Malaysian Journal of Social Sciences and Humanities (MJSSH) (e-ISSN : 2504-8562)
2025, Volume 10, Issue 4, e003338
DOI: https://doi.org/10.47405/mjssh.v10i4.3338

RESEARCH ARTICLE

Enhancing Teacher Self-Assessment Instruments for Evaluating Collaborative Learning in Multicultural EFL Classrooms

Mohd Rashid Bin Mohd Saad[1*], Cuojie Leng[2], Zihe Wu[3], Jingjing Wang[4]

[1]Department of Language and Literacy, Faculty of Education, Universiti Malaya, 50603, Kuala Lumpur, Malaysia
Email: msaadmr@um.edu.my
[2]Department of Language and Literacy, Faculty of Education, Universiti Malaya, 50603, Kuala Lumpur, Malaysia
Email: s2119313@siswa.um.edu.my
[3]Department of Language and Literacy, Faculty of Education, Universiti Malaya, 50603, Kuala Lumpur, Malaysia
Email: s2189469@siswa.um.edu.my
[4]Department of Language and Literacy, Faculty of Education, Universiti Malaya, 50603, Kuala Lumpur, Malaysia
Email: s2129034@siswa.um.edu.my

CORRESPONDING AUTHOR (*):
Mohd Rashid Bin Mohd Saad (msaadmr@um.edu.my)

KEYWORDS:
Collaborative learning
EFL classrooms
Teacher self-assessment
Multicultural education
Culturally responsive teaching

CITATION:
Mohd Rashid Mohd Saad, Leng, C., Wu, Z., & Wang, J. (2025). Enhancing Teacher Self-Assessment Instruments for Evaluating Collaborative Learning in Multicultural EFL Classrooms. *Malaysian Journal of Social Sciences and Humanities (MJSSH), 10*(4), e003338. https://doi.org/10.47405/mjssh.v10i4.3338

ABSTRACT

This study addresses the need for culturally sensitive teacher evaluation instruments in multicultural English as a Foreign Language (EFL) classrooms in Qinghai Province, China, where diverse ethnic groups (Han Chinese, Tibetan, Hui, and Mongolian) learn together. The research employed a quantitative research method to design and validate a teacher self-assessment checklist for evaluating collaborative learning approaches in these settings. Grounded in Classical Reliability Theory, Transformational Leadership Theory, and Intercultural Communication Theory, the 28-item instrument encompasses six dimensions: CLA Understanding and Teacher Preparedness, Classroom Implementation, Student Engagement and Feedback, Adaptation and Reflection, Professional Development, and Challenges and Solutions. Data collected from 300 EFL teachers through stratified random sampling were analyzed using Cronbach's alpha and Confirmatory Factor Analysis (CFA). Results demonstrated excellent internal consistency reliability (α = 0.92) and strong construct validity, with CFA fit indices confirming the proposed six-factor structure. All factors showed satisfactory convergent and discriminant validity. The validated checklist addresses a critical gap in existing evaluation practices and contributes to multicultural education by providing a reliable instrument that positions cultural diversity as an asset rather than a challenge in EFL education, while promoting teacher reflection and facilitating professional development in multicultural collaborative learning environments.

Contribution/Originality: This study contributes to multicultural EFL education by

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

developing and validating one of the first teacher self-assessment instruments specifically designed for collaborative learning in ethnically diverse classrooms. The research provides a psychometrically sound tool that positions cultural diversity as an asset rather than a challenge in language education.

1. Introduction

Teacher assessment instruments have a significant role in ascertaining the efficacy of EFL teaching methodologies. In the last few years, the Collaborative Learning Approach (CLA) gained much popularity in English as a Foreign Language (EFL) settings due to its potential to enhance participation, promote learner interaction, and enhance language acquisition through authentic communication. CLA, with its learners working collaboratively in groups of two or a small group to achieve common learning goals, is a shift of tempo away from the traditional teacher-centric teaching to more learner-centric, interactive learning spaces (Alzubi et al., 2024).

Within multicultural settings, the use of CLA presents unique opportunities along with unique challenges. Such multicultural learning settings exist in locations like Qinghai Province in China, with the classrooms having a mix of different ethnic groups like Han Chinese, Tibetans, Hui, and Mongolians as students. Such a mix of cultures enriches the learning environment but also calls for culturally responsive teaching and assessment tools that acknowledge and adjust to diverse communications styles, learning styles, and cultural values (Liao & Li, 2023).

Traditional approaches to CLA implementation tend to assume cultural homogeneity, oblivious to the impact of cultural diversity on cooperative dynamics. Eliyahu, (2023) identifies that collaborative learning could be experienced differently by students of various cultures, with cultures valuing individual achievement and others valuing group harmony and group achievement. note that collaborative learning exercises that do not consider cultural differences have the potential to marginalize students with different cultural norms than the dominant educational paradigm.

Despite the extensive use of CLA in multicultural EFL classrooms, a troubling shortage of reliable and culturally responsive instruments to gauge its implementation persists. Such a shortage raises several related issues:

Firstly, traditional teacher evaluation instruments fail to consider the complex dynamic between collaborative learning approaches and culture. Such instruments usually measure teaching practices in terms of standard criteria that do not always account for the necessary accommodations to effectively apply CLA to diverse cultures. Such a mismatch, as argued by Snyder and Fenner (2021) could lead to the erroneous measurement of teacher effectiveness and could also discourage culturally responsive adaptations of collaborative approaches.

Secondly, without culturally responsive assessment tools, teachers in multicultural environments like Qinghai Province lack guided directions to consider and promote the practice of CLA. Yildirm, (2019) observe that teachers need to have models of evaluation that address specifically how to deal with cultural differences in collaborative environments, manage various communication styles, and achieve inclusive participation among students with varying ethnic backgrounds.

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

Thirdly, the unavailability of reliable self-assessment instruments bars systemized professional development to enhance CLA in multicultural settings. Franco et al. (2024) note, professional development requires validated instruments to assist teachers in identifying accurate areas of need in the application of culturally responsive collaborative practices.

These issues collectively reveal an urgent need in the field of educational assessment: the lack of validated, culturally responsive self-assessment tools for teachers implementing CLA in multicultural EFL classrooms (Boruah, 2022). Such a need is particularly urgent in regions like Qinghai Province, where the learners' cultural heterogeneity calls for thoughtful accommodations of collaborative learning methods.

This study fills this gap by developing and piloting a teacher self-assessment checklist intended to address implementation of CLA in multicultural EFL classrooms. By including considerations of culture on multiple dimensions of implementation of collaborative learning, this checklist aims to provide a tool for measurement that is responsive to the complex interplay between collaborative dynamics and culture.

The objectives of this research are:

i. To design a teacher self-assessment checklist for evaluating the implementation of collaborative learning approaches in multicultural EFL classrooms.
ii. To evaluate the reliability and validity of the self-assessment checklist.
iii. To provide recommendations for the effective use of the checklist in professional development and teacher reflection.

2. Literature Review

2.1. Theoretical Framework

The study is grounded in the classical theory of dependability with an emphasis on: Internal measurement consistency, Inter-rater dependability criteria and Consistency of the measurements in different contexts.

2.1.1. Reliability Theory in Educational Assessment

The study is grounded in classical theory of reliability, the cornerstone of the understanding of the fundamental principles of educational measurement. At the heart of the theory of reliability are key concepts that enhance the quality of educational tests and facilitate more precise interpretations of pupil performance (Adams, 2005).

Internal Consistency of Measurement: As Meyer, (2010) stated that Internal consistency is the level to which different items or tasks within an evaluation measure the same construct. Internal consistency is vital in ensuring that measurements yield reliable and consistent results. In the case of EFL learning, highly internally consistent tests have the potential to better measure the language understanding and teamwork competencies of the students, thereby enabling the teachers to develop more effective instructional methods that are culturally responsive and inclusive.

Inter-rater Standards of Reliability: When more than one rater is involved in the evaluation process, inter-rater reliability is required. Inter-rater reliability determines the extent to which different raters give the same performance the same rating. Inter-

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

rater reliability is particularly critical in multicultural classrooms, as language proficiency among the students with varying backgrounds may vary significantly. Adhering to strict inter-rater reliability standards ensures that collaborative learning activity assessments remain unbiased and fair, and thus more reliable in general (Livingston & Carlson, 2018).

