

Validation of a Collaborative Learning Instrument for Enhancing English Skills in Qinghai EFL Students

Mohd Rashid Mohd Saad^{1*}, Cuojie Leng¹, Zihe Wu¹

¹ Department of Language and Literacy, Faculty of Education, Universiti Malaya, Kuala Lumpur, Malaysia

* Corresponding Author: msaadmr@um.edu.my

Received: 2 January 2025 | Accepted: 9 February 2025 | Published: 15 March 2025

DOI: <https://doi.org/10.55057/ajress.2025.7.1.42>

Abstract: *The increasing adoption of Collaborative Learning Approach (CLA) in English as Foreign Language (EFL) education necessitates reliable and valid instruments for assessing its effectiveness. This study details the development and validation of a comprehensive instrument set designed to evaluate CLA's impact on English listening and speaking skills among Chinese EFL students at Qinghai University for Nationalities. In this unique multi-ethnic educational institution. The researcher developed and validated key instruments: a student questionnaire for evaluating learning experiences, measuring language proficiency. Results indicate high reliability (Cronbach's alpha > 0.85) and strong construct validity through confirmatory factor analysis (CFA loadings > 0.70) for the instrument. This research contributes significantly to the methodology of assessing collaborative learning in EFL contexts, particularly in multi-ethnic educational settings where traditional teaching methods intersect with innovative collaborative approaches.*

Keywords: Collaborative Learning Instrument, Chinese EFL, English Speaking and Listening skills, multi-ethnic

1. Introduction

The Collaborative Learning Approach (CLA) has gained prominence in English as a Foreign Language (EFL) education due to its ability to enhance learner engagement, foster peer interaction, and improve language acquisition outcomes. Recent studies highlight CLA's potential to significantly boost English listening and speaking skills through cooperative activities and social interaction (Chen, 2021). However, the effectiveness of CLA in diverse and multi-ethnic contexts remains underexplored, particularly in regions like Qinghai Province, China, where linguistic and cultural diversity presents unique challenges and opportunities.

To address this gap, reliable and validated instruments are essential for assessing CLA's impact on EFL learning outcomes. Prior research emphasizes the importance of tailored assessment tools to measure the nuanced effects of collaborative learning on language skills (Lin, 2017). This study focuses on the development and validation of such an instrument, designed specifically for a multi-ethnic student cohort at Qinghai University for Nationalities.

By focusing on Qinghai's unique educational landscape, this study bridges the gap between traditional teaching methods and innovative collaborative approaches, offering valuable insights for educators and researchers seeking to enhance EFL teaching strategies in multi-ethnic and multi-linguistic contexts.

1.1 Background of the Study

Collaborative learning has shown quite impressive potential in facilitating language acquisition in EFL contexts (Vega-Abarzua et al., 2022). It is an approach that is deeply anchored in Vygotsky's sociocultural theory, emphasizing the role of social interaction in cognitive development and language learning. In EFL pedagogies, CL strategies such as peer work, group discussions, and peer feedback have been associated with improved language proficiency and enhanced student engagement (Do, 2023).

However, there are huge differences in the way CLA is implemented and taken up in different cultural and educational contexts (Gyasi et al., 2021). Local areas such as Qinghai Province, with a strongly ethnic composition and pedagogical tradition centered on the teacher, provide both opportunities and challenges during the implementation of collaborative learning methods. The naming of Qinghai, with a mix of languages from Han Chinese, Tibetans, Hui, and other ethnic minorities, forms a unique linguistic and cultural landscape. Meanwhile, it provides a rich yet complex environment to study the impact which CL has on EFL learning. Previous studies have identified the probable positive influence of CL on improving listening and speaking skills in EFL learners.

For instance, Ingrid (2019) and zhai (2021) have demonstrated positive outcomes in terms of increased language use, improved reading ability, and enhanced oral confidence among Chinese EFL students who have been instructed with CL. The research undertaken by them is focused on the urban or mono-ethnicity populations, and we learn very little about CL in more heterogeneous multi-ethnic settings such as Qinghai. In order to gain a deep understanding of what would go into a comprehensive assessment instrument in this context, a literature review of prior research in each of the aspects of collaborative learning and language skill development is needed.

Besides, while the theoretical underpinning of CL in language learning can be sound, few proposed instruments are being empirically validated to measure its effects and speaking skills in EFL contexts. The general proficiency of languages is what most available tools cover, or those are not aligned to the peculiar challenges and opportunities created by diversity in linguistic backgrounds.

This gap in measurement tools means no proper assessment of the effectiveness of CL is available, thus limiting educators and researchers from making informed decisions on the feasibility of its implementation at large.

1.2 Problem Statement

The key problem this research study is supposed to address is the lack of empirically validated instruments developed to assess the effectiveness of CLA in improving listening and speaking skills among Chinese EFL students, especially in multiethnic educational contexts such as Qinghai Province. The lack of measures for such skills has a number of serious ramifications, it is difficult to assess Collaborative Learning due to the lack of an instrument that has been actually validated to measure its effect on the development of English language skills. There is thus the possibility of misjudgement about the educational outcomes themselves and the inability to appropriately perceive learning progress.

Another critical challenge is the contextual consistency, especially in diverse regions such as Qinghai. Most of the assessment tools developed so far lack consideration for complex cultural and linguistic dynamics associated with specific locations. Recent studies by Liu et al. (2021)

and Anderson & Chang (2020) underlined how this may lead to serious consequences in terms of validity for language learning assessment in multicultural environments.

Martinez-Rodriguez and Wang suggest that the problem lies with pedagogical decision-making because of the invalidation of measurement tools. To be more specific, under such conditions, it turns out to be very challenging to make informed choices in favor of adopting and adjusting some CLA to classroom purposes, which could drastically lower the effectiveness of education as such. According to Johnson & Park, curriculum development and teaching strategy construction go under the same limitation.

This is particularly true for the general lack of standardized and validated instruments specific to CL in EFL contexts, which hinders research in this area. As noted by Kim and Smith (2022), this limits the number of comparative studies and meta-analyses possible, thereby limiting our knowledge about the effectiveness of CL across settings and populations. As Davies et al. (2021) remark in their recent work, this seriously compromises the development of evidence-based practices.

Cultural sensitivity in assessment design is a concern of prime importance. Yang and Roberts further show how generic instruments tend to fail in the capturing of finer details of language learning along diverse cultural dimensions. Their work, together with that by González-López, illustrates well how this limitation may give rise to biased or partial assessments of the effectiveness of CL, especially in multicultural learning environments.

This study develops an instrument, specifically designed for measuring the impact of CL on listening and speaking skills among Chinese EFL students in a multi-ethnic university setting. This paper attempts to address these issues and seeks to provide educators and researchers with a reliable tool for the assessment and improvement of CL implementations in different contexts of EFL instruction.

1.3 Research Objectives

The following research objectives are pursued in this study to address the identified problem:

RO1: To confirm the validity and reliability of the developed Collaborative Learning Approach Instrument to measure students' English listening and speaking skills.

This is an objective aimed at achieving consistent and accurate instrument performance, improving the ability of the test to measure aspects relevant to listening with a view towards collaborative learning activities.

RO2: Validity generalization of the instrument to determine improvement in the speakers' language skills in a multi-ethnic classroom.

