank you all for sharing your experiences! Your feedback is very valuable in helping us improve FiF. Let's move on to discuss your satisfaction with FiF and the importance of its features.

## **Part 2: Satisfaction with FiF**

**The researcher:** Next, let's discuss your overall satisfaction with FiF. Could you each share your reasons for being satisfied or dissatisfied, and compare FiF with other automated English speaking evaluation systems available in the market?

**Student A:** I'm very satisfied with FiF overall. I particularly like the feedback function, which helps me quickly understand my pronunciation and grammar mistakes. This kind of instant feedback is highly effective, especially when I'm practising speaking, as it allows me to correct my mistakes right away.

**Student B:** I'm also very happy with FiF. Compared to other tools on the market, I find its interface much more user-friendly and the operation is simpler. Most speaking evaluation systems have complex interfaces that make it hard to get started, but FiF has done a great job in this regard.

**Student C:** However, I do have a few complaints. For example, sometimes FiF's scoring isn't very accurate. There are times when I feel like I've pronounced things well, but the system gives me a low score. By comparison, some of the other speaking evaluation tools I've used also have scoring problems and troubles, but their feedback is more targeted.

**Student D:** I've had a similar experience. Sometimes FiF doesn't accurately identify errors, which leaves me confused. Additionally, the feedback often feels generic and lacks personalised guidance. I think that in this area, some popular tools in the market, such as certain well-known speaking apps, provide more intelligent correction features that can more specifically point out the source of pronunciation and grammar errors.

**Student E:** I also feel that FiF's feedback lacks personalisation. After each practice session, I do receive feedback, but it's often too vague and not tailored enough to help me truly understand the root cause of my mistakes.

**Student F:** Speaking of dissatisfaction, I'd also like to mention that while FiF's interaction function is good, it sometimes lacks flexibility in real-time discussions and exchanges. Compared to other tools in the market, some applications offer stronger social features, allowing us to communicate and discuss more easily with classmates.

**Student G:** But I still believe FiF is highly beneficial for improving fluency and pronunciation. Despite some shortcomings, my speaking level has definitely improved while using FiF. This improvement is something that other automated evaluation systems can't quite match, especially when it comes to the flexibility of independent practice.

**Student H:** I strongly agree with G's point. Although FiF has its flaws, it's still one of the few tools I've found that can genuinely help improve my spoken expression. While other tools have their own strengths, they often fall short when it comes to the timeliness and effectiveness of feedback compared to FiF.

**Student I:** I think FiF's interaction feature is a highlight, as practising with classmates is a great way to learn. However, as mentioned earlier, some other tools might be better in this aspect. For example, some applications offer more robust real-time discussion features, which make it easier for students to interact and support each other.

**Student J:** I'd also like to mention that FiF lacks strategic guidance, which is something I'm not satisfied with. Many tools on the market provide learning strategies and suggestions to help students solve problems more effectively, but FiF offers relatively less support in this area.

**Student K:** Despite these shortcomings, I think FiF motivates me to study. Its real-time feedback and progress tracking give me more control over my learning. In contrast, some other automated evaluation systems don't offer this kind of motivational mechanism.

**Student L:** I feel that overall, FiF is designed with the user experience in mind. Although it has some technical problems and troubles and lacks in feedback quality, I still think it's better suited to our learning mode than many other popular tools on the market.

**Student M:** I'd like to add that, even though there are many automated evaluation systems on the market, FiF feels more personalised to me. I can find practice methods that suit me, rather than just following the system's steps.

**Student N:** Speaking of personalisation, I hope FiF can improve in this area. While its feedback system meets my needs to some extent, it would be ideal if it could provide customised suggestions based on my learning habits and weak points.

**The researcher:** Your feedback is extremely valuable. From your discussion, I can see that FiF has its strengths, but there are also areas that need improvement. Next, we can continue discussing the importance of FiF's features and suggestions for specific improvements.

