

# Multidimensional Construction and Tiered Development of English Teachers' Assessment Literacy in the Context of New Liberal Arts: A Study Based on English Teachers in Universities of Yunnan Province

Na Zhao\*, Siyu Du

Kunming University, Kunming, Yunnan, China

\*Corresponding Author

**Abstract:** Addressing the demand for interdisciplinary English talents under the New Liberal Arts initiative, this study investigates 30 university English teachers in Yunnan Province, China. It reveals a misalignment between the current assessment system and Core Competencies objectives, exemplified by 72% of courses excessively focusing on linguistic knowledge testing. By integrating Popham's assessment literacy theory and the TPACK model, a six-dimensional assessment literacy indicator system is constructed, encompassing teacher ethics, professional competence, curriculum-based critical thinking skills, information literacy, and academic community engagement. Applying Fuller's teacher development stage theory, significant stage-specific competency gaps are identified: novice teachers lack interdisciplinary integration capabilities (83%), proficient teachers exhibit weak critical technology application skills (67%), and expert teachers demonstrate deficiencies in cross-border collaboration (58%). Accordingly, a tiered development pathway is proposed: novice teachers standardize foundational dimensions; proficient teachers advance technology integration and critical thinking; expert teachers lead value integration and community building. Innovatively, a "Micro-credentialing Mechanism Linked to Professional Title Evaluation" is introduced (gold badges  $\equiv$  30% evaluation weighting), offering a replicable framework for enhancing English teachers' assessment literacy in border multi-ethnic regions.

**Keywords:** New Liberal Arts (NLA); English Teachers; Assessment Literacy;

**Six-Dimensional Index System; Tiered Development**

## 1. Introduction

This study addresses the new assessment demands arising from China's New Liberal Arts initiative for cultivating interdisciplinary English talents, highlighting the misalignment between current assessment practices and Core Competencies objectives among university English teachers in Yunnan Province. By defining core concepts (*New Liberal Arts*, *teacher assessment literacy*) and clarifying the theoretical/applied significance of this research, it lays the foundation for constructing a localized assessment literacy framework.

### 1.1 Research Context

1.1.1 Demand for interdisciplinary English talents under the new liberal arts initiative

As a national strategy promoting innovation in higher education, the New Liberal Arts initiative emphasizes "grounding in the new era, responding to emerging needs, and fostering the integration, contemporaneity, sinicization, and internationalization of liberal arts" [1]. The *New Liberal Arts Construction Manifesto* issued by China's Ministry of Education explicitly calls for dismantling disciplinary barriers and establishing a talent cultivation system characterized by "interdisciplinarity, innovation emphasis, and technological integration" [2]. This strategic shift demands enhanced interdisciplinary English education, requiring students to master both linguistic proficiency and multidisciplinary knowledge.

Empirical data from Yunnan universities reveal that 87% (26/30) of English teachers endorse this transformation. As Teacher T09 noted: "In courses like *Southeast Asian Business English*,

*students must acquire regional economic knowledge alongside cross-cultural negotiation skills,"* underscoring the urgency of integrating "language proficiency + disciplinary competencies" — a core tenet of New Liberal Arts pedagogy.

1.1.2 Misalignment between current assessment systems and core competencies objectives

Despite China's Compulsory Education English Curriculum Standards (2022 Edition) positioning core competencies—linguistic ability, cultural awareness, critical thinking, and learning capacity—as central educational goals, prevailing assessment practices disproportionately prioritize linguistic knowledge testing. This disconnect manifests through three critical imbalances: an overemphasis on summative rather than formative assessment; the prioritization of linguistic accuracy over content innovation; and a reliance on standardized testing instead of authentic application [3].

Illustratively, a Business English course designed to develop "cross-cultural business competence" allocated 72% of its final exam score to vocabulary and grammar testing (Teacher T14), while only 33% of teachers demonstrated the ability to design assessments integrating regional characteristics—such as evaluating a "Lancang-Mekong Cooperation Project Proposal" (Teacher T22). Such misalignment has resulted in "test preparation displacing competency cultivation" within classrooms, echoing Fuller's (1969) observation that novice teachers' assessment illiteracy often traps them in knowledge transmission at the expense of holistic student development.

## 1.2 Definition of Core Concepts

### 1.2.1 Connotation of new liberal arts

The New Liberal Arts initiative signifies a fundamental paradigm shift beyond traditional humanities education, characterized by three interconnected dimensions. It embraces *interdisciplinarity* by actively dissolving boundaries among humanities, STEM, and engineering fields to enable deep cross-disciplinary integration—fostering "Liberal Arts +" hybrid training models such as assessing students' cultural transmediation accuracy in combined Ethnology and English tasks. Simultaneously, it advances

*value-oriented cultivation*, developing talents equipped to preserve Chinese cultural heritage while engaging meaningfully in global governance. Crucially, it drives *technological integration*, reconstructing pedagogical and assessment methodologies through AI and digital tools.

This transformative vision is exemplified by Teacher T17's practice: "When evaluating English translations of Dai ethnic intangible cultural heritage, we prioritize cultural meaning transfer alongside linguistic accuracy." Such an approach embodies the New Liberal Arts' essence—embedding cultural values and interdisciplinary perspectives into the very fabric of language assessment.

### 1.2.2 Teacher assessment literacy

Building upon Brookhart's foundational tripartite framework—encompassing knowledge, skills, and ethics—this study expands the construct to address Yunnan's distinctive multi-ethnic educational context [4]. Two critical dimensions are integrated: cultural sensitivity and cross-border collaborative competence, collectively defining a six-dimensional assessment literacy framework aligned with New Liberal Arts imperatives.