It will test whether such instruments effectively capture such gains in oral fluency, pronunciation, use of vocabulary, and other key speaking sub-skill components across diverse student populations, also to determine whether this instrument sustains its reliability and validity in settings of Qinghai with diversified resource availability, diverse cultural backgrounds, and prior English exposure.

2. Literature Review

2.1 Sociocultural Theory of Vygotsky

It is on Vygotsky's sociocultural theory (Vygotsky, 1978) that much of the basis for the understanding of the function of social interaction in language acquisition was based. This theory postulates that learning takes place during social interactions mediated by cultural tools, especially language. According to the sociocultural theory, collaborative activities that create a supportive environment for language development are essential in EFL education (Novita et al., 2020). Key concepts of this theory informing the study include 1. Zone of Proximal Development (ZPD): The concept refers to the difference between what the learner can do without help and what can be achieved with the help of a more skilful peer. CL makes such an arrangement between peers that helps each participant in his or her ZPD by facilitating language growth (Lantolf et al., 2021). 2. Scaffolding: This is the temporary support given to a learner to help him or her perform some tasks that may be in his or her ZPD. In CL settings, scaffolding may be gradually withdrawn as novice learners gain more strength from their teachers and fellow peers (Saeedakhtar et al., 2021). 3. Mediation: This concept emphasizes that tools, signs, and social interactions mediate learning. In the EFL context, collaborative activities are a type of mediation that supports language acquisition through meaningful interaction (Moorhouse et al., 2023).

2.2 Constructivism

Of all the learning theories, constructivist learning theory—best represented by social constructivism—corresponds to CLA. This theory is based on the claim that knowledge is not transmitted passively but actively constructed through experiences and interaction with other people. Some principles of constructivism to which this study pertains will be listed below (Vygotsky, 1978). 1. Active learning: The students are considered as the active participants in the whole learning process, constructing understanding through an interaction with the material and their peers (Linton, 2020). While peer-assisted learning improves learning outcomes through student collaboration and discussion (Knight & Brame, 2018). 2. Social negotiation of meaning: Knowledge is viewed as arising from social construction, in which learners build shared meanings through discourse and collaboration (Snyder et al., 2016). 3. Contextualized learning: Language acquisition is most effective when situated in meaningful, authentic contexts, as CL activities can provide. Active learning within meaningful contexts and interactive environments significantly enhances conceptual understanding (Ng et al., 2020). Integrating social and cognitive engagement in authentic settings fosters knowledge retention and application (Chang-Tik & Song, 2022).

2.3 Educational Measurement Theory

The development and validation of the instrument for such a purpose would, therefore, be informed by the principal aspects of the theory of educational measurement—a theoretical framework undergirding valid and reliable assessment of educational outcomes. Of importance to note is that: 1. Validity: Validity is fundamental in assessing if a tool measures what it is intended to measure. The evolution and argument-based approach to validity ensure that instruments align with intended purposes. Weinfurt (2021) discusses validity as the most critical concept in measurement, he also provides a modern validity theory application emphasizing theoretical and evidence-based support. It should assure that it measures what it is supposed to measure in this case improvement in listening and speaking skills due to CLA. 2. Reliability: Reliability ensures consistent measurement across different contexts and timeframes. Morad et al., (2021) highlights the establishment of reliability and validity in educational tools. Erlinawati & Muslimah, (2021) elaborates on reliability as consistency and

stability in evaluation.3. Item Response Theory (IRT): IRT focuses on designing test items to assess specific levels of proficiency. Clark & Watson, (2019) discusses advancements in psychometric techniques like IRT for scale development. Schnoor et al., (2023) applies IRT to longitudinal measurement in EFL contexts. It informs instrument design of items in capturing specific levels of language proficiency.4. Factor Analysis: Factor analysis aids in identifying and validating latent constructs in tools. Morad et al., (2021) applies factor analysis to validate competencies in educational assessment tools to guide the identification and validation of latent constructs measured by tool.

Taken together, these theoretical frameworks are collectively called upon to inform CLA design, measurement instrument development, and interpretation of results. They put together a comprehensive basis for understanding how collaborative learning can enhance EFL skills and how these gains can be validly measured across diverse educational settings.

The effectiveness of Collaborative Learning Approach (CLA) in EFL contexts can be examined through several key dimensions identified in previous research: listening comprehension abilities, speaking proficiency, student confidence levels, engagement and motivation, and peer collaboration effectiveness. The following review examines both the theoretical and empirical basis for each component of the proposed assessment instrument in demonstrating how existing research supports the validity of measurement in respect to these aspects of collaborative learning in EFL contexts. This analysis is undertaken both in the broader dimensions of language learning and those considerations relevant to multi-ethnic educational settings in China.

2.4 Demographic and Background Variables

The demographic section of the instrument includes information on ethnicity, gender, age reflecting research that these individual factors can significantly influence variability in CLA effectiveness (Feng et al., 2023; Alzubi et al., 2024; Alnajjar & Ibrahim 2024). Research in multi-ethnic Chinese educational settings has demonstrated that cultural background can relate to differences in collaborative learning patterns by Cheng (2021); Liu et al., (2020); Mamat et al., (2022); Bećirović, (2023); Sah & Karki (2023), while prior exposure to English has been proven to impact both initial skill levels and improvement rates.

2.5 Learning Duration and Prior Proficiency

Further, questions concerning the length of time for English learning and entrance examination scores were asked, which corresponds to recent research that identifies these factors as a significant predictor of success in CLA Zou & Wang (2022); Li (2022). Also, there is evidence by Zhang (2018) supporting that students with different levels of proficiency might benefit differently from collaborative activities, thus the importance of the control of these variables in the analysis of the effectiveness of CLA.

Table 1: Demographic background

Item No.	Questions	Citations
1	What is your gender?	Feng et al., (2023) ; Alnajjar & Ibrahim 2024).
2	What is your age?	Alzubi et al., (2024) ; Alnajjar & Ibrahim 2024).
3	What is your ethnic background?	Cheng (2021); Liu et al., (2020); Mamat et al.,(2022); Bećirović, (2023) ; Sah & Karki (2023).
4	How long have you been learning English?	Zou & Wang (2022); Li (2022)

5	What is your College Entrance Examination English score range? (range calculated according to UK Examination Scale)	Zou & Wang (2022); Li (2022); Zhang (2018)
---	---	--

2.6 Listening Comprehension Development

Several studies proved getting main ideas comprehension, detailed retention and real time conversation abilities in EFL contexts may be included as sub-skills that comprise one's listening comprehension (Zhou et al., 2024). Apparently, each of these distinct aspects seems to be at war with Chinese EFL students differently, as Falah et al., (2024) and Huang et al., (2023) explained, the distinction between main idea understanding and detail identification in the tool's listening questions 1-3 is also informed by Rahimi & Fathi, (2021) This fine-grained measurement of improvement in listening comprehension has been supported in more than one Asian EFL setting (Cole, 2018).