### **Part 3: The Importance of FiF's Features**

**The researcher:** Now, let's discuss the importance of FiF's features. Please share your views on the evaluation, feedback, and interaction functions in terms of English speaking learning and assessment, and how these features have influenced the development of your speaking skills.

**Student A:** I believe FiF's evaluation function is very important. It allows me to clearly track my learning progress. After each practice session, the system provides detailed scores, which show how I'm performing in terms of pronunciation, fluency, and grammar. This objective evaluation is really helpful in guiding my learning and showing me which areas I need to improve.

**Student B:** I also place great value on the evaluation function. Not only does it allow me to see my progress, but it also motivates me to set new learning goals. For example, I aim to get a higher score in the next practice session, and this sense of competition encourages me to work harder.

**Student C:** As for the feedback function, I think its usefulness lies in providing instant information. After each practice, I receive detailed feedback, especially on my pronunciation and grammar errors. This timely feedback allows me to quickly adjust my practice strategy, making it highly effective.

**Student D:** The feedback function helps me become aware of my shortcomings. Sometimes, after recording, I'll notice errors that I hadn't realised. The instant feedback allows me to correct these mistakes in the next practice session, which helps improve my speaking ability.

**Student E:** I particularly like the interaction function. Being able to practise with classmates is a great advantage. We can correct each other and share learning experiences, which boosts my confidence in speaking. I've found that through discussing with classmates, I can better understand different ways of expressing myself, and this has greatly helped my learning.

**Student F:** I agree with E. The interaction function plays an important role in facilitating learning. Through real-time exchanges with classmates, we can motivate each other, and the learning atmosphere becomes more dynamic.

**Student G:** The evaluation function helps me set clear learning goals. Every time I see my score, it pushes me to improve, particularly in fluency and grammar expression.

**Student H:** Regarding feedback, I think its effectiveness lies in showing me specific errors and areas for improvement. This targeted feedback allows me to focus on correcting my mistakes in future practice sessions, especially in pronunciation and grammar usage.

**Student I:** The interaction function makes learning more enjoyable for me. By communicating with classmates, I can gain different perspectives that help me understand the course content more comprehensively. This interactive learning approach not only helps me improve my speaking skills faster but also makes learning more fun.

**Student J:** I feel that FiF's evaluation function does more than just provide a score; it sets a standard for my learning. Seeing my scores improve over time boosts my confidence, and this positive reinforcement is very important for me.

**Student K:** The practicality of the feedback function lies in its ability to show me my performance quickly. After each practice, I can see the feedback immediately. This timeliness allows me to correct mistakes quickly and improve the accuracy of my spoken expression.

**Student L:** The interaction function serves as a motivator for me. Through collaborative practice with classmates, I can learn various ways of expressing myself, which gives me more confidence in my speaking.

**Student M:** I believe the evaluation function is critical for my learning development. By setting clear evaluation standards, I can continuously adjust my learning strategy and work towards my goals.

**Student N:** The effectiveness of the feedback function keeps me looking forward to each practice session. After each recording, I'm eager to check the feedback. This feedback makes me feel supported by the system in improving my skills.

**Student O:** I think the interaction function helps build stronger learning relationships. Through group discussions, we not only improve our speaking skills but also deepen our understanding of each other. This learning environment makes me feel very comfortable.

**Student P:** As for the evaluation function, I think it helps me identify my weaknesses. I regularly check my scores and evaluations to understand which areas need strengthening, and then I can create more effective learning plans.

**Student Q:** The instant nature of the feedback function allows me to quickly adjust my learning direction. I'd like to see more specific feedback, especially on the use of complex sentences, which would help me better understand how to improve.

**Student R:** The presence of the interaction function makes me feel supported during learning. Talking with classmates makes my spoken expression more natural, and through discussions, I can keep improving.

**Student S:** I believe that FiF's three main functions complement each other. Without proper evaluation, I wouldn't know my goals; without timely feedback, I wouldn't be able to identify problems; and without interaction, the learning process would become monotonous.