Stability Across Contexts: Stability is the level to which the results of the evaluation do not change over a period and across different contexts. Stability in EFL contexts is necessary to measure the language learning and collaboration skills of the learners in a holistic manner. With an effective reliability model, instructors can design assessments that remain valid and usable regardless of the language or culture environment, thus promoting an inclusive evaluative system.

2.1.2. Transformational Leadership Theory

Supermane et al. (2018) finds out transformational leadership theory emerges as a key model for facilitating innovation and inclusivity in educational assessment practice. The theory centers on the teacher as a change agent who brings about change in the classroom by working collaboratively and being responsive to other people's needs.

Teacher Leadership in the Facilitating of Collaborative Learning: Transformational leadership requires teachers to take the lead in facilitating collaborative learning environments. Not only do they facilitate group working among students, but they also coach other teachers in culturally responsive assessment. When teachers exercise transformational leadership, they are able to effectively facilitate a culture that is inclusive of diversity and centered on the group's success (Ogbonnaya, & Chukwudebelu, 2020).

Adaptive Assessment Strategies: At the heart of transformational leadership is the assumption that the strategies of the leader must adapt to the needs of the people being led (Kwan, 2020). In multicultural classrooms, this means that the assessment practices must also adapt to fit different cultural orientations and learning styles. With the implementation of adaptive assessment strategies, teachers can enhance participation and foster a learning culture that embraces and understands all students.

Cultural Sensitivity in the Evaluation Process: Transformational leaders promote the culture of cultural sensitivity within the evaluation process. Through the identification and integration of students' different cultural backgrounds into the evaluation tools, teachers can ensure that evaluations are equitable and reflective of the unique conditions that guide students' learning (Van Knippenberg, 2013). Such correspondence guarantees the validation of different teaching methods, thereby enhancing the educational texture of EFL classrooms.

2.1.3. Intercultural Communication Theory

Intercultural communication theory offers profound insight into the dynamics of communication in multicultural classrooms. Strategies of effective communication that honor and celebrate the differences of cultures are the central focus of intercultural communication theory.

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

Cross-Cultural Communication Patterns: The theory calls for the identification of various patterns of communication that may differ across cultures. Since EFL settings have students with a wide variety of backgrounds, teachers must have the competence to identify such differences to facilitate effective collaborative learning (Csillik, 2022). Employing methods that cater to various patterns of communication, teachers can make learning more inclusive and exciting.

Intercultural communication theory also verifies the necessity of cultural sensitivity in the process of assessment. Mahon and Cushner (2020), cultural differences based assessments have the potential to result in misconceptions regarding the capabilities and competencies of students. Cultural awareness in evaluation models ensures that the assessments match the realities and environments of the students, therefore yielding more authentic and credible assessments.

Multicultural Classroom Dynamics: The above theoretical model also addresses the complexities of multicultural classroom dynamics. Teachers with a grasp of the dynamics of the classroom will more likely use assessments that address the different needs of the students to have all voices represented and heard. Not only is it more effective in collaborative learning but also in fostering a sense of belonging among students of diverse backgrounds.

2.2. Current Evaluation Practices

There are some researches describes some of the key themes that characterize widespread evaluation practice in multicultural EFL contexts, and in particular in regions like Qinghai Province.

Traditional Assessment Procedures Lack Sensitivity to Culture. Most studies establish that traditional assessment procedures lack sensitivity to diversity in culture and make unjust assumptions regarding students' performance (Anderson et al., 2022). Traditional practices exclude nonconventional students with alternative cultures. Failure of such assessment procedures to consider culture reduces the efficiency of the assessment process to uncover true EFL learners' abilities and requires a transition to more inclusive and culturally responsive assessment tools (Awad et al., 2016; Nortvedt et al., 2020).

Limited Account for the Dynamics of Collaborative Learning: Most of the traditional methods of assessment fail to account for the fact that learning is a collaborative endeavor. As collaborative learning is increasingly recognized as a more desirable pedagogy, it is important that methods of assessment account for the collaborative dynamics of learning. Traditional methods of assessment are more focused on individual performance without regard for shared contributions and participation that is typical of collaborative learning.

Insufficiency of Proper Reliability Measures in Multicultural Contexts: Conventional tests' effectiveness is also undermined by the insufficiency of proper reliability measures in multicultural contexts (Byram, 2020). Most of the existing instruments do not take into account the heterogeneity of the students' experiences and capabilities, leading to inconsistency and unfair judgments. Through the harmonization of assessment mechanisms with the heterogeneous cultural contexts of the students, the

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

judgments' reliability could be enhanced so that they give a true reflection of the quality of collaborative learning.

Ultimately, as the educational future of EFL teaching moves forward, the incorporation of culturally responsive models of evaluation that have a sound theoretical foundation is most important. Through the aligning of evaluation practice with multiple environments of culture, instructors can increase the validity and justice of assessments as well as to provide inclusive learning environments in which all learners have the chance to succeed. Through this literature review, the need for future research and development of evaluation methods that honor cultural diversity as a strength becomes evident, ultimately optimizing the effectiveness of collaborative learning in multicultural settings (Ghavifekr, 2020).

3. Methodology

The methodology focuses on developing and testing a teacher self-assessment checklist. Development of the checklist was done through a meticulous examination of CLA principles and was grounded on criteria such as classroom dynamics in multicultural classrooms, teacher training, and student engagement. Items were constructed to cover various dimensions of CLA such as teacher training, student engagement, and application of collaborative activities. Validation of the checklist involved obtaining responses from seasoned teachers with multicultural classroom experience and a pilot test to establish reliability and clarity.

3.1. Research Design

The study employed a quantitative research approach to establish the teacher self-assessment checklist that was created to assess CLA in multicultural EFL classrooms was reliable and valid. This was to ensure that the instrument produces consistent and authentic results when used by various teachers in various classroom settings. The objective of this quantitative study was to rigorously test the efficacy of the checklist and whether it can measure constructs of CLA validly and provide solid data to support its possible application in classrooms.

This study was conducted in Qinghai Province, China, this was selected due to its rich ethnic and cultural diversity. This province is home to multiple ethnic groups including Han Chinese, Tibetan, Hui, and Mongolian communities, making it an ideal setting for investigating multicultural classroom dynamics. Educational institutions in this region provided an authentic context for studying collaborative learning approaches in classrooms with diverse student compositions, as teachers in this region regularly navigate intercultural communication challenges while implementing collaborative learning strategies.

3.2. Sampling and Procedure

The target population comprised EFL teachers working in higher education institutions across Qinghai Province. Based on provincial educational statistics, approximately 1300 EFL teachers were employed across various institution types in the region at the time of data collection. These included public universities, private universities, community colleges, vocational schools, and language training centers. A stratified random sampling technique was employed to ensure representative distribution across key variables in

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

Qinghai Province's EFL teacher population. The process involved two-stage stratification: first by institutional type (public universities, private universities, vocational schools, community colleges, and language training centers), and second within each institution by demographic characteristics including gender, age ranges, ethnic background (Han Chinese, Tibetan, Hui, Mongolian, and others), and teaching experience levels. This comprehensive stratification approach ensured proportional representation of the diverse teacher population, significantly enhancing the generalizability of the study's finding.

Data collection was conducted over three months (September-November 2024), beginning with securing institutional approvals from participating higher education institutions. The self-checklist questionnaire was digitized on a secure online platform and distributed through institutional WeChat group systems and professional EFL teaching associations in Qinghai Province. Participants had 2 weeks to complete the survey. From 342 distributed questionnaires, 300 complete responses were received (87.7% response rate), with incomplete questionnaires (missing >10% of items) excluded from analysis.

Data analysis employed multiple statistical techniques, including descriptive statistics for demographic variables and checklist items, Cronbach's alpha coefficient to assess internal consistency reliability (threshold $\alpha \geq 0.70$), and Confirmatory Factor Analysis (CFA) using AMOS 28.0 to evaluate construct validity through multiple fit indices (Chi-square/df ratio, CFI, TLI, RMSEA, SRMR). Convergent validity was assessed through factor loadings and Average Variance Extracted, while discriminant validity was evaluated by comparing the square root of AVE with inter-factor correlations, with all statistical analyses conducted using SPSS 28.0 and AMOS 28.0 software packages.