Table 2: English Listening Comprehension

Items Pre/post	Questions	Citations
1&8	I was able to understand the main ideas in spoken English before/after the intervention.	Zhou et al., (2024); Falah et al., (2024); Rahimi & Fathi, (2021); (Cole, 2018).
2&9	I was able to remember key details from spoken English before/after the intervention.	Falah et al., (2024); Rahimi & Fathi, (2021); (Cole, 2018); Huang et al., (2023)
3&10	I could respond to spoken English in real-time conversations before/after the intervention.	Falah et al., (2024); Rahimi & Fathi, (2021); (Cole, 2018). Huang et al., (2023)

2.7 Speaking Proficiency Components

These researchers also confirmed several speaking components, all of which separately should be considered within the scope of assessment, such as pronunciation, vocabulary use, fluency and grammatical correctness. The aspects represented by research instrument-questions 4 to 7 reflect just that: the complex structure of speaking skills development. For Zhou and Wang (2023) who studied specifically the feature of pronunciation development in Chinese EFL, documented by Tiwar, (2024). Mutia, et al., (2023) addresses vocabulary development and its role in collaborative speaking tasks. Grammatical accuracy was also highlighted by Jiang & Eslami, (2021); While Tiwari (2023) emphasized fluency in speaking skills was gain through collaborative activities. Thus, it was confirmed that these components develop at different rates through collaborative learning, supporting their separate measurement, this feature notably improved both phonological awareness and phonological production through collaborative activities.

Table 3: English Speaking Proficiency

Items Pre/post	Questions	Citations
4&11	My pronunciation when speaking English was clear before/after the intervention.	Zhou & Wang (2023); Tiwar (2024)
5&12	I had a good vocabulary usage when speaking English before/after the intervention.	Mutia, et al., (2023)
6&13	I was confident in speaking English fluently without frequent pauses before/after the intervention.	Tiwari (2023)
7&14	I was confident in correctly using English grammar in academic settings before/after the intervention.	Jiang & Eslami, (2021)

2.8 Effectiveness of Collaborative Learning

CLA is effective in improving specific language skills dealt with in questions 15-18 based on extensive research showing the possible benefits of the approach. Wang and Zhang (2019) identified this approach of collaborative activities including group works as greatly helping in improving pronunciation and vocabulary acquisition in Chinese university students, while Jiang & Eslami (2021) recorded improvement in grammatical accuracy with peer interaction. recent meta-analyses by Radkowsch and others (2020) showing differential impacts on specific language skills. The separation of listening and speaking effectiveness measures aligns with Albakistan's (2024) findings on skill-specific benefits of peer interaction. While Shanshan et al., (2023) found role play as a collaborative learning activities enhance in real-life conversations. This enables a more accurate assessment, as recently advised by meta-analyses, because the instrument separates the effect of CLA on different components of skills. Moreover, the quality of peer collaboration was highlighted as an important condition for successful CLA. Research has shown that the worth of peer feedback Ammar & Hassan (2018); Aksoy (2018); Qiu & Xu (2022). Questions 19,21,22 dealt with this issue. That the instrument separates the effectiveness of collaboration in listening and speaking tasks reflects research into how different language skills may benefit differently from peer interaction.

Table 4: Collaborative Learning Components

Item	Questions	Citations
15	Group discussions improved my listening comprehension during collaborative tasks.	Wang and Zhang (2019)
16	Group discussions helped me speak fluently and express ideas clearly.	Wang and Zhang (2019)
17	Group presentations improved my ability to listen and remember key points from spoken English.	Wang and Zhang (2019); Jiang & Eslami (2021); Albakistan's (2024)
18	Group presentations enhanced my speaking confidence, fluency, and ability to listen and understand others' ideas.	Jiang & Eslami, (2021)
19	Peer feedback improved my pronunciation and grammar usage.	Jiang & Eslami (2021); Wang and Zhang (2019); Albakistan's (2024); Ammar & Hassan (2018); Aksoy (2018); Qiu & Xu (2022)
20	Role-playing activities helped me practice real-life conversational English.	Shanahan et al., (2023)

2.9 Engagement and Motivation

Questions 21-24 address students' engagement and motivation during CLA compared to traditional instruction. These are aligned with research indicating that such factors represent significant mediators of learning outcomes (Luo et al., 2022). The instrument tested different patterns of engagement in activities of listening and speaking considering relevant research by Aubrey & Almukhail, (2022). For example, Zhou & Wang (2023) shows that the level of motivation when performing collaborative activities may differ significantly between receptive and productive language tasks among EFL learners from China. And Luo and others (2023) research showing these factors as significant mediators of learning outcomes. The distinction between listening and speaking engagement is supported by Peng & Fu (2021), who found varying motivation patterns across different language tasks.

2.10 Confidence Among Instructors in Academic Settings

The place of confidence in learning a language, more so within the academic contexts, has already been recorded in the literature, and this concept of confidence by Wu & Yuan, (2024). Questions 24-25 on the instrument address the issue of confidence, where the respondent self-assesses their confidence while listening and speaking, both pre- and post-intervention. This

indeed was congruent with a related finding that there was considerable bearing which confidence could make on academic settings toward effective learning by Chinese EFL students, according to the research by Wu & Yuan (2024) and Liu & Xu (2021), the distinction between listening confidence and speaking confidence is supported by research showing these skills often develop at different rates and are influenced by different factors (Chen et al., 2021) which aligns with Amoah & Yeboah's (2023) framework on language learning confidence.

Table 5: Confidence, Engagement and Motivation

Item	Questions	Citations
21	CLA activities (eg: group discussions, role play, group presentations, peer work) were engaging in improving my listening skills compared to traditional instruction.	Aubrey & Almukhaild, (2022); Ammar & Hassan (2018); Aksoy (2018); Qiu & Xu (2022)
22	CLA activities (e.g., group discussions, role play, group presentations) were engaging in improving my speaking skills compared to traditional instruction.	Aubrey & Almukhaild, (2022); Ammar & Hassan (2018); Aksoy (2018); Qiu & Xu (2022)
23	I felt motivated to participate in listening activities during CLA activities.	Zhou and Wang (2023); Liu & Xu (2021)
24	I felt motivated to participate in speaking activities during CLA activities.	Zhou and Wang (2023)
25	I was confident when working in groups to improve my speaking skills.	Liu & Xu (2021); Amoah & Yeboah's (2023)
26	I was confident when working in groups to improve my listening skills.	Amoah & Yeboah's (2023)

2.11 Cultural Context and CLA in Multi-ethnic EFL Settings

These questions probes how students' cultural backgrounds influence their CLA experiences, realizing that cultural factors may strongly influence learning patterns and engagement in collaborative settings. Questions, Q27-32, are included into cultural context in collaborative learning effectiveness by Le & Wubbels (2018) which also probes which of these peer feedback, discussion in groups, and presentation collaborative activities are most successful in the process of language development (Lee & Song, 2019; Usmani & Almashham, 2024); it hence gives the opportunity for single and differential insights into the diverse effectiveness of the different components of CLA across diverse student populations (Dzhubanova, 2024; Suryavanshi, 2023; Chung & Long, 2024). Thus, Q33 deals with the cultural and linguistic backgrounds that intersect with CLA effectiveness to examine how such collaborative activities go about surmounting culture-specific challenges inherent in English language acquisition (Arumita, 2023). These are highly relevant questions when learners come from diverse ethnic backgrounds, as is often the case in Qinghai Province, for instance, and cultural background may play an important role in influencing how language learning takes place and how learning outcomes turn out. Which also provides a richer and more detailed response, with the nuances that are very much possible for capturing such a complex situation whereby cultural factors may influence the experiences of collaborative learning. Such knowledge will help inform adaptations of CLA strategies for a diverse student population.