**Student T:** I hope that in the future, FiF will continue to enhance the intelligence of these functions, so that every student can enjoy the best learning experience, further promoting the development of our speaking skills.

**Student U:** Overall, I think FiF performs very well in these three areas. It has helped me make significant progress in my speaking practice and has also made learning the language enjoyable.

**Student V:** I really look forward to seeing FiF improve its evaluation, feedback, and interaction features in future versions, so that we can receive more support and assistance during our learning.

**The researcher:** Thank you all for sharing! Your views on FiF's evaluation, feedback, and interaction functions are very insightful. From your discussions, I can see that FiF has clear strengths, but it also highlights areas for improvement. Next, we can continue discussing specific suggestions for improvements to better meet your learning needs.

#### **Part 4: Problems and troubles Encountered**

**The researcher:** Now let's talk about the problems and troubles you've encountered while using FiF. Could each of you share any technical problems, content challenges, and how these problems and troubles have affected your user experience?

**Student A:** When using FiF, I've encountered some technical problems and troubles, particularly with unstable internet connections. Every time I record, the system occasionally freezes, forcing me to start over. This has been frustrating, especially when I'm in the middle of a practice session. These technical problems and troubles not only slow down my learning progress but also make completing the practice feel incredibly irritating.

**Student B:** I've also experienced problems and troubles with internet delays. Especially during peak times, my recordings often fail to save, which makes me feel like my time has been wasted. These technical problems make me anxious before starting practice because I don't know what might go wrong next.

**Student C:** The technical issue I've encountered mainly involves system crashes. Once, while I was in the middle of an important practice session, the system suddenly shut down, and I had almost finished recording a great session. Not only did I lose the chance to practise, but it also made me doubt the system's stability.

**Student D:** Apart from technical problems and troubles, content challenges are also a big problem. Sometimes, I feel that the exercises provided by the system are not entirely relevant to what I'm learning in class. This mismatch in content leaves me confused during practice and prevents me from fully applying what I've learned.

**Student E:** I agree with D. The practice content on FiF can be dull at times, especially when I want to delve deeper into a specific topic, and the system doesn't provide suitable content. This makes me feel a bit lost during my learning journey.

**Student F:** In terms of content, I believe that a lack of variety is also a challenge. Sometimes, the practice format feels repetitive, which makes learning a bit tedious. I would love to see more scenario-based conversations or group discussion exercises, which would allow me to apply what I've learned in more practical settings.

**Student G:** The technical problems have definitely affected my learning experience. There are times when I have to try multiple times just to get a recording to work, which can be frustrating. This sense of frustration makes me hesitant to use the tool the next time.

**Student H:** I feel that technical problems and troubles have directly affected my confidence. Whenever the system malfunctions, I start to worry that I might be using the tool incorrectly, and this uncertainty makes me feel anxious during practice.

**Student I:** Regarding content relevance, I think FiF should better align its practice exercises with our classroom learning content. Sometimes, the practice topics don't match the knowledge points I've been studying, which makes me feel lost during practice.

**Student J:** I've also encountered challenges with the content. I hope FiF can adjust the difficulty level of its exercises based on our learning progress, so that each person can practise at an appropriate level and avoid ineffective exercises.

**Student K:** As for technical problems and troubles, I find that the system's response speed can sometimes be unsatisfactory. For example, sometimes the feedback takes a long time to load, which affects my learning progress and efficiency.

**Student L:** The repetitive content makes learning feel boring at times. Occasionally, I wish for more

challenging exercises, so I can truly improve my speaking ability.

**Student M:** I feel that technical problems and troubles have affected my motivation to learn. Each time the system crashes, I have to start all over again, which leaves me feeling discouraged and often makes me want to give up on practice.

**Student N:** In terms of content challenges, I've found that some of the exercises are not flexible enough. I wish for more options, allowing me to choose exercises that meet my specific needs.

**Student O:** Regarding technical problems and troubles, I hope FiF will improve system stability in future updates. This way, I can record my practice with confidence, without worrying about technical failures.