3.3. Instrumentation

3.3.1. Components of the Teacher Evaluation Instrument

The study instrumentation was designed with care to elicit rich data on the implementation of CLA in multicultural EFL classrooms. The revised teacher evaluation instrument has two main sections: see Table 1: Demographic information (11 items) and Table 2: Teacher Self-Assessment Checklist (28 items). Demographic information collects important background information on the teachers, including gender, age, ethnic origin, education background, experience in teaching, institution setting, and classroom profile. These variables provide important context for interpreting the different settings in which CLA is being implemented and how teacher characteristics might influence implementation practices. The Teacher Self-Assessment Checklist, organized around six different sections (A-F), captures a wide range of CLA implementation areas: understanding and teacher preparedness, classroom implementation, student engagement and student feedback, adaptation and reflection, teacher development and ongoing improvement, and challenges and solutions. Each was designed with care using existing theoretical frameworks including Classical Reliability Theory, Transformational Leadership Theory, and Intercultural Communication Theory to ensure content validity. Items were worded using a 5-point Likert scale to measure frequency, extent or degree of agreement with statements on CLA practices to offer room for fine-grained quantitative analysis. This two-part instrument allows for rich exploration of the complex interplay between teacher characteristics, classroom contexts, and CLA implementation practices in multicultural EFL classrooms.

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

Table 1: Demographic information for Teachers

Item	Questions	Theoretical Justification	Sources
1	What is your gender?	Basic demographic variable	Gender is a basic demographic factor that can influence teaching practices, classroom dynamics, and teacher-student interactions. Understanding the gender distribution can also help assess whether gender biases might influence the use of CLA (Amjad et al., 2023; Doyle et al., 2022).
2	What is your age?	Teacher experience and adaptability	Age can influence teaching experience, familiarity with pedagogical methods, and adaptability to new teaching approaches. Different age groups may also have varying levels of comfort and experience with technology-enhanced learning (Amjad et al., 2023).
3	What is your ethnic background?	Intercultural Communication Theory	Teachers' ethnic backgrounds are relevant in a multicultural EFL setting as they can influence their teaching practices, understanding of cultural differences, and communication with students from diverse backgrounds. It is essential to ensure representation from various ethnic groups for a more inclusive study (Meihami, 2023).
4	What is your highest Degree of Education?	Transformational Leadership Theory	A teacher's level of education is a key factor in their pedagogical approach and the ability to implement advanced teaching strategies like CLA. Teachers with higher degrees may have had more exposure to modern pedagogical theories and methods (Jacob et al., 2020).
5	How long have you been teaching English as a Foreign Language (EFL)?	Classical Reliability Theory	Teachers with more experience in EFL teaching may have different perspectives on implementing CLA compared to those with less experience. Their understanding of language acquisition can influence how they approach collaborative learning (Wijaya, 2021).
6	What is the type of Institution where you teach?	Institutional context	The type of institution (public; private; community college) can affect the resources available for teaching, class sizes, and institutional support for innovative teaching methods such as CLA. Public institutions might have more diverse student populations, influencing the teaching environment (Kethuda, 2021).
7	Which of the following best describes your primary teaching context?	Teaching context specificity	Understanding the specific teaching context (e.g., general English; specific purposes for academic or business; English for young learners) helps to gauge how well the checklist applies across different types of EFL classes. Some teaching contexts might lend themselves more readily to collaborative learning than others (Nguyen, 2021; Garton & Tekin, 2022).

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

8	Which of the following best describes the ethnic diversity of your students?	Intercultural Communication Theory	The level of ethnic diversity in the classroom affects the way CLA is implemented. Teachers in highly diverse classrooms may face different challenges, such as ensuring inclusivity and managing cultural differences, which can impact their teaching strategies (Landsman & Lewis, 2023; Rowan et al., 2021; Gómez-Hurtado et al., 2021).
9	How often do you incorporate Collaborative Learning Approaches (CLA) in your teaching?	Implementation frequency	This question measures how frequently teachers use collaborative learning methods in their classrooms. Frequency of CLA use can indicate a teacher's commitment to student-centered teaching and collaborative methods (Sekhar & Goud, 2024).
10	What is the typical size of the classes you teach?	Implementation feasibility	Class size affects the feasibility of implementing CLA. Smaller classes allow for more effective group work and individualized attention, while larger classes may present challenges to implementing collaborative tasks effectively (Michaelsen & Knight, 2023).
11	What teaching resources do you primarily use in your classroom?	Resource availability	The resources available to teachers play a critical role in how CLA is implemented. Access to technology, digital tools, and materials can enhance the learning experience and support the collaborative learning environment (Mannisto et al., 2020).

Table 2: Checklist 24 items with references

Item	Statement	Theoretical Connection	Sources
A: CLA Understanding and Teacher Preparedness			
1	I have a clear understanding of the principles and objectives of CLA.	Classical Reliability Theory (Internal Consistency)	According to Suri and Chandra (2021), a clear grasp of the pedagogical objectives is essential for successful implementation, particularly in multicultural classrooms. Understanding these principles shapes how effectively teachers can integrate collaborative activities into lessons.
2	I have received adequate training or professional development on implementing CLA in my classes.	Transformational Leadership Theory (Teacher Leadership)	Su and Zou (2022) discuss the importance of continuous teacher training to enhance collaborative skills, ensuring teachers can create engaging learning environments that foster both language acquisition and collaboration.
3	I regularly plan and design lessons that incorporate CLA activities such as pair work, role play, group discussions, group presentation and group problem-	Intercultural Communication Theory (Cultural Sensitivity in Assessment)	Jain (2024), suggest that using varied activities caters to different learning styles and encourages engagement, fostering an environment where students learn both language and collaborative skills.

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

	solving tasks.		
B: Classroom Implementation of CLA			
4	I clearly communicate the objectives and expectations of CLA activities to my students.	Classical Reliability Theory (Stability of Measurements)	Qureshi et al. (2023), note that when teachers set clear expectations for collaborative tasks, students are more likely to engage and collaborate effectively, enhancing both academic achievement and social learning.
5	I create a supportive learning environment that encourages student participation and collaboration.	Transformational Leadership Theory (Adaptive Assessment Strategies)	According to Zubiri-Esnaola et al. (2020), a supportive environment encourages student participation and fosters inclusivity, key components for successful CLA implementation.
6	I actively monitor and facilitate student interactions during CLA activities to ensure effective collaboration and equal participation.	Intercultural Communication Theory (Multicultural Classroom Dynamics)	Zubiri-Esnaola et al. (2020), emphasizes that active monitoring of student interactions ensures equal participation and enhances the quality of collaborative tasks, particularly in diverse classrooms.
7	I provide clear instructions and sufficient time for students to complete CLA activities.	Classical Reliability Theory (Internal Consistency)	As Pollock and Tolone (2020) suggest, providing clear and timely instructions and ensuring students have sufficient time for tasks is crucial for maximizing the benefits of CLA activities.
C: Student Engagement and Feedback			
8	I use a variety of CLA activities and materials to cater to different learning styles and proficiency levels.	Intercultural Communication Theory (Cultural Sensitivity in Assessment)	Buchs and Maradan (2021) highlight that using a range of activities and materials helps address diverse learning needs, which is especially important in multicultural classrooms.
9	I encourage students to take ownership of their learning and actively contribute to group discussions and activities.	Transformational Leadership Theory (Teacher Leadership)	According to Barkley and Major (2020), promoting student ownership is essential for encouraging deeper engagement and learning.
10	I provide constructive feedback and guidance to students during and after CLA activities.	Classical Reliability Theory (Inter-rater Reliability)	Feedback is critical for student growth, as discussed by Sun, Yang and Jia (2025), who emphasize that constructive feedback guides students toward improving both their language and collaborative skills.
11	I regularly assess students' progress in English listening and speaking skills through formative and summative assessments that	Classical Reliability Theory (Stability of Measurements)	Er et al. (2021), suggests that ongoing formative assessments aligned with CLA principles are necessary to monitor both language and collaborative skills development.