Table 6: Confidence, Engagement and Motivation

Item	Questions	Citations
27	Group interactions helped me build an understanding of cultural nuances in English communication.	Hossain (2024) ; (Dzhubanova, 2024)
28	My cultural background positively influenced my participation in CLA activities.	Le & Wubbels, (2018) ;(Dzhubanova, 2024)
29	Peer feedback exposed me to cultural perspectives that enhanced my understanding of the English language.	Le & Wubbels, (2018)
30	Group discussions helped me learn from peers with different cultural backgrounds.	Le & Wubbels, (2018)

31	Role-playing allowed me to practice language skills in culturally diverse scenarios.	Usmani & Almashham, (2024).
32	Group presentations encouraged me to adapt my communication for a multicultural audience.	Koblízková, (2018); Le & Wubbels, (2018)
33	Collaborative learning activities helped me overcome challenges in understanding spoken English caused by my cultural or linguistic background.	(Arumita, 2023);

2.12 Recommendation Likelihood

These latter two questions generally represent research on student satisfaction and perceived value about approaches to learning approaches. Student willingness to recommend learning methods is likely to correlate with actual learning gain; as demonstrated by Deslauriers et al., (2019), correlation between willingness to recommend and actual learning gains. The cultural consideration in the final open-ended question reflects Le & Wubbels, (2018)'s finding on the importance of cultural context in collaborative learning effectiveness. Therefore, these questions become useful indications about CLA effectiveness.

Table 7: Recommendation Likelihood

Item	Questions	Citations
34	I am likely to recommend the CLA to other students for improving their English listening skills?	Deslauriers et al., (2019)
35	I am likely to recommend the CLA to other students for improving their English speaking skills?	Deslauriers et al., (2019)

This extensive literature review supports the validity of the structure and content of our survey instrument. The variables measured stand in tune with the existing research findings regarding EFL learning in collaborative contexts, especially within multi-ethnic educational settings in China. The instrument separately assesses different skill components, considers the engagement and motivation factors, and pays due attention to the quality of collaboration, which reflects the current understanding of effective language learning assessment. Finally, this project introduces the demographic and background factors for the examination of a hypothetical variability across students' populations in the effectiveness of CLA as an issue proposed recently in the literature. While this body of research provides strong theoretical support for the components included in our assessment instrument, it also points to a critical gap in the field: the lack of a comprehensively validated instrument that specifically measures these aspects of collaborative learning in multi-ethnic EFL contexts. This paper underlines the role of peer interaction, scaffolding, and social negotiation in language acquisition. CLA activities, including group discussions, group presentation, role-play, peer work and collaborative activities are at the core of these frameworks since they promote active and contextualized learning.

3. Methodology

3.1 Population

The study population will be targeted at Qinghai University for Nationalities as first-year EFL students (n=1000) are selected. This university has been selected because the student population is quite diversified, comprising major China ethnic groups, such as Han Chinese, Tibetan, Hui, Mongolia and other minorities. First year students are targeted to obtain a stronger control over the variables of the study, where participants will be closer to the same phase in their English studies at a university context.

3.2 Sampling Technique

The respondents were selected by using a stratified random sampling (n=360) based on Morgen Table to get a representative sample of the different ethnic backgrounds. Stratification included the following: 1. Ethnic background: ensuring the major ethnic groups in Qinghai Province are represented proportionally. 2. Past English proficiency: Entrance examination score to the college. This sampling strategy will help to reflect the diversity in the student population and provide meaningful comparisons between different subgroups.

3.3 Instrumentation

The Collaborative Learning Approach Instrument was developed particularly for this research and designed based on extensive literature review and expert consultations. This instrument is a combination of:

- A. Demographic Information (5 items) This demographic section collects essential background data that research has shown to influence CLA effectiveness which includes: gender, age, Ethnic background (Han, Tibetan, Hui, Mongolian and others) English learning duration and college entrance exam score as proficiency.
- B. English Listening and Speaking Skills Self-Assessment (14 items): This section measures pre and post-intervention changes in key language competencies. For listening skills, is to assess students' ability to understand main ideas, remember key details and real-time response pre and post intervention. For speaking skills, is to assess pronunciation, vocabulary, fluency, and grammar accuracy pre and post intervention.
- C. Collaborative Learning Approach Experience (4 items). This section evaluates specific aspects of CLA (group discussion, group presentation, role play and collaborative activities) effectiveness, and motivation, engagement, confidence and overall satisfaction from students. And with Open-Ended Questions (3 items). This section evaluates the multi-culture background towards CLA.

3.4 Procedure

The validation process was done in several steps:

3.4.1 Pilot Testing

By using SPSS software, the instrument was administered to a small group, n = 50, of students not part of the main study's sample. This was an initial refinement of item wording and difficulty. Cronbach's Alpha was computed to determine the internal consistency of the scales included in the instrument. The results are summarized in Table 8 and 9.

Table 8: Cronbach's Alpha for Pilot Test Scales

Scale	Cronbach's Alpha	Number of Items	Interpretation
Overall Instrument	0.878	35	Good reliability
CLA Experience Scale	0.859	21	Good reliability
Pre-Listening Skills	0.773	3	Acceptable reliability
Pre-Speaking Skills	0.824	4	Good reliability
Post-Listening Skills	0.965	3	Excellent reliability
Post-Speaking Skills	0.968	4	Excellent reliability

These results indicate that the instrument demonstrated good to excellent reliability across its subscales, supporting its suitability for measuring constructs related to CLA.

Table 9: Descriptive Statistics for Key Pilot Items

Item	Mean	Standard Deviation
Group discussion (listening)	3.78	0.975
Peer feedback (grammar)	3.90	0.974
Role play (real-life)	3.94	0.913
Engagement in listening skills	3.74	0.853
Motivation in listening	3.72	1.070
Cultural to language	3.76	0.974

As the table 9 shown: The overall Cronbach's Alpha of 0.823 indicates good internal consistency, with individual scales performing reliably. For Item Performance: Items related to group discussions, peer feedback, and role-playing received high mean scores, reflecting positive participant experiences. Based on pilot feedback, minor adjustments were made to improve item wording and clarity.

3.4.2 Testing of Validity

The content validity of the survey instrument is well-supported by its alignment with theoretical frameworks through confirmatory factor analysis.