**Student P:** As for content relevance, I hope to see more exercises related to real-life conversations and communication. This would better prepare me for real-world English usage scenarios.

**Student Q:** I think that both the technical problems and troubles and content relevance have directly impacted my learning outcomes. Sometimes I spend too much time troubleshooting technical problems, which prevents me from focusing on improving my speaking skills.

**Student R:** I've also noticed that technical problems and troubles have affected the interaction function. For example, during group discussions, if one of the classmates has a poor internet connection, it affects the efficiency of the whole group's communication.

**Student S:** The content challenges leave me feeling confused during certain exercises. At times, I don't even know which areas I need to focus on, which leaves me directionless during practice.

**Student T:** I think the technical problems and troubles with FiF have made me feel less confident during practice. Every time the system fails, I worry about whether I'll be able to complete the practice session successfully.

**Student U:** Regarding content relevance, I hope the system can recommend exercises that match our learning progress. This would help me keep up with the pace of the lessons.

**The researcher:** Thank you for sharing! Your views on the technical problems, content challenges, and their impact on your user experience are very comprehensive. Next, we can discuss specific improvement suggestions to better meet your learning needs.

## **Part 5: Areas for Improvement in FiF**

**The researcher:** Now let's discuss the areas where FiF could improve. Please share your suggestions for improving its features, focusing on the accuracy of scoring, the immediacy of evaluation, the accuracy and personalisation of feedback, strategy guidance, user-friendliness of the interface, and the relevance and adaptability of content.

**Student A:** I think the scoring accuracy of FiF needs improvement. Sometimes, I feel that my pronunciation is quite good, but the system gives me a low score. I hope they can use smarter algorithms to improve the scoring accuracy, so we can better understand our actual level.

**Student B:** I agree with A. The immediacy of the scoring is another area I hope to see improved. Sometimes, after finishing a practice session, I have to wait a long time to receive feedback, which prevents me from adjusting my practice promptly.

**Student C:** Regarding feedback, I think the accuracy needs to be improved. Sometimes the system only gives a general error message without explaining the specific issue. If it could provide more personalised feedback, it would make it easier for me to understand my shortcomings.

**Student D:** I'd also like to see strategy guidance. Often, I know I've made mistakes but don't know how to correct them. I hope FiF can offer specific practice suggestions to help us improve.

**Student E:** As for the user interface, I think FiF's interface is simple, but it's not always intuitive. Sometimes I have to spend time searching for specific functions. I hope the interface design can be improved to make functions more accessible.

**Student F:** I agree with E's view. The user interface should be more user-friendly, especially for new users.

Clearer navigation and guidance would greatly improve the user experience.

**Student G:** I think content relevance is also an issue. Sometimes, the exercises provided don't align with what we are learning in class, which leaves me feeling confused during practice. I hope the content can be more adaptive, offering exercises that are aligned with our learning progress.

**Student H:** Regarding content adaptability, I hope FiF can adjust according to my learning level and needs. Sometimes I need more challenging content rather than basic exercises to truly improve my speaking ability.

**Student I:** On the technical side, both scoring accuracy and feedback timeliness have a direct impact on my learning experience. I hope the system can use more advanced technology to improve both scoring accuracy and feedback speed.

**Student J:** The accuracy of the feedback is important to me. Sometimes the system's feedback is too general and doesn't analyse my specific errors in detail. If it could offer more targeted suggestions, it would significantly help me improve my speaking ability.

**Student K:** I also think personalisation in feedback is key. I hope FiF can provide feedback tailored to my past performances, allowing me to practise more effectively with targeted suggestions.

**Student L:** I believe the user interface's user-friendliness is a crucial aspect of using FiF. Being able to quickly find the necessary functions would make learning smoother, rather than getting distracted by operational problems and troubles.

**Student M:** In terms of content relevance, FiF sometimes provides exercises that don't align with our learning objectives. I hope more real-life conversation scenarios can be included in the content to make the practice more practical.