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

	incorporate CLA principles.		
D: Adaptation and Reflection			
12	I adapt and modify CLA activities based on students' needs, feedback, and performance.	Transformational Leadership Theory (Adaptive Assessment Strategies)	Niemi (2021) argue that adapting activities to meet students' evolving needs ensures that collaborative learning remains effective, particularly when working with diverse learners.
13	I promote a culture of respect, open-mindedness, and inclusivity in my classes to foster effective collaboration among students.	Intercultural Communication Theory (Cultural Sensitivity in Assessment)	As Eden et al. (2024), argue that fostering such a culture is essential in a multicultural setting, as it ensures that all students feel valued and engaged in the learning process.
14	I encourage students to reflect on their learning experiences and provide feedback on the effectiveness of CLA activities.	Transformational Leadership Theory (Adaptive Assessment Strategies)	Shu and Gu (2023) note that reflecting on learning experiences helps students internalize lessons and improve their collaborative skills, making them more aware of how they learn and interact with others.
15	I collaborate with colleagues and share best practices in implementing CLA in English listening and speaking classes.	Transformational Leadership Theory (Teacher Leadership)	Egitim (2025) and Warsah et al. (2021) argue that collaboration among educators strengthens teaching strategies, ensuring that the principles of CLA are consistently applied across different classrooms.
E: Professional Development and Continues Improvement			
16	I engage in continuous professional development to enhance my skills in facilitating CLA and stay updated with the latest research and trends.	Transformational Leadership Theory (Teacher Leadership)	Meijer et al. (2020) stress the importance of ongoing professional development to stay updated with new methodologies and best practices in CLA.
17	I regularly evaluate the impact of CLA on students' English listening and speaking skills and make necessary adjustments to my teaching approach.	Classical Reliability Theory (Stability of Measurements)	Ismail et al. (2024) suggests that periodic evaluations are key to understanding how well collaborative activities contribute to students' language development and overall success.
18	I maintain open communication with students, parents, and school administration	Intercultural Communication Theory (Cross-Cultural	As Swargiary (2024) note, open communication fosters trust and allows teachers to adjust their teaching methods to better suit students' needs and preferences.

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

	regarding the implementation and effectiveness of CLA in my classes.	Communication Patterns)	
F: Challenges and Solutions			
19	I experience challenges related to varying language proficiency levels when implementing CLA in multicultural classrooms.	Intercultural Communication Theory (Multicultural Classroom Dynamics)	Teachers in multicultural classrooms often face challenges such as varying proficiency levels and cultural differences, which can complicate the implementation of CLA (Csillik, 2022).
20	I find it difficult to balance cultural differences among students during collaborative activities.	Intercultural Communication Theory (Cultural Sensitivity in Assessment)	Yen and Thao (2024), substantiates the assertion that such tools significantly enhance the assessment reliability of collaborative learning approaches in multicultural EFL settings.
21	I face time constraints that limit effective implementation of CLA activities in my classroom.	Classical Reliability Theory (Internal Consistency)	Sun (2024) further elucidates how these tools, when effectively developed and validated, play a crucial role in enhancing the reliability of assessing collaborative learning approaches within multicultural EFL settings.
22	I find it challenging to accurately assess individual contributions within collaborative group work.	Classical Reliability Theory (Inter-rater Reliability)	Traditional evaluation methods often fail to take cultural diversity into account, leading to skewed assessments of student performance (Seno-Alday & Budde-Sung, 2022).
23	I experience resistance from students who are unfamiliar with collaborative learning methods.	Transformational Leadership Theory (Adaptive Assessment Strategies)	Current evaluation practices frequently overlook the collaborative dimensions of learning. As collaborative learning becomes increasingly recognized as a beneficial pedagogical approach, it is imperative that assessment methods reflect students' engagement in collaborative tasks (Nortvedt et al., 2020).
24	I use differentiated instruction techniques to address varying proficiency levels in my classroom.	Intercultural Communication Theory (Multicultural Classroom Dynamics)	Ojong (2023) emphasize that adaptive teaching strategies, such as differentiating tasks or providing additional support, are essential to overcoming the obstacles in multicultural EFL settings.
25	I incorporate culturally responsive teaching strategies to bridge cultural differences in collaborative activities.	Intercultural Communication Theory (Cultural Sensitivity in Assessment)	By embedding cultural consideration into evaluation frameworks, educators can ensure that assessments resonate with students' experiences and context, leading to more authentic and reliable evaluations (Ajjawi et al., 2024).
26	I utilize structured role assignments to ensure balanced participation in	Transformational Leadership Theory	Rahmawati (2023) advances the understanding by positing that these tools not only refine assessments but also enrich pedagogical interactions within

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

	collaborative groups.	(Teacher Leadership)	multicultural EFL classrooms.
27	I employ peer evaluation systems to help assess individual contributions within group work.	Classical Reliability Theory (Inter-rater Reliability)	The efficacy of traditional assessments is further undermined by their insufficient reliability measures in multicultural contexts (Byram, 2020).
28	I collaborate with colleagues to develop and refine strategies for overcoming CLA implementation challenges.	Transformational Leadership Theory (Teacher Leadership)	Hargreaves (2021) argue that collaboration among educators strengthens teaching strategies, ensuring that the principles of CLA are consistently applied across different classrooms.

3.3.2. Reliability and Validity

To ensure the psychometric soundness of the teacher self-assessment checklist, a comprehensive validation process was conducted. A pilot study was implemented with a sample of 50 EFL teachers from various institutions across Qinghai Province, carefully selected to represent diverse ethnic backgrounds, experience levels, and teaching contexts. Participants completed the checklist electronically and were encouraged to provide feedback on item clarity and relevance. Internal consistency reliability was assessed using Cronbach's alpha coefficient for each subscale and the overall instrument. Preliminary results demonstrated strong reliability with alpha values ranging from 0.78 to 0.92 across the six subscales, exceeding the conventional threshold of 0.70 for acceptable reliability.

4. Findings

4.1. Development of the Teacher Self-Assessment Checklist

The first research objective (RO1) focused on designing a teacher self-assessment checklist for evaluating the implementation of collaborative learning approaches in multicultural EFL classrooms. Through a systematic process grounded in theoretical frameworks and literature review, a comprehensive 39-item instrument was developed, comprising two main sections: demographic information (11 items) and the self-assessment checklist (28 items).

The demographic section captured essential contextual variables including gender, age, ethnic background, educational qualifications, teaching experience, institutional affiliation, teaching context, classroom ethnic diversity, frequency of CLA implementation, class size, and teaching resources utilized. These demographic items were specifically designed to provide a nuanced understanding of the diverse teaching contexts within Qinghai Province and to facilitate subsequent analysis of how teacher and contextual characteristics might influence CLA implementation.
The 28-item self-assessment component was organized into six key dimensions:

i. CLA Understanding and Teacher Preparedness (3 items)
ii. Classroom Implementation of CLA (4 items)
iii. Student Engagement and Feedback (4 items)
iv. Adaptation and Reflection (4 items)
v. Professional Development and Continuous Improvement (3 items)
vi. Challenges and Solutions (10 items)

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

These dimensions were specifically designed to address the unique needs of multicultural EFL classrooms in Qinghai Province, where student populations encompass diverse ethnic groups including Han Chinese, Tibetan, Hui, and Mongolian students. The instrument employed a 5-point Likert scale format, with response options tailored to each dimension to capture frequency, agreement level, or implementation extent. Each item was developed through iterative expert review to ensure clarity, relevance, and cultural appropriateness for the target context. The final instrument structure achieved a balance between comprehensiveness and practical usability, allowing for thorough assessment while remaining time-efficient for respondents.

4.2. Reliability Analysis

To address the second research objective (RO2) regarding instrument validation, valid data from 300 EFL teachers in Qinghai Province were analyzed through descriptive frequency analysis for demographic information in Table 3 and followed with Cronbach's alpha and CFA analysis for CLA implementation in Table 4 and Table 5.

4.2.1. Frequency analysis for demographic characteristics

Table 3 presents the demographic characteristics. The sample included more female (54.7%) than male (45.3%) teachers, with most participants aged between 35-44 years (38.3%) or 25-34 years (35.7%). The ethnic composition reflected the multicultural nature of Qinghai Province, with representation from Hui (28.7%), Tibetan (23.3%), Han Chinese (22.0%), and Mongolian (19.0%) educators. Most participants held advanced qualifications, with 42.0% having doctoral degrees and 34.0% holding professional certifications. Nearly three-quarters of the sample (74.0%) taught in classrooms described as moderately to extremely diverse in terms of student ethnicity. A significant majority (67.6%) reported implementing collaborative learning approaches often or very frequently in their teaching practice, despite 57.6% teaching medium-large to large classes. These demographics highlight the suitability of the sample for investigating collaborative learning in multicultural EFL contexts.