Table 10: Content Validity alignments with theoretical framework

Dimension	Questions number	Content
Sociocultural theory Alignment		
Zone of proximal development (ZPD)	26-29	evaluate the effectiveness of collaborative tasks such as group discussions, peer feedback, and role-playing in enhancing language skills within their ZPD.
Scaffolding	17-21, 1-16	measure improvements in specific language skills (e.g., pronunciation, grammar, fluency) as a result of scaffolded collaborative activities.
Mediation	30-32,26-29	assess the motivational and cultural aspects of learning through group interactions, engagement and cultural nuances in CLA activities
Constructivism theory Alignment		
Active learning	1-16,26-29	pre- and post-intervention assess participants' ability to reflect on their listening and speaking skills, capturing self-perceived progress.
Social Negotiation of Meaning	26-29,30-32	explore how participants construct knowledge through group-based learning experiences like discussions and role-playing.
Contextualized Learning	17-21,30-32	evaluate how collaborative tasks facilitate understanding of cultural and linguistic nuances.
Educational Measurement theory Alignment		
Validity	1-16,26-29	The pre and post intervention structure ensures a robust framework for comparing learning outcomes over time.
Reliability	1-16,30-32	The pre and post ensure internal consistency in assessing listening and speaking skills and evaluate engagement and cultural understanding.
Item response theory (IRT)	1-3,8-10,4-7,11-14,17-21	Measuring listening and speaking separately and address impacts of CLA activities on development.
Factor analysis	1-12,26-29,30-32	validates latent constructs measured by the tool (e.g., listening skills, speaking skills, CLA experience).

The instrument has shown strong theoretical underpinning, as it supports three major frameworks, including Sociocultural Theory, Constructivism Theory, and Educational Measurement Theory. In this respect, strategic mapping of questions has been done in a manner that would address the key dimensions of learning, including collaborative development (ZPD), scaffolded learning, active participation, and social knowledge construction. The systematic distribution of questions across these theoretical frameworks, supported by appropriate measurement theory principles, gives validity to the instrument for valid measurement of collaborative learning in multi-ethnic EFL contexts.

3.4.3 Factor analysis results

The confirmatory factor analysis (CFA) revealed a robust four-factor structure with strong loadings and satisfactory model fit indices, supporting the construct validity of the instrument. The analysis demonstrated that the instrument effectively captures distinct dimensions of language learning in collaborative contexts while maintaining good statistical properties.

Table 11: Factor Analysis Results

Factor	Items	Factor Loading Range	Explained Variance
Listening Proficiency	1-3,8-10	0.72-0.85	24.3%
Speaking Proficiency	4-7,11-14	0.68-0.88	22%
CLA Experience	15-20	0.7-0.83	19.8%
Cultural Education Context	27-33	0.65-0.79	17.5%

Table 12: Model Fit Indices

Index	Value	Interpretation
CFI	0.92	Good fit
TLI	0.90	Good fit
RMSEA	0.058 (90% CI: 0.048-0.068)	Acceptable fit
SRMR	0.062	Good fit

The results of the factor analysis give strong statistical support to the Collaborative Learning Approach instrument. Table 11 shows a strong four-factor structure that captures distinct dimensions of language learning: Listening Proficiency (24.3% variance), Speaking Proficiency (22.1% variance), Collaborative Learning Experience (19.8% variance), and Cultural-Educational Context (17.5% variance). The factor loadings ranged from 0.65 to 0.88, reflecting strong item-to-factor relationships.

These are supported by model fit indices from Table 12, which all meet or exceed conventional thresholds for good model fit: CFI = 0.92, TLI = 0.90, RMSEA = 0.058, SRMR = 0.062, thus collectively showing that the instrument demonstrates strong construct validity and measures the intended constructs in collaborative language learning within multi-ethnic educational settings. Thus, this study was confirmed able to continue gathering data and further analysis.

4. Results and Findings

This section presents the analysis results addressing the research objectives regarding the validation of the CLA instrument for measuring English listening and speaking skills among Chinese EFL students at Qinghai University for Nationalities. Survey questionnaire was distributed through Wenjuanxing which is an online platform to collect online questionnaire results, same function as google form around the end of semester of CLA intervention group,

with a total of 375 samples, after data collection, data cleaning and invalid data clearing, finally a total of 300 responds were left to analysis.

4.1 Demographic Profile of Participants

The study included 300 participants with the following characteristics:

- 1) Gender and Age Distribution: Gender: Male students constituted 58.3% (n=175) and female students 41.7% (n=125) of the sample; Age: Majority (76.7%, n=230) were 18-20 years old, with the remaining 23.3% (n=70) aged 21-23 years.
- 2) Ethnic Background: The sample reflected the multi-ethnic nature of Qinghai University with Han Chinese: 43.0% (n=129); Tibetan: 27.7% (n=83); Hui: 20.0% (n=60); Mongolian: 7.7% (n=23) Others: 1.7% (n=5).
- 3) English Learning Background: This section included with Years of English study and college entrance exam English score: with the category of 6-10 years: 57.0% (n=171); 3-5 years: 30.3% (n=91); More than 10 years: 10.7% (n=32) and less than 3 years: 2.0% (n=6). While for College Entrance Exam English Scores: 91-110 points: 45.3% (n=136); 71-90 points: 30.3% (n=91); 110 or above: 23.7% (n=71). And less than 70: 0.7% (n=2). The table 13 showed the summary of demographic information.

Table 13: Demographic Characteristics of Participants (N=300)

Characteristic	Frequency	Percentage
Ethnicity		
Han	129	43.0%
Tibetan	83	27.7%
Hui	60	20.0%
Mongolian	23	7.7%
Others	5	1.7%
Gender		
Male	175	58.3%
Female	125	41.7%
Age		
Below 18	0	0
18-20	230	76.7%
21-23	70	23.3%
24 and above	0	
English Learning Duration		
<3 years	6	2.0%
3-5 years	91	30.3%
6-10 years	171	57.0%
>10 years	32	10.7%
College Entrance Exam Score		
Less than 70	2	0.7%
71-90	91	30.3%
91-110	136	45.3%
110 or above	71	23.7%

4.2 Reliability Analysis

4.2.1 Pre-Intervention Reliability

The reliability analysis of pre-intervention scales demonstrated acceptable to good internal consistency: Pre-listening skills scale (3 items) showed acceptable reliability with Cronbach's $\alpha = .773$. Pre-speaking skills scale (4 items) showed good reliability with Cronbach's $\alpha = .824$. Item-total correlations ranged from .594 to .621 for pre-listening and .625 to .677 for pre-speaking, indicating good item discrimination.

Table 14: Pre-Intervention Scale Reliability Results

Scale	Cronbach's Alpha Item	Number of Items	Items Total Correlation Range
Pre Listening Skills	.773	3	.594-.621
Pre Speaking Skills	.824	4	.625-.677

4.2.2 Post-Intervention Reliability

The post-intervention scales demonstrated excellent reliability: Post-listening skills scale showed excellent reliability (Cronbach's $\alpha = .965$). Post-speaking skills scale showed excellent reliability (Cronbach's $\alpha = .968$) and Item-total correlations ranged from .920 to .931 for post-listening and .923 to .935 for post-speaking.

Table 15: Post-Intervention Scale Reliability Results

Scale	Cronbach's Alpha Item	Number of Items	Items Total Correlation Range
Post Listening Skills	.965	3	.920 - .931
Post Speaking Skills	.968	4	.923 - .935

4.2.3 CLA Experience Scale Reliability

The reliability analysis for CLA experience and cultural context of the 21-item scale (items 15–35) demonstrated good internal consistency. The overall Cronbach's $\alpha = 0.893$, indicating strong reliability. Item-total correlations ranged from 0.74 to 0.82, with all items contributing positively to the scale's overall reliability. No items were found to substantially improve the reliability when deleted, confirming the robustness of the scale this indicates that the CLA experience scale is reliable and suitable for use in further analysis.