**Student N:** I've noticed that technical problems and troubles sometimes affect the user experience, such as slow system responses. This makes me feel frustrated while practising. I hope FiF can enhance the system's stability so we can complete each practice session smoothly.

**Student O:** Regarding content, I also hope FiF can recommend exercises that are more suited to our learning progress. At times, I feel I'm progressing slowly in certain areas, but the system doesn't provide enough practice to address those weaknesses.

**Student P:** As for scoring accuracy, I hope FiF can incorporate more intelligent analysis into the scoring system, ensuring that each student receives a fair score.

**Student Q:** When it comes to feedback immediacy, I believe the system should provide feedback immediately after each practice session. This way, I can address my mistakes right away, without having to wait.

**Student R:** I think user-friendliness of the interface is essential. A clear and simple interface would reduce confusion during operation and allow us to focus more on our speaking practice.

**Student S:** I hope FiF's content can be more diverse. For instance, adding more exercises related to real-life communication would make the content more engaging and help me apply what I've learned in real situations.

**Student T:** In terms of functionality, the lack of strategy guidance is the area I most hope to see improved. If FiF could include specific practice suggestions in the feedback, it would greatly help me improve my speaking ability.

**Student U:** Regarding content relevance and adaptability, I hope FiF can adjust the content based on my learning progress. If the system could recommend exercises targeting my weak points, it would greatly enhance my learning efficiency.

**Student V:** To summarise, I think improving the accuracy of scoring, feedback timeliness, user interface friendliness, and content adaptability would significantly enhance the user experience with FiF.

**The researcher:** Thank you all for your insightful and comprehensive suggestions on improving FiF's functionality, user interface, and content. Next, we'll continue discussing other topics to better meet your learning needs.

## **Part 6: User Expectations**

**The researcher:** Now, let's discuss your expectations for FiF in the future. Please share your thoughts on what you hope to see in future developments of FiF, any new features you'd like to have, and what kind of learning support and resources you need.

**Student A:** My expectation for FiF in the future is that it can improve its scoring system. I hope that in future versions, more advanced technology can be used to increase the accuracy of the scoring, helping us better understand our speaking abilities.

**Student B:** I'd like FiF to add more personalised features, especially in terms of feedback. I hope the system can tailor feedback based on my learning progress and performance, allowing me to conduct more targeted practice sessions.

**Student C:** I hope FiF can offer real-time strategy guidance. By analysing my recordings, the system could provide specific suggestions for improvement, giving me more direction in my practice.

**Student D:** For me, I hope future versions of FiF can add social features, allowing us to communicate more easily with classmates and teachers. I think that through discussions and sharing, we can improve our speaking skills more quickly.

**Student E:** I'd like to see FiF offer more learning resources, such as video and audio materials that directly relate to what we're learning. These materials could help us apply what we've learned in real conversations more effectively.

**Student F:** I hope FiF can introduce more cultural background learning content. Understanding the culture behind the language can help me better comprehend and use English, and improve my cross-cultural communication skills.

**Student G:** In terms of new features, I'd like to see FiF add a community feature where students can support each other and share their learning experiences. This sense of community would motivate us to engage more actively in learning.

**Student H:** I also hope future versions of FiF can integrate more learning resources, such as related online courses and study materials, helping us improve our speaking skills more comprehensively.

**Student I:** I expect FiF to introduce AI technology and provide more intelligent practice methods. For example, using speech recognition technology could analyse our pronunciation in real-time and offer instant feedback.

**Student J:** In terms of learning support, I hope FiF can offer more simulated conversation exercises, allowing us to practise real-world communication skills in a safe environment. This would help us feel more confident when facing real-life conversations.

**Student K:** I'd like to see FiF add more interactive exercises in the future, such as group discussions and role-playing activities. These activities would give me more opportunities to practise speaking in real-life situations.

**Student L:** I hope to see more personalised learning paths. Based on my learning progress, the system could recommend exercises that suit me. This adaptability would make me feel like the tool is truly catering to my needs.