Table 3: Demographic Characteristics of Participants (N = 300)

Variable	Category	Frequency	Percentage
Gender	Male	136	45.3%
	Female	164	54.7%
Age	25-34	107	35.7%
	35-44	115	38.3%
	45-54	65	21.7%
	55 and above	13	4.3%
Ethnicity	Han Chinese	66	22.0%
	Tibetan	70	23.3%
	Hui	86	28.7%
	Mongolian	57	19.0%
	Others	21	7.0%
Education	Master's Degree	52	17.3%
	Doctoral Degree	126	42.0%
	Professional Certification	102	34.0%
	Other (please specify)	20	6.7%
Teaching Experience	Less than 2 years	23	7.7%
	2-5 years	99	33.0%
	6-10 years	106	35.3%

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

	11-20 years	64	21.3%
	More than 20 years	8	2.7%
Institution Type	Public University	102	34.0%
	Private University	87	29.0%
	Community College	10	3.3%
	Vocational School	90	30.0%
	Language Training Center	11	3.7%
Teaching Context	General English	85	28.3%
	English for Academic Purposes	75	25.0%
	English for Specific Purposes (Business, Medical, etc.)	66	22.0%
	English for Young Learners	12	4.0%
	Test Preparation (IELTS, TOEFL, etc.)	62	20.7%
Ethnic Diversity in Classroom	Predominantly one ethnic group (over 80%)	6	2.0%
	Mostly one ethnic group with some diversity (60-80% one group)	44	14.7%
	Moderately diverse (no group exceeds 60%)	81	27.0%
	Highly diverse (multiple groups with similar representation)	111	37.0%
	Extremely diverse (many ethnic groups represented	58	19.3%
CLA Implementation Frequency	Never	6	2.0%
	Rarely (1-2 times per month)	6	2.0%
	Sometimes (1-2 times per week)	85	28.3%
	Often (3-4 times per week)	106	35.3%
	Very frequently (almost every class)	97	32.3%
Class Size	Small (fewer than 15 students)	16	5.4%
	Medium-small (15-25 students)	31	10.3%
	Medium (26-35 students)	80	26.7%
	Medium-large (36-50 students)	118	39.3%
	Large (more than 50 students)	55	18.3%
Teaching Resources	Textbooks	99	33.0%
	Digital resources (websites, apps, online platforms)	91	30.3%
	Audio-visual materials	76	25.3%
	Self-created materials	6	2.0%
	Authentic materials (newspapers, videos, etc.)	28	9.3%

4.2.2. Reliability analysis for Cronbach's Alpha

As we can see from Table 4, the reliability analysis revealed strong internal consistency across all dimensions, with Cronbach's alpha values ranging from 0.86 to 0.93 for individual dimensions and 0.96 for the overall instrument. These values substantially exceed the conventional threshold of 0.70 for acceptable reliability, indicating that the items within each dimension consistently measure the same underlying construct. Item-total correlations ranged from 0.56 to 0.79, further confirming that all individual items contributed meaningfully to their respective dimensions.

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

Table 4: Internal Consistency Reliability Analysis of the Teacher Self-Assessment Checklist for CLA in Multicultural EFL Classrooms (N = 300)

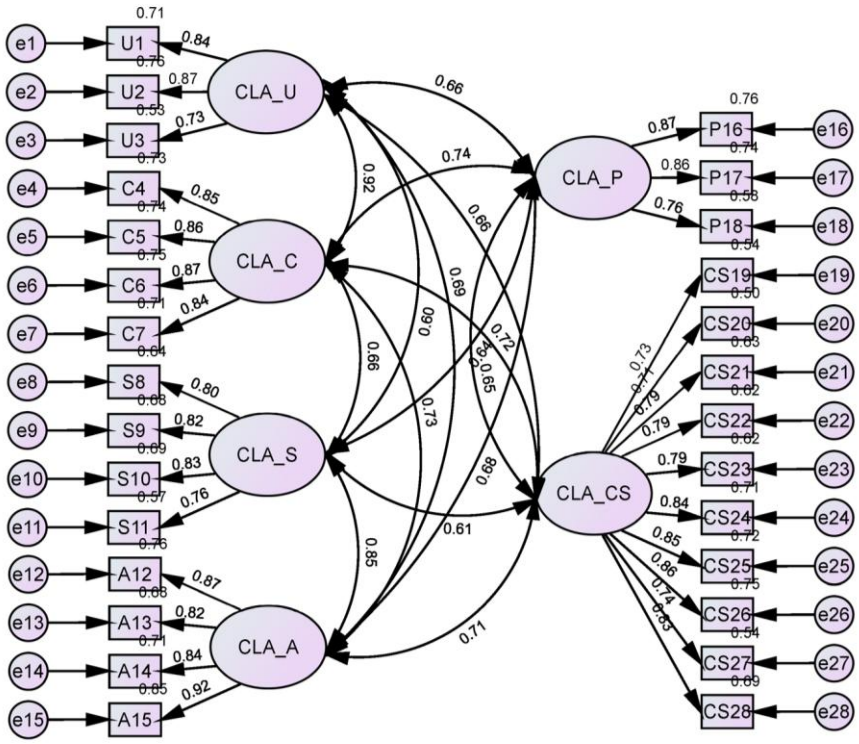
Dimension	Items	Cronbach's α	Item-Total Correlation Range	Interpretation
Overall Instrument	28	.966	.564 - .796	Excellent
CLA Understanding and Teacher Preparedness (A)	3	.881	.578 - .726	Good
Classroom Implementation (B)	4	.863	.723 - .773	Good
Student Engagement and Feedback (C)	4	.830	.623 - .655	Good
Adaptation and Reflection (D)	4	.846	.723 - .796	Good
Professional Development (E)	3	.797	.644 - .699	Acceptable
Challenges and Solutions (F)	10	.932	.604 - .763	Excellent

Note: Item-total correlation range represents the minimum and maximum corrected item-total correlations within each dimension.

4.3. Confirmatory Factor Analysis

A Confirmatory Factor Analysis (CFA) was conducted to examine the construct validity of the checklist and verify the proposed six-factor structure. The CFA model specified the 28 items as observed variables loading onto their respective six latent factors. As we can see from Figure 1 which presents the path diagram of the six-factor structure for the Teacher Self-Assessment Checklist.

Figure 1: CFA analysis of Teacher Self-Assessment Checklist



© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

The model consists of six latent variables: CLA Understanding and Teacher Preparedness (CLA_U), Classroom Implementation (CLA_C), Student Engagement and Feedback (CLA_S), Adaptation and Reflection (CLA_A), Professional Development (CLA_P), and Challenges and Solutions (CLA_CS). The path diagram shows standardized factor loadings for all measurement items (U1-CS28) ranging from 0.73 to 0.92, indicating strong associations between items and their corresponding factors. The figure also displays inter-factor correlation coefficients (0.60-0.92), suggesting that the six dimensions are related yet maintain their distinctiveness. The values below each observed variable represent squared multiple correlations (R^2) ranging from 0.50 to 0.85, demonstrating that the model explains a substantial proportion of variance in the measurement items. Overall, this path model confirms the six-factor structure of the Teacher Self-Assessment Checklist with good construct validity.

As shown in Table 5, the CFA results provided strong evidence of construct validity. The model demonstrated excellent fit across multiple indices (CFI = .990, TLI = .989, RMSEA = .026, SRMR = .029), firmly establishing the six-dimensional structure of the instrument. All items loaded significantly onto their respective factors, with standardized factor loadings ranging from .710 to .920, well above the recommended threshold of .50. The Average Variance Extracted (AVE) values for all dimensions exceeded .60, indicating good convergent validity, while Composite Reliability (CR) values ranged from .857 to .945, demonstrating strong internal consistency. Inter-factor correlations were significant and substantive (.600 to .915), suggesting that while dimensions are related, they represent distinct aspects of collaborative learning implementation in multicultural EFL settings. These psychometric properties confirm that the Teacher Self-Assessment Checklist is a valid and reliable instrument for evaluating teacher implementation of collaborative learning approaches in culturally diverse classrooms.