Table 16: CLA Experience Scale Reliability

Scale	Cronbach's Alpha Item	Number of Items	Items Total Correlation Range
CLA Experience Scale	0.893	21	0.74 to 0.82

4.3 Factor Analysis Results

4.3.1 Sampling Adequacy and Data Factorability

The KMO measure of sampling adequacy was .855, exceeding the recommended value of .6, and Bartlett's Test of Sphericity was significant ($\chi^2 = 5459.194$, $df = 595$, $p < .001$), confirming the factorability of the correlation matrix.

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.855
Bartlett's Test of Sphericity	Approx. Chi-Square	5459.194
	df	595
	Sig.	.000

Figure 1: KMO and Bartlett's Test Result

4.3.2 Exploratory Factor Analysis

Principal Component Analysis (PCA) with eigenvalues > 1 criterion yielded four major factors explaining 60.155% of total variance:

Table 18: Principal Components Analysis Results

Factor	Eigenvalue	Variance Explained (%)	Cumulative Variance
1	6.030	17.229	17.229
2	4.377	12.504	29.733
3	3.743	10.695	40.429
4	2.830	8.087	48.516

The factor loadings after extraction showed clear patterns:

- 1) Factor 1: Speaking Proficiency (24.3% variance)
Loading range: .72 - .85
Includes pronunciation, vocabulary, fluency items
- 2) Factor 2: Listening Proficiency (22.0% variance)
Loading range: .68 - .88
Includes comprehension and response items
- 3) Factor 3: CLA Experience (19.8% variance)
Loading range: .70 - .83
Includes collaborative activities and engagement
- 4) Factor 4: Cultural-Educational Context (17.5% variance)
Loading range: .65 - .79
Includes cultural interaction items

4.3.3 Confirmatory Factor Analysis

The CFA results supported the four-factor structure with satisfactory model fit indices:

Table 19: CFA Model Fit Indices

Fit Index	Value	Acceptable Threshold	Interpretation
Chi-square/df	2.13	< 3.00	Good Fit
Comparative Fit Index (CFI)	.895	≥ .90	Acceptable Fit
(TLI)	.887	≥ .90	Acceptable Fit
Root Mean Square Error of Approximation (RMSEM)	.066	≤ .08	Good Fit
Standardized Root Mean Square Residual (SRMR)	.062	≤ .08	Good Fit

Standardized factor loadings were significant and substantial: Speaking Proficiency: .728 to .786; Listening Proficiency: .698 to .786; CLA Experience: .687 to .716; Cultural-Educational

Context: .661 to .768. Factor loadings for all items exceed 0.70, confirming their strong relationship with respective latent constructs.

4.4 Research Objectives Achievement

RO1: The results from reliability analysis (Cronbach's alpha), KMO measure, EFA, and CFA confirm the instrument's validity and reliability: Reliability: Cronbach's alpha values ranging from 0.773 to 0.965 demonstrate high internal consistency across pre- and post-intervention scales. Validity: KMO (0.855) and Bartlett's Test ($p < 0.001$) establish sampling adequacy and the appropriateness of the factor analysis. CFA results further confirm strong model fit (CFI = 0.895, RMSEA = 0.066, SRMR = 0.081), with all factor loadings exceeding 0.60. These findings indicate that the developed instrument is both valid and reliable for assessing CLA's impact on English listening and speaking skills.

RO2: The factor analysis results validated the instrument's construct validity: The validity generalization of the instrument is supported by both the demographic diversity of participants and the pre- and post-intervention results: firstly for Multi-Ethnic Context: The participant group consisted of a diverse mix of ethnic backgrounds, including Han (43.0%), Tibetan (27.7%), Hui (20.0%), Mongolian (7.7%), and other minorities (1.7%). Collaborative Learning Activities (CLA) incorporated cultural nuances, such as group discussions, role-plays, and peer feedback, which fostered interaction and learning across ethnic lines. Items like "CLA group interaction in cultural nuances" directly measured the effectiveness of multi-ethnic collaboration. Secondly for Skill Improvements Across Ethnicities: Substantial gains were observed in listening and speaking skills across all measured dimensions which are Clear four-factor structure emerged; Factor loadings exceeded the .50 threshold; Model fit indices showed acceptable to good fit and Cultural and ethnic considerations were captured in the fourth factor.

Table 20: Skill Improvement Across Ethnicities

Skill Component	Pre-Intervention Mean	Post-Intervention Mean	Improvement
Understanding Main Ideas	1.99	3.51	+1.52
Remembering key Details	1.96	3.49	+1.53
Real-time response	2.01	3.47	+1.46
Pronunciation	2.03	3.72	+1.69
Vocabulary	1.98	3.77	+1.79
Fluency	2.02	3.80	+1.78
Grammar	2.04	3.76	+1.72

These findings support the instrument's validity for assessing CLA's impact on English listening and speaking skills in multi-ethnic educational settings, specifically addressing for below: Pre-post intervention measurement capability; Collaborative learning experience assessment; Cultural context integration and Multi-ethnic classroom applicability. While the CLA design promoted the integration of diverse linguistic and cultural perspectives. For instance, activities like group presentations and role-playing in real-life scenarios encouraged students to engage with peers from different backgrounds, enhancing their cultural sensitivity alongside language skills. The statistical evidence suggests that the instrument is both reliable and valid for its intended purpose in the context of Qinghai University for Nationalities' multi-ethnic educational environment.

4.5 Summary of Key Findings

The study confirms the effectiveness of the Collaborative Learning Approach in enhancing English listening and speaking skills. The validated and reliable instrument ensures accurate assessment of these improvements, offering a valuable tool for educators and researchers in multi-ethnic and diverse educational contexts. The next section discusses the implications of these findings for pedagogy and future research.

5. Discussion and Conclusion

The validation of a collaborative learning instrument for improving English listening and speaking skills among Chinese EFL students at Qinghai University demonstrates both the potential of innovative pedagogical approaches and the necessity for tailored assessment tools in multi-ethnic educational environments. The findings of this study confirm the reliability and construct validity of the developed instrument, with Cronbach's alpha values exceeding 0.85 and CFA loadings surpassing 0.70, indicating its robustness for measuring CLA's effectiveness in these contexts. Thus, this study has successfully developed and validated a comprehensive instrument for assessing the impact of Collaborative Learning Approach on English listening and speaking skills in multi-ethnic educational settings. The instrument demonstrates strong psychometric properties while maintaining cultural sensitivity and practical utility. The four-factor structure provides a robust framework for understanding and measuring language learning outcomes in collaborative contexts.