**Student M:** In terms of learning support and resources, I'm looking forward to having more practice materials, especially those related to current events and trending topics. This would help me improve my speaking skills with more up-to-date content.

**Student N:** My expectation for FiF's future development is that it can be better integrated with our classroom learning. This would make it easier for us to find practice exercises related to what we're learning in class.

**Student O:** I hope FiF can offer topic-specific course packs that allow us to focus on specific areas where we need to improve.

**Student P:** In terms of new features, I'd like to see more detailed feedback options that help me understand

my performance in different areas, such as grammar, pronunciation, and fluency. This would allow me to make more targeted improvements.

**Student Q:** I'd like FiF to improve its mobile functionality so that we can practise on our phones. This would make learning more flexible and accessible.

**Student R:** I hope future versions of FiF will support more languages, offering multi-lingual practice options to help students from different backgrounds.

**Student S:** I expect FiF to provide more resources for learning communities, such as study groups and online tutoring, where we can collaborate and learn from each other.

**Student T:** In terms of learning support, I'd love to see professional tutors available on the platform who can provide online guidance to help us address any problems and troubles we encounter during our learning journey.

**Student U:** Lastly, I hope FiF continues to update itself regularly, providing the latest learning resources and content so we can keep our learning up-to-date in an ever-changing language environment.

**Student V:** Overall, my expectation for FiF is that it continues to improve its functionality, content, and learning support in order to meet our evolving learning needs.

**The researcher:** Thank you for sharing your thoughts! Your expectations for the future development of FiF, new features, and learning support resources are very comprehensive.

## **Appendix 8 Sample of Participants' Consent Form**

### **An exploratory study on a web-based formative assessment tool for developing EFL learners' oral proficiency at a Chinese higher vocational college**

---

*I would like to invite you to contribute to the above-mentioned academic study as a participant to response to a questionnaire (about 5-10 closed questions) via email or an app. Before you decide whether to take part, please go through the brief description of its goals and procedures detailed below.*

#### **Background to the research**

This research involves an exploratory investigation to learners' perception and impact of a web-based tool used in formative assessment at a higher vocational college. It follows a mixed method approach by investigating EFL learners' perception of web-based technology and finding the evidence of the FIF formative system affecting EFL learners' oral proficiency. The questionnaire is about your attitude towards web-based formative assessment.

#### **What is the purpose of the research?**

The study explores

- Learners' attitude towards web-based formative assessment
- Learners' feelings about applying FIF system
- Impact of FIF system on EFL learners' oral proficiency
- The reliability and validity of FIF system in the formative assessment for oral proficiency

#### **How will my data be used?**

- The study will comply with all standard ethical procedures.
- Interview recordings will be stored securely on a password-protected computer.
- Data will be transcribed and analysed for identifying patterns/themes.
- All participants will be anonymised in analysis and reporting of the data.
- Data will be kept strictly confidential and destroyed after 5 years.

#### **How do I comply with the ethical regulations and guidelines?**

This research involves approximately 100-120 students who are majored in English in grade one at JTC and their data. I ensure that I will conduct the research within the guidelines and recommendations of JTC, and comply with the regulations of UWTSD/UK, BERA ethical guidelines for educational research, fourth edition (2018).

#### **Do I have to take part?**

Participation is voluntary and requires the signing of the attached consent form.

#### **Is there any impact on my academic achievements and awards assessment?**

Your performances and testing results in this study will not influence their academic achievements and

awards assessment.

**How do I discuss my participation or withdraw from the study?**

To discuss any aspect of your participation in this study, please contact me at [309882865@qq.com](mailto:309882865@qq.com).