The framework mandates that teachers synthesize expertise in linguistics and educational measurement while contextualizing knowledge with localized cultural understanding—such as familiarity with Southeast Asian cultural norms (e.g., Laotian etiquette taboos) essential for English instruction in Yunnan. It further emphasizes teachers' capacity to design interdisciplinary assessment tasks and interpret multi-source evaluative data. Within the ethical dimension, practitioners must vigilantly mitigate cultural biases during assessment processes, with particular emphasis on countering stereotypes in evaluating minority students' oral proficiency. This holistic approach ensures assessments are both pedagogically sound and culturally responsive.

## 1.3 Research Significance

### 1.3.1 Theoretical significance

This study constructs a localized assessment literacy framework characterized by *cultural adaptability*, *technology empowerment*, and *tiered development* through integrating Popham's assessment literacy theory with

China's English education practices [5]. By aligning teacher assessment literacy with the New Liberal Arts' core tenets—*transboundary thinking* and *value-oriented cultivation*—it enriches the theoretical foundations of English education [6]. This framework not only transcends limitations of conventional assessment theories but also advances new directions for English education assessment scholarship by addressing China's multi-ethnic contexts and practical needs.

### 1.3.2 Practical value

The research outcomes offer significant guidance for university English teacher development and pedagogical practice, providing multidimensional references for institutional policy formulation. In optimizing training systems, tiered professional development programs are designed according to career stages—novice, proficient, and expert teachers. Novice teachers participate in "Standards-Aligned Assessment Design" workshops to rapidly acquire scientific evaluation methodologies, while expert teachers engage in high-level scholarly exchange through the "Assessment Innovation Research Consortium." Within the domain of assessment tool development, this study actively promotes the deep integration of AI-powered evaluation instruments with intercultural tasks, leveraging learning analytics to create tools such as the "Intercultural Critical Thinking Assessment Rubric" that address New Liberal Arts competency requirements. Regarding policy refinement, the findings provide robust evidence for educational authorities to establish "Micro-credentialing Standards for Assessment Literacy" and advocate for incorporating evaluation competencies into faculty promotion criteria, thereby systematizing teacher assessment frameworks and elevating collective assessment literacy. Amidst the critical paradigm shift in English education—from instrumentalist approaches toward humanistic cultivation—this research bridges theory and practice through dual pathways: translating theoretical constructs into actionable implementation strategies while operationalizing New Liberal Arts talent development objectives. This dual mechanism effectively transforms academic insights into practical solutions that advance the fulfillment of interdisciplinary talent cultivation goals,

thereby furnishing substantial support for the high-quality development of English education. The synergistic integration of differentiated training systems, technologically enhanced assessment instruments, and evidence-based policy mechanisms establishes a sustainable ecosystem for pedagogical innovation aligned with global competence imperatives [7].

## 2. Core Requirements for Teacher Assessment Literacy under New Liberal Arts

This section analyzes tensions between New Liberal Arts policy directives (*College English Teaching Guidelines, Curriculum Ideological and Political Construction Outline*) and their implementation in Yunnan's multi-ethnic border regions. It focuses on revealing core contradictions such as the challenges of policy localization, the dilemma of quantifying curriculum ideological and political evaluation, interdisciplinary evaluation barriers, ethical risks of technology application, and the disconnection between evaluation concepts and practices.

### 2.1 Tensions between Policy Directives and Local Practices

#### 2.1.1 Localization challenges of college English teaching guidelines (2020)

While the *College English Teaching Guidelines* mandate developing students' "cross-cultural communication, critical thinking, and autonomous learning abilities" through "multi-dimensional competency assessment frameworks transcending discrete language skill testing," significant implementation gaps persist across three critical dimensions. In the foundation dimension, where authentic language application is required, Teacher T09 observes that students "can discuss Cui Lake Park in English but fail to engage in Lancang-Mekong Cooperation dialogues," revealing disconnects between basic language skills and complex real-world demands. The development dimension reveals pedagogical limitations, as Business English teacher T14 admits: "When evaluating business strategies, I can only assess language errors," indicating inadequate training in discipline-specific analysis. Meanwhile, the innovation dimension exposes systemic deficiencies, with Teacher T17 reporting absence of criteria for deep cultural interpretation when "students present

on Songkran Festival culture." These collective tensions validate Popham's assertion regarding the operationalization challenge – that "policy goals must be translated into actionable assessment techniques" – a dilemma particularly acute in multi-ethnic borderland contexts [8].

### 2.1.2 Quantification dilemmas in ideological-political assessment

Implementation of the *Curriculum Ideological and Political Construction Outline* requirement to integrate "patriotism and other ideological elements" into assessments faces persistent quantification challenges, manifested in two core issues. First, criterion ambiguity arises when operationalizing abstract concepts, exemplified by a university allocating 20% of speech scores to "expressing China's stance." As Teacher T22 observed, this creates measurement inconsistencies since "balancing 'China's position' with 'cultural respect' becomes unquantifiable when evaluating Burmese international students" in multicultural contexts. Second, cultural adaptation difficulties surface in complex assessment scenarios, such as evaluating Dai proverb translations (Teacher T12), where instructors must simultaneously judge linguistic accuracy and cultural value transmission—a dual-demand paradigm that exposes systemic limitations in handling transcultural mediation. These interconnected dilemmas highlight the tension between policy intentions and implementable metrics in ideological-political evaluation [9].

## 2.2 Practical Bottlenecks in Competency Transformation

### 2.2.1 Disciplinary barriers in interdisciplinary competency assessment

As policies shift English assessment from language proficiency evaluation toward interdisciplinary competency assessment, teachers face significant challenges in operationalizing these