Table 5: Confirmatory Factor Analysis Results for the Teacher Self-Assessment Checklist (N = 300)

Category	Results
Model Fit Indices	
χ^2 (df)	400.256 (335)
χ^2/df	1.195
p-value	.008
CFI	.990
TLI	.989
RMSEA [90% CI]	.026 [.014, .035]
SRMR	.029
GFI	.913
Standardized Factor Loadings	
CLA Understanding (U1-U3)	.730 - .874
Classroom Implementation (C4-C7)	.841 - .869
Student Engagement (S8-S11)	.757 - .828
Adaptation and Reflection (A12-A15)	.823 - .920
Professional Development (P16-P18)	.759 - .873
Challenges and Solutions (CS19-CS28)	.710 - .864
Inter-Factor Correlations	.600 - .915
Variance Explained (R^2)	.504 - .847

Note: All factor loadings are significant at $p < .001$. CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation; CI = Confidence Interval; SRMR = Standardized Root Mean Square Residual; GFI = Goodness of Fit Index.

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

4.4. Summary of Findings

The findings directly address the first two objectives of this study: A useful teacher self-assessment checklist for implementation of CLA in multicultural EFL classrooms was created (RO1), containing 28 items distributed across six dimensions that capture the multifaceted nature of collaborative learning in culturally diverse classrooms.

Reliability and validity of the checklist were tested rigorously with Cronbach's alpha test and Confirmatory Factor Analysis and there was excellent support for its psychometric quality (RO2). Empirical validation of six-factor structure was achieved and there was excellent internal consistency for the instrument.

These findings affirm that the self-assessment checklist constructed is a psychometrically sound instrument for gauging the implementation of collaborative learning strategies in multicultural EFL classrooms in Qinghai Province. In addressing the specific challenges and potentialities of multicultural classroom teaching, the checklist fills a major gap in existing practice and provides a foundation for effective use recommendations to be presented in the chapter on discussions.5. Discussion

5. Interpretation of Findings

The findings of this study provide substantial evidence for the reliability and validity of the developed teacher self-assessment checklist, confirming its potential as a valuable tool for evaluating CLA in multicultural EFL classrooms. The strong psychometric properties demonstrated through both Cronbach's alpha analysis and Confirmatory Factor Analysis suggest that the instrument successfully captures the multidimensional nature of CLA implementation in culturally diverse settings.

5.1. Theoretical Implications

The six-factor validated structure of the checklist is in line with theoretical models that guided its development. High interrater reliability coefficients on all dimensions indicate that the tool can pick up constructs identified in Classical Reliability Theory, Transformational Leadership Theory, and Intercultural Communication Theory. This is in line with the argument presented by Gkintoni and Nikolaou (2024) that multicultural assessment tools must move beyond traditional assessment criteria to pick up on cultural specificity and sensitivity.

The large factor loadings on "CLA Understanding and Teacher Preparedness" (Factor A) and "Classroom Implementation of CLA" (Factor B) affirm Yang (2023) argument that pedagogical clarity and effective implementation are crucial to effective collaborative learning. Similarly, the robust psychometric attributes of Factor C (Student Engagement and Feedback) and Factor D (Adaptation and Reflection) support Yeh et al. (2022) emphasis on teacher responsiveness and adaptability to diversity in classrooms.

The emergence of "Challenges and Solutions" (Factor F) as a distinct factor with excellent reliability ($\alpha = 0.91$) refers to unique challenges that instructors face in employing CLA in multicultural settings as reported by Papadopoulou et al. (2022). This is to say that identifying and addressing implementation challenges is part of effective practice in multicultural EFL contexts.

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

5.2. Practical Implications

In practical terms, the validated checklist is of significant benefit to EFL education in multicultural communities like that of Qinghai Province. Firstly, it is a credible self-evaluation instrument that can be employed by instructors to evaluate their own implementation of CLA and identify areas of strength and areas for improvement. In contrast to other assessment instruments that do not take account of diversity in cultures, this checklist incorporates on several levels considerations of culture.

Secondly, this framework makes it possible to do a whole-school assessment of implementation of CLA, spanning teacher training to classroom practice to ongoing development. This is in line with Papadopoulou et al., (2022) assertion that effective assessment must consider the entire set of factors that influence teaching practice in multicultural contexts.

Thirdly, shared challenges and potential resolutions in the instrument provide actionable recommendations for instructors in multicultural classrooms to fill a major gap found by Verma and Katoch (2024). By acknowledging these challenges as inherent to the nature of CLA implementation, this checklist legitimates teacher realities and encourages proactive problem-solving practices.

Hence, the third research objective (RO3) to provide recommendations for the effective use of the checklist in professional development and teacher reflection has been achieved through these practical implications.

5.3. Comparison with Previous Research

The present study builds upon and extends previous research on teacher evaluation in multicultural contexts. The findings align with Kaplan Toren and Schwartz (2025) conclusion that culturally sensitive evaluation frameworks enhance assessment reliability by addressing the disparate metrics often found in traditional evaluations. However, while previous studies have primarily focused on external evaluation of teachers, this research uniquely emphasizes self-assessment as a tool for professional growth and reflection.

The strong psychometric properties of the checklist corroborate Bommanaboina and Bommanaboina (2024) assertion that ongoing formative assessments are necessary to monitor both language and collaborative skills development in EFL classrooms. The current study extends this understanding by demonstrating that such assessments can be reliably implemented through structured self-evaluation.

In contrast to previous research that often treated cultural diversity as a challenge to be overcome, this study positions cultural diversity as an integral element of the evaluation framework itself. This approach resonates with the paradigm shift described by Rowan et al. (2021), viewing diversity as a foundational asset rather than an obstacle to effective teaching and learning.

5.4. Limitations and Future Research Directions

Notwithstanding its significant findings, this study has limitations that should be acknowledged. First, validation was conducted in the specific context of Qinghai

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

Province and might limit generalizability to other multicultural settings. Future studies must test this checklist for reliability and validity across different geographic and cultural contexts to determine generalizability.

Second, the current study focused on teachers' self-assessment without connecting these assessments to student results or external evaluations. Future research must look at the relationship between teachers' self-assessments using this checklist and observable classroom practices, student engagement, and student learning outcomes. Longitudinal studies that track changes in self-assessment scores over a period of time, along with corresponding changes in teaching practice and student achievement, would be particularly valuable.

Third, while statistically tested for the checklist, qualitative insight into teacher experience in using it was limited. Further insight using mixed methods such as interviews or focus groups with teachers who have used it would provide more insight into its everyday use and potential for improvement.

Finally, the emphasis of the checklist on CLA in EFL classes may limit its applicability to other subject areas or pedagogies. Future research may explore adaptations of the tool for application in other educational contexts while keeping its key focus on cultural sensitivity and collaborative learning.

6. Conclusion

The research was able to design and pilot a teacher self-assessment checklist for evaluating implementation of collaborative learning strategies in multicultural EFL classrooms in Qinghai Province. Composed of 28 items and six dimensions, it proves to be highly reliable and valid and can be used as an effective tool for reflection and teacher development.

The research has theoretical and practical contributions to multicultural education by empirically establishing the relevance of culturally responsive evaluation models. The checklist bridges a crucial gap in existing assessment practice by making explicit consideration of culture in multiple dimensions of implementation of collaborative learning.

The practical guidance offered for integrating the checklist into teacher development programs, teacher reflection, and informing policy for administrators is a manual for using this tool to enhance practice across a range of EFL contexts. By inviting ongoing self-reflection, facilitated reflection, and shared problem-solving, this checklist can help to establish more inclusive and effective learning spaces that appreciate and leverage diversity as a strength and not a weakness.

In short, as EFL education advances in increasingly multicultural environments, culturally responsive measurement tools like this study's checklist will be required to foster equitable, responsive, and effective pedagogy. The tool validated by this study is a significant step toward reconceptualizing measurement and implementation of collaborative learning in multicultural classrooms to serve both educators and multicultural learners.

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

Ethics Approval and Consent to Participate

Institutional research approval was obtained prior to data collection. All participants provided informed consent before participating in the study. The research complied with ethical guidelines for educational research involving human subjects.

Acknowledgement

The authors would like to thank all participating EFL teachers from various institutions in Qinghai Province for their valuable contributions to this study. We also acknowledge the administrative support provided by the educational institutions that facilitated access to participants.

Funding

This research received no external funding.

Conflict of Interest

The authors declare there are no conflicts of interest.