**Consent Form**

**An exploratory study on a web-based formative assessment tool for  
developing EFL learners' oral proficiency at a Chinese higher vocational  
college**

**Name, position, and contact address of researcher:**

Yixian, Lyu, Vice professor in Foreign Language Development, Jiangxi Teachers College, China; A doctoral student in English Language Education and English Literature, Institute of Education and Humanities, University of Wales Trinity Saint David, UK.  
[309882865@qq.com](mailto:309882865@qq.com)

**Participant Name:**

**Department and Institution:**

**Date of answering:**

**Please initial all boxes**

1. I confirm that I have read and understand the information sheet for the above study and have had the opportunity to ask questions.
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving a reason.
3. I agree to take part in the above study.

4. I agree to answer the questions in the questionnaire.  
☐ Yes ☐ No
5. I agree to the use of anonymised quotations in publications.  
☐ Yes ☐ No
6. I agree that my data gathered in this study may be stored for five years securely (after it has been anonymised) and may be used for future research.  
☐ Yes ☐ No

***Name of Participant***

***Date***

***Signature***

***Name of Researcher***

***Date***

***Signature***

## Appendix 9 Screenshot of the Email Request for Ethical Approval to the JTC Research Department

The screenshot shows a QQ email client interface. The left sidebar contains navigation icons for writing, receiving, and sending emails, as well as a list of folders including '收件箱 (2461)', '星标邮件 (4)', '群邮件 (883)', '草稿箱 (10)', '已发送', '已删除', '垃圾箱 (6)', '我的收藏夹', '其他邮箱', '日历', '记事本', '简历', '发件助手', '在线文档', '每日悦读', '附件收藏', '文件中转站', and '贺卡'. The main content area displays an email dated '28, March, 2022' with the subject 'Applying for ethical approval'. The email is addressed to the 'Director of JTC academic research administrative department'. The body of the email includes a greeting, a statement of the sender's role and purpose, a background to the research, the purpose of the research, the study's focus, how the data will be used, and how the sender complies with ethical regulations. The email concludes with a sign-off and contact information.

28, March, 2022

### Applying for ethical approval

Dear Director of JTC academic research administrative department,

I am a longstanding member of the teaching community at Jiangxi Teachers College, I am currently conducting my doctoral research and I am writing this letter to apply for ethical approval from you. I will introduce my research in brief, and tell you why ethical approval is required, the anticipated risks, and my compliance with ethical standards in detail as below.

#### Background to the research

This research involves a exploratory investigation of learners' perception and impact of a web-based tool used in formative assessment at a higher vocational college. It follows a mixed method approach by investigating EFL learners' perception of web-based technology and finding the evidence of the FIF formative system affects EFL learners' oral proficiency.

#### What is the purpose of the research?

The study explores

- Learners' attitude towards web-based formative assessment
- Learners' feelings about applying FIF system
- Impact of FIF system on EFL learners' oral proficiency
- The reliability and validity of FIF system in the formative assessment for oral proficiency

#### How will my data be used?

- The study will comply with all standard ethical procedures.
- Interview recordings will be stored securely on a password-protected computer.
- Data will be transcribed and analysed for identifying patterns/themes.
- All participants will be anonymised in analysis and reporting of the data.
- Data will be kept strictly confidential and destroyed after 5 years.

#### How do I comply with the ethical regulations and guidelines?

This research involves approximately 100-120 students who are majored in English in grade one at JTC and their data. I ensure that I will conduct the research within the guidelines and recommendations of JTC, and comply with the regulations of UWTSD UK, BERA ethical guidelines for educational research, fourth edition (2018). I will gain consent approval from each participant in a proper way, they have the right to refuse to participate and withdraw from this research at any time. They will be told that their performances and testing results in this study will not influence their academic achievements and awards assessment. I ensure that research participants' data and details are to be kept confidential and that they cannot be disadvantaged by participating in my research. I will behave in a strictly professional manner, remain calm, and be consistent with all groups within normal and usual channels in contacting participants. And all research activities will be conducted during working hours in semi-public places, leaving doors open.

Kind regards.

Yours sincerely, Yixian, Lyu.

Foreign Language Department, Jiangxi Teachers College.

Contact email: [309882865@qq.com](mailto:309882865@qq.com)