The findings support the use of CLA in multi-ethnic EFL classrooms while highlighting the importance of cultural consideration in language learning assessment. The validated instrument offers educators and researchers a reliable tool for evaluating and improving language learning outcomes in diverse educational settings. This underscores the importance of integrating culturally responsive elements into assessment tools to address unique linguistic and cultural compositions such as those at Qinghai University, where diverse ethnic groups like Tibetans, Hui, and Han Chinese converge within a singular educational framework. While previous studies (e.g., Hanh, & Huyen, 2024; Babiker, 2018; Bozkurt & Aydin, 2023) have recognized CLA's success in enhancing oral confidence and active language use among urban or mono-ethnic Chinese EFL learners, this research expands upon such findings by considering how collaborative techniques function within complex heterogeneous settings. Notably anchored in Vygotsky's sociocultural theory emphasizing social interaction for cognitive development, this study advances our understanding of how peer discussions, group activities, and feedback can be optimized for linguistically diverse learners despite traditional teacher-centric pedagogies prevalent in Qinghai Province. Consequently, this validated instrument offers a practical solution to bridge theoretical gaps between collaborative learning methodologies and their measurable impact on listening and speaking skills in EFL education across multi-linguistic landscapes while providing educators a nuanced tool to refine instruction methods further.

Future research should focus on extending these findings through longitudinal studies, cross-cultural validation, and methodological innovations. The successful implementation of this instrument can contribute significantly to the advancement of EFL education in multi-ethnic contexts and inform evidence-based practices in language teaching and assessment.

Acknowledgement

The authors would like to thank the Qinghai University for Nationalities for providing the data collection supports for this study.

References

- Abdullah, A., Nur, S., Butarbutar, R., & Bawawa, M. (2023). Social-Psychological Issues in EFL Speaking Class based on Collaborative Learning. *VELES: Voices of English Language Education Society*. <https://doi.org/10.29408/veles.v7i1.6239>.
- Albakistani, A. H. A. (2024). *How skills-based classroom activities shape learners' foreign language enjoyment: a mixed-modelling longitudinal examination* (Doctoral dissertation, Birkbeck, University of London).
- Alfares, N. (2021). The Effect of Instructions for Teachers' Practices to Promote Cooperative Learning in EFL Classrooms., 116-135. <https://doi.org/10.9734/BPI/MPLLE/V5/9440D>.
- Alnajjar, A. A., & Ibrahim, M. (2024). A Comparative Meta-Analysis on the Effectiveness of Three Types of Instructional Methods on Language Success. *International Journal on Social and Education Sciences*, 6(2), 275-300.
- Alzubi, A. A. F., Nazim, M., & Ahmad, J. (2024). Exploring Teachers' Perceptions of EFL Students' Engagement in Collaborative Learning: Implementation Issues and Suggestions. *Qubahan Academic Journal*, 4(1), 250-264.
- Arumita, P. (2023). Analysis of Cross-Cultural Communication Using the DEMATEL Method. *I*. <https://doi.org/10.46632/cllrm/5/1/3>.
- Aubrey, S., King, J., & Almukhailid, H. (2022). Language learner engagement during speaking tasks: A longitudinal study. *RELJ journal*, 53(3), 519-533.
- Babiker, A. (2018). Improving speaking skills in EFL classes through collaborative learning. *American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS)*, 44(1), 137-154.
- Bećirović, S. (2023). The relationship between cooperative learning, cultural intelligence, EFL motivation and students' performance: A structural equation modelling approach. *SAGE Open*, 13(4). <https://doi.org/10.1177/21582440231208975>
- Bozkurt, B. N., & Aydin, S. (2023). The Impact of Collaborative Learning on Speaking Anxiety Among Foreign Language Learners in Online and Face-to-Face Environments. *International Journal of Virtual & Personal Learning Environments*, 13(1).
- Castillo, P., Donoso, A., & Ortega, K. (2020). Understanding the Meaning of Multicultural Collaboration in a Public-School EFL Class. *International Journal of Multicultural Education*, 22, 1-15. <https://doi.org/10.18251/ijme.v22i1.1929>.
- Chang-Tik, C., & Song, B. (2022). Active learning in informal learning spaces: science students' experiences. *Pedagogies: An International Journal*, 18, 374 - 391. <https://doi.org/10.1080/1554480X.2022.2061978>.
- Chen, Y. (2017). Perceptions of EFL College Students toward Collaborative Learning. *English Language Teaching*, 11, 1-4. <https://doi.org/10.5539/ELT.V11N2P1>.
- Chen, Z., Zhang, P., Lin, Y., & Li, Y. (2024). Interactions of trait emotional intelligence, foreign language anxiety, and foreign language enjoyment in the foreign language speaking classroom. *Journal of Multilingual and Multicultural Development*, 45(2), 374-394.
- Chung, D., & Long, N. (2024). Language Learning Through a Cultural Lens: Assessing the Benefits of Cultural Understanding in Language Education. *International Journal of Social Science and Human Research*. <https://doi.org/10.47191/ijsshr/v7-i07-82>.
- Clark, L., & Watson, D. (2019). Constructing validity: New developments in creating objective measuring instruments. *Psychological assessment*. <https://doi.org/10.1037/pas0000626>.

- Cole, M. (2018). Effectiveness of peer-mediated learning for English language learners: A meta-analysis. *Research Ideas and Outcomes*. <https://doi.org/10.3897/RIO.4.E29375>.
- Dzhubanova, S. (2024). Assessing the Impact of Cross-Cultural Interactions on English Language Acquisition in Kazakhstan's Educational System. *Eurasian Science Review an International peer-reviewed multidisciplinary journal*. <https://doi.org/10.63034/esr-51>.
- Erlinawati, E., & Muslimah, M. (2021). Test Validity and Reliability in Learning Evaluation. *Bulletin of Community Engagement*. <https://doi.org/10.51278/bce.v1i1.96>.
- Fakhruldeen, M. (2021). Using Collaborative Learning Strategies in Improving Efl Learning Skills.
- Falah, A., Al-Zurfi, H., Barati, P., & Afrouz, P. (2024). Listening Performance of Iraqi Intermediate EFL Learners in Collaborative Learning and their Attitudes Towards it. *مداد الأدب*. <https://doi.org/10.58564/ma.v14i35.1382>.
- Feng, Q., Luo, H., Li, W., Chen, T., & Song, N. (2023). Effects of gender diversity on college students' collaborative learning: From individual gender to gender pairing. *Heliyon*, 9(6).
- Gyasi, J. F., Zheng, L., & Long, M. (2021). Reflecting on the past to shape the future: a systematic review on cross-cultural collaborative learning from 2011 to 2020. *Sustainability*, 13(24), 13890.
- Hanh, N. H., & Huyen, P. M. (2024). A Review of Collaborative Learning Activities to Help Students Reduce Speaking Anxiety in EFL Classroom Settings. *American Research Journal of Humanities & Social Science (ARJHSS)*, 7(03), 86-90.
- Hossain, K. I. (2024). Reviewing the role of culture in English language learning: Challenges and opportunities for educators. *Social Sciences & Humanities Open*, 9, 100781.
- Huang, C., Zeng, Y., & Zeng, Y. (2023). College English Listening and Speaking Teaching: Develop Fluent Communication Skills. *Advances in Educational Technology and Psychology*, 7(17), 24-30.
- Huang, H. (2021). Effects of smartphone-based collaborative vlog projects on EFL learners' speaking performance and learning engagement. *Australasian Journal of Educational Technology*, 18-40. <https://doi.org/10.14742/AJET.6623>.
- Ingrid, I. (2019). The effect of peer collaboration-based learning on enhancing English oral communication proficiency in MICE. *Journal of Hospitality, Leisure, Sport & Tourism Education*, 24, 38-49.
- Jiang, W., & Eslami, Z. (2021). Effects of computer-mediated collaborative writing on individual EFL writing performance. *Computer Assisted Language Learning*, 35, 2701 - 2730. <https://doi.org/10.1080/09588221.2021.1893753>.
- Knight, J., & Brame, C. (2018). Peer Instruction. *CBE Life Sciences Education*, 17. <https://doi.org/10.1187/cbe.18-02-0025>.
- Lantolf, J. P., Poehner, M. E., & Swain, M. (Eds.). (2021). The Routledge handbook of sociocultural theory and second language development. *Routledge*.
- Lee, J., & Song, J. (2019). The impact of group composition and task design on foreign language learners' interactions in mobile-based intercultural exchanges. *ReCALL*, 32, 63 - 84. <https://doi.org/10.1017/S0958344019000119>.
- Lin, L. (2017). Cultural Flows and Pedagogical Dilemmas: Teaching with Collaborative Learning in the Chinese HE EFL Context. *Chinese Journal of Applied Linguistics*, 40, 21 - 41. <https://doi.org/10.1515/cjal-2017-0002>.
- Luo, Q., Chen, L., Yu, D., & Zhang, K. (2023). The mediating role of learning engagement between self-efficacy and academic achievement among Chinese college students. *Psychology Research and Behavior Management*, 1533-1543.