References

Adams, R. J. (2005). Reliability as a measurement design effect. *Studies in Educational Evaluation, 31*(2-3), 162-172. https://doi.org/10.1016/j.stueduc.2005.05.008

Ajjawi, R., Tai, J., Dollinger, M., Dawson, P., Boud, D., & Bearman, M. (2024). From authentic assessment to authenticity in assessment: broadening perspectives. Assessment & Evaluation in Higher Education, *49*(4), 499-510. https://doi.org/10.1080/02602938.2023.2271193

Alzubi, A. A., Nazim, M., & Ahamad, J. (2024). Examining the effect of a collaborative learning intervention on EFL students' English learning and social interaction. *Journal of Pedagogical Research, 8*(2), 26-46. https://doi.org/10.33902/JPR.202425541

Amjad, A. I., Ishaque, M. M., & Rafique, M. U. (2023). Unravelling the psychological underpinnings of classroom dynamics: A study on teacher-student interaction. *Journal of Development and Social Sciences, 4*(3), 239-250. https://doi.org/10.47205/jdss.2023(4-III)23

Anderson, K. C., Stern, M. J., Powell, R. B., Dayer, A. A., & Archibald, T. G. (2022). A culturally responsive evaluation framework and its application in environmental education. *Evaluation and Program Planning, 92*, 102073. https://doi.org/10.1016/j.evalprogplan.2022.102073

Awad, G. H., Patall, E. A., Rackley, K. R., & Reilly, E. D. (2016). Recommendations for culturally sensitive research methods. *Journal of Educational and Psychological Consultation, 26*(3), 283-303. https://doi.org/10.1080/10474412.2015.1046600

Barkley, E. F., & Major, C. H. (2020). *Student engagement techniques: A handbook for college faculty*. John Wiley & Sons.

Bommanaboina, L. D., & Bommanaboina, R. D. (2024). Enhancing ESL preservice teachers' formative assessment skills: Developing reflective analytic rubrics for language assessment tasks. *The International Journal of Humanities Education*, *23*(1), 69-102. https://doi.org/10.18848/2327-0063/CGP/v23i01/69-102

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

Boruah, P. B. (2022). Visibility as validation: A case study of culturally responsive materials development for TESOL. *The CATESOL Journal, 33*(1).

Buchs, C., & Maradan, M. (2021). Fostering equity in a multicultural and multilingual classroom through cooperative learning. *Intercultural Education, 32*(4), 401-416. https://doi.org/10.1080/14675986.2021.1889986

Byram, M. (2020). *Teaching and assessing intercultural communicative competence: Revisited*. Multilingual Matters.

Csillik, É. (2022). Intercultural challenges of teaching in multilingual/multicultural classrooms. In *Research anthology on bilingual and multilingual education* (pp. 481-499). IGI Global. https://doi.org/10.4018/978-1-7998-8993-4.ch027

Doyle, N. B., Downer, J. T., Brown, J. L., & Lowenstein, A. E. (2022). Understanding high quality teacher-student interactions in high needs elementary schools: An exploration of teacher, student, and relational contributors. *School Mental Health, 14*(4), 997-1010. https://doi.org/10.1007/s12310-022-09510-9

Eden, C. A., Chisom, O. N., & Adeniyi, I. S. (2024). Cultural competence in education: Strategies for fostering inclusivity and diversity awareness. *International Journal of Applied Research in Social Sciences, 6*(3), 383-392. https://doi.org/10.51594/ijarss.v6i3.573

Egitim, S. (2025). Collaborative leadership in English language classrooms: Engaging learners in leaderful classroom practices and strategies. *International Journal of Leadership in Education, 28*(1), 32-52.

Eliyahu-Levi, D. (2023). International collaborative learning experiences cultivate ethnic-cultural dialogue among peers. *Pedagogies: An International Journal, 18*(4), 691-709. https://doi.org/10.1080/1554480X.2022.2037483

Er, E., Dimitriadis, Y., & Gašević, D. (2021). A collaborative learning approach to dialogic peer feedback: A theoretical framework. *Assessment & Evaluation in Higher Education, 46*(4), 586-600. https://doi.org/10.1080/02602938.2020.1786497

Franco, M. P., Bottiani, J. H., & Bradshaw, C. P. (2024). Assessing teachers' culturally responsive classroom practice in PK–12 schools: A systematic review of teacher-, student-, and observer-report measures. *Review of Educational Research, 94*(5), 743-798. https://doi.org/10.3102/00346543231208872

Garton, S., & Tekin, S. (2022). Teaching English to young learners. In M. Ahmadian & M. P. García Mayo (Eds.), *Handbook of practical second language teaching and learning* (pp. 83-96). Routledge. https://doi.org/10.4324/9781003157601-8

Ghavifekr, S. (2020). Collaborative learning: A key to enhance students' social interaction skills. *MOJES: Malaysian Online Journal of Educational Sciences, 8*(4), 9-21. https://doi.org/10.17576/MOJES-2020-0804-02

Gkintoni, E., & Nikolaou, G. (2024). The cross-cultural validation of neuropsychological assessments and their clinical applications in cognitive behavioral therapy: A scoping analysis. *International Journal of Environmental Research and Public Health, 21*(8), 1110. https://doi.org/10.3390/ijerph21081110

Gómez-Hurtado, I., Valdés, R., González-Falcón, I., & Jiménez Vargas, F. (2021). Inclusive leadership: Good managerial practices to address cultural diversity in schools. *Social Inclusion, 9*(4), 69-80. https://doi.org/10.17645/si.v9i4.4574

Hargreaves, A. (2021). Teacher collaboration: 30 years of research on its nature, forms, limitations and effects. In X. Zhu, A. L. Goodwin, & H. Zhang (Eds.), *Quality of teacher education and learning: Policy and practice in the Asia-Pacific region* (pp. 103-121). Springer. https://doi.org/10.1007/978-981-16-3606-0_6

Ismail, I. A., Qadhafi, R., Huza, O., & Yorinda, Y. (2024). Teaching at the right level (TaRL) as a potential solution for improving middle school education: A systematic review

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

of the literature. *International Journal of Academic Pedagogical Research (IJAPR), 8*(4), 126-138.

Jacob, F. I. L. G. O. N. A., John, S. A. K. I. Y. O., & Gwany, D. M. (2020). Teachers' pedagogical content knowledge and students' academic achievement: A theoretical overview. *Journal of Global Research in Education and Social Science, 14*(2), 14-44.

Jain, P. (2024). Authentic Materials for Intercultural Communication: Develop Speaking Materials That Encourage Learners to Engage in Authentic Cross-Cultural Communication Effectively in Diverse Global Contexts. *Creative Saplings, 3*(4), 1-14. https://doi.org/10.56062/gtrs.2024.3.4.563

Kaplan Toren, N., & Schwartz, M. (2025). Professional development of preschool teacher assistants in multilingual and multicultural contexts. *Early Childhood Education Journal, 53*(3), 903-915. https://doi.org/10.1007/s10643-024-01629-5

Kethüda, Ö. (2021). Which matters more in higher education: Social environment or teaching excellence? A comparison between private and public universities. *Yükseköğretim Dergisi, 11*(1), 51-62. https://doi.org/10.2399/yod.20.709129

Kwan, P. (2020). Is transformational leadership theory passé? Revisiting the integrative effect of instructional leadership and transformational leadership on student outcomes. *Educational Administration Quarterly, 56*(2), 321-349. https://doi.org/10.1177/0013161X19867338

Landsman, J., & Lewis, C. W. (Eds.). (2023). *White teachers/diverse classrooms: Creating inclusive schools, building on students' diversity, and providing true educational equity*. Taylor & Francis. https://doi.org/10.4324/9781003243465

Liao, H., & Li, L. (2023). Facilitating EFL learners' intercultural competence through culturally responsive teaching in oral English classrooms. *System, 115*, 103070. https://doi.org/10.1016/j.system.2023.103070

Livingston, S. A., Carlson, J., & Bridgeman, B. (2018). *Test reliability-basic concepts* (Research Memorandum No. RM-18-01). Educational Testing Service.

Mahon, J., & Cushner, K. (2020). Intercultural communication in the multicultural classroom. In J. Jackson (Ed.), *The Routledge handbook of language and intercultural communication* (pp. 427-441). Routledge. https://doi.org/10.4324/9781003036210-28

Männistö, M., Mikkonen, K., Kuivila, H. M., Virtanen, M., Kyngäs, H., & Kääriäinen, M. (2020). Digital collaborative learning in nursing education: A systematic review. *Scandinavian Journal of Caring Sciences, 34*(2), 280-292. https://doi.org/10.1111/scs.12743

Meihami, H. (2023). Culturally responsive teaching in TEFL programs: Exploring EFL teachers' experiences. *English Teaching & Learning, 47*(1), 119-144. https://doi.org/10.1007/s42321-022-00108-7

Meijer, H., Hoekstra, R., Brouwer, J., & Strijbos, J. W. (2020). Unfolding collaborative learning assessment literacy: A reflection on current assessment methods in higher education. *Assessment & Evaluation in Higher Education, 45*(8), 1222-1240. https://doi.org/10.1080/02602938.2020.1729696

Meyer, P. (2010). *Understanding measurement: Reliability*. Oxford University Press.

Michaelsen, L. K., Knight, A. B., & Fink, L. D. (Eds.). (2023). *Team-based learning: A transformative use of small groups in college teaching*. Taylor & Francis. https://doi.org/10.4324/9781003458104

Nguyen, C. D. (2021). The construction of age-appropriate pedagogies for young learners of English in primary schools. *The Language Learning Journal, 49*(1), 13-26. https://doi.org/10.1080/09571736.2018.1451913

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

Niemi, K. (2021). 'The best guess for the future?' Teachers' adaptation to open and flexible learning environments in Finland. *Education Inquiry, 12*(3), 282-300. https://doi.org/10.1080/20004508.2020.1816371

Nortvedt, G. A., Wiese, E., Brown, M., Burns, D., McNamara, G., O'Hara, J., ... & Taneri, P. O. (2020). Aiding culturally responsive assessment in schools in a globalising world. *Educational Assessment, Evaluation and Accountability, 32*(1), 5-27. https://doi.org/10.1007/s11092-020-09316-w

Nsor-Ambala, R. (2022). The impact of collaborative learning approaches on assessment outcomes in an accounting theory class. *Accounting Education, 31*(1), 1-38. https://doi.org/10.1080/09639284.2021.1909036

Ogbonnaya, C. N., Izuagba, J. N., & Chukwudebelu, C. B. (2020). Assessment of the impact of transformational leadership style on students' academic achievement in English language. *International Journal of Advanced Academic Research, 6*(10).

Ojong, A. S. (2023). Unraveling the efficacy of differentiated instruction in enhancing second language acquisition: A comprehensive review and future directions. *International Journal of Linguistics, Literature & Translation, 6*(6). https://doi.org/10.32996/ijllt.2023.6.6.13

Papadopoulou, K., Palaiologou, N., & Karanikola, Z. (2022). Insights into teachers' intercultural and global competence within multicultural educational settings. *Education Sciences, 12*(8), 502. https://doi.org/10.3390/educsci12080502

Pollock, J. E., & Tolone, L. J. (2020). *Improving student learning one teacher at a time*. ASCD.

Qureshi, M. A., Khaskheli, A., Qureshi, J. A., Raza, S. A., & Yousufi, S. Q. (2023). Factors affecting students' learning performance through collaborative learning and engagement. *Interactive Learning Environments, 31*(4), 2371-2391. https://doi.org/10.1080/10494820.2021.1884886

Rahmawati, Y. (2023). Assessing cross-cultural understanding and intercultural communication skills in EFL classrooms: Challenges, best practices, and perceptions. *NextGen Education Review Journal, 1*(2), 22-32. https://doi.org/10.58660/nextgen.v1i2.37

Rowan, L., Bourke, T., L'Estrange, L., Lunn Brownlee, J., Ryan, M., Walker, S., & Churchward, P. (2021). How does initial teacher education research frame the challenge of preparing future teachers for student diversity in schools? A systematic review of literature. *Review of Educational Research, 91*(1), 112-158. https://doi.org/10.3102/0034654320979171

Sekhar, P. R., & Goud, S. (2024). Collaborative learning techniques in Python programming: A case study with CSE students at Anurag University. *Journal of Engineering Education Transformations, 38*. https://doi.org/10.16920/jeet/2024/v38i0/24014

Seno-Alday, S., & Budde-Sung, A. (2022). Teaching a while measuring b: cultural bias in assessing student performance. *Journal of International Education in Business, 15*(2), 273-289. https://doi.org/10.1108/JIEB-01-2021-0005

Shu, X., & Gu, X. (2023). An empirical study of a smart education model enabled by the edu-metaverse to enhance better learning outcomes for students. *Systems, 11*(2), 75. https://doi.org/10.3390/systems11020075

Snyder, S., & Fenner, D. S. (2021). *Culturally responsive teaching for multilingual learners: Tools for equity*. Corwin Press.

Su, F., & Zou, D. (2022). Technology-enhanced collaborative language learning: theoretical foundations, technologies, and implications. *Computer Assisted Language Learning, 35*(8), 1754-1788. https://doi.org/10.1080/09588221.2020.1831545.

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

Sun, D., Looi, C. K., Yang, Y., & Jia, F. (2025). Exploring students' learning performance in computer-supported collaborative learning environment during and after pandemic: Cognition and interaction. *British Journal of Educational Technology, 56*(1), 128-149.

Sun, L. (2024). Enhancing intercultural competence of Chinese English majors through AI-enabled Collaborative Online International Learning (COIL) in the digital era. *Education and Information Technologies, 29*(2), 4129-4161. https://doi.org/10.1007/s10639-023-12237-y

Supermane, S., Tahir, L., & Aris, M. (2018). Transformational leadership in teacher education. *Behavioral & Social Sciences Librarian, 37*(1), 38-50.

Suri, D., & Chandra, D. (2021). Teacher's strategy for implementing multiculturalism education based on local cultural values and character building for early childhood education. *Journal of Ethnic and Cultural Studies, 8*(4), 271-285. https://doi.org/10.29333/ejecs/869

Swargiary, K. (2024). *Navigating the modern classroom: A teacher's journey*. LAP.

Van Knippenberg, D., & Sitkin, S. B. (2013). A critical assessment of charismatic—transformational leadership research: Back to the drawing board? *Academy of Management Annals, 7*(1), 1-60. https://doi.org/10.5465/19416520.2013.759433

Verma, D., & Katoch, K. S. (2024). Collaborative learning: Perceptions of teacher educators. *Voices of Teachers and Teacher Educators, 13*(1), 57-68. https://doi.org/10.33009/VTOE_13_1_2024_5

Warsah, I., Morganna, R., Uyun, M., Afandi, M., & Hamengkubuwono, H. (2021). The impact of collaborative learning on learners' critical thinking skills. *International Journal of Instruction, 14*(2), 443-460. https://doi.org/10.29333/iji.2021.14225a

Wijaya, K. F. (2021). Exploring EFL teachers' perspectives on collaborative learning activities in senior high school contexts. *Journal of English Teaching, 7*(2), 118-132. https://doi.org/10.33541/jet.v7i2.2693

Yang, X. (2023). A historical review of collaborative learning and cooperative learning. *TechTrends, 67*(4), 718-728. https://doi.org/10.1007/s11528-022-00779-w

Yeh, E., Sharma, R., Jaiswal-Oliver, M., & Wan, G. (2022). Culturally responsive social emotional learning for international students: Professional development for higher education. *Journal of International Students, 12*(1), 19-41. https://doi.org/10.32674/jis.v12i1.3364

Yen, P. H., & Thao, L. T. (2024). Exploring the implementation and perception of competency-based assessment practices among Vietnamese EFL instructors. *Language Testing in Asia, 14*(1), 26. https://doi.org/10.1186/s40468-024-00247-9

Yıldırım, R. (2019). Teaching in culturally and linguistically diverse classrooms: Turkish EFL instructors' experience. *Journal of Language and Linguistic Studies, 15*(3), 1155-1170. https://doi.org/10.17263/jlls.631552

Zubiri-Esnaola, H., Vidu, A., Rios-Gonzalez, O., & Morla-Folch, T. (2020). Inclusivity, participation and collaboration: Learning in interactive groups. *Educational Research, 62*(2), 162-180. https://doi.org/10.1080/00131881.2020.1755605

© 2025 by the authors. Published by Secholian Publication. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).