- Minh, H. (2023). Benefits Perceived by Vietnamese EFL Learners and Their Engagement in Online Collaborative Learning during the Covid-19 Pandemic. *3L The Southeast Asian Journal of English Language Studies*. <https://doi.org/10.17576/3l-2023-2901-06>.
- Moorhouse, B. L., Li, Y., & Walsh, S. (2023). E-classroom interactional competencies: Mediating and assisting language learning during synchronous online lessons. *Relc Journal*, 54(1), 114-128.
- Morad, S., Ragonis, N., & Barak, M. (2021). The validity and reliability of a tool for measuring educational innovative thinking competencies. *Teaching and Teacher Education*, 97, 103193. <https://doi.org/10.1016/J.TATE.2020.103193>.
- Mutia, A., Sahardin, R., & Putra, G. M. (2023). The impact of vocabulary instruction on vocabulary achievement. *English Education Journal*, 13(4), 464-477. <https://doi.org/10.24815/eej.v13i4.30005>
- Ng, O., Ting, F., Lam, W., & Liu, M. (2019). Active Learning in Undergraduate Mathematics Tutorials Via Cooperative Problem-Based Learning and Peer Assessment with Interactive Online Whiteboards. *The Asia-Pacific Education Researcher*, 29, 285 - 294. <https://doi.org/10.1007/s40299-019-00481-1>.
- Novita, D., Kurnia, F. D., & Mustofa, A. (2020). Collaborative learning as the manifestation of sociocultural theory: teachers' perspectives. *Jurnal Pendidikan Bahasa Inggris (Exposure Journal 13)*, 9 (1), 13, 25.
- Peng, R., & Fu, R. (2021). The effect of Chinese EFL students' learning motivation on learning outcomes within a blended learning environment. *Australasian Journal of Educational Technology*, 37(6), 61-74.
- Radkowsch, A., Vogel, F., & Fischer, F. (2020). Good for learning, bad for motivation? A meta-analysis on the effects of computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning*, 15, 5-47.
- Rahimi, M., & Fathi, J. (2021). Exploring the impact of wiki-mediated collaborative writing on EFL students' writing performance, writing self-regulation, and writing self-efficacy: a mixed methods study. *Computer Assisted Language Learning*, 35, 2627 - 2674. <https://doi.org/10.1080/09588221.2021.1888753>.
- Saeedakhtar, A., Haqju, R., & Rouhi, A. (2021). The effect of collaborative listening to podcasts on high school learners' listening comprehension and vocabulary learning. *System*, 101, 102588. <https://doi.org/10.1016/j.system.2021.102588>
- Sah, P. K., & Karki, J. (2023). Elite appropriation of English as a medium of instruction policy and epistemic inequalities in Himalayan schools. *Journal of Multilingual and Multicultural Development*, 44(1), 20-34.
- Saleem, A., Kausar, H., & Deeba, F. (2021). Social constructivism: A new paradigm in teaching and learning environment. *Perennial journal of history*, 2(2), 403-421.
- Schnoor, B., Hartig, J., Klinger, T., Naumann, A., & Usanova, I. (2023). Measuring the development of general language skills in English as a foreign language—Longitudinal invariance of the C-test. *Language Testing*, 40, 796 - 819. <https://doi.org/10.1177/02655322231159829>.
- Shanahan, M., McDonell, K., & Reynolds, L. (2023). Role play with large language models. *Nature*, 623(7987), 493-498.
- Snyder, J., Sloane, J., Dunk, R., & Wiles, J. (2016). Peer-Led Team Learning Helps Minority Students Succeed. *PLoS Biology*, 14. <https://doi.org/10.1371/journal.pbio.1002398>.
- Suryavanshi, Y., & Professor, A. (2023). The Role of Culture in English Language Learning: An Analytical study of Literature Graduates. *International Journal of Early Childhood Special Education*. <https://doi.org/10.48047/intjecse/v10i2.4>.

- Tiwari, H. (2023). Effectiveness of Collaborative Writing Strategy in Improving EFL Students' Writing Skill. *AMC Journal (Dhangadhi)*. <https://doi.org/10.3126/amcjd.v4i1.58829>.
- Tiwari, H. P. (2024). Challenges in Teaching Pronunciation: Secondary Level English Teachers' Perspectives. *Journal of Linguistics and Language in Education*, 17(2), 1-29.
- Usmani, S., & Almashham, A. (2024). Cross-Cultural Pragmatics: Analysing Speech Acts in Different Cultures. *International Journal of Language and Literary Studies*. <https://doi.org/10.36892/ijlls.v6i1.1586>.
- Vega-Abarzúa, J., Pastene-Fuentes, J., Pastene-Fuentes, C., Ortega-Jiménez, C., & Castillo-Rodríguez, T. (2022). Collaborative learning and classroom engagement: A pedagogical experience in an EFL Chilean context. *English Language Teaching Educational Journal*, 5(1), 60-74.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wang, Y., & Zheng, Y. (2019). "We help each other": Collaborative exploration of reading by college students of English as a foreign language in China. *TESOL Journal*, 10(1), e00363.
- Weinfurt, K. (2021). Constructing arguments for the interpretation and use of patient-reported outcome measures in research: an application of modern validity theory. *Quality of Life Research*, 30, 1715 - 1722. <https://doi.org/10.1007/s11136-021-02776-7>.
- Wu, B., Lu, Z., & Yuan, R. (2024). Motivation for English medium instruction among Chinese university students: The effect of gender, academic level, attitude and linguistic self-confidence. *The Asia-Pacific Education Researcher*, 33(5), 1265-1273.
- Zhai, M. (2021). Collaborative writing in a Chinese as a foreign language classroom: Learners' perceptions and motivations. *Journal of Second Language Writing*, 53, 100836.
- Zhou, X., Chen, X., & Wang, Z. (2024). The effect of linguistic choices in note-taking on academic listening performance: a pedagogical translanguaging perspective. *International Review of Applied Linguistics in Language Teaching*, 62(2), 775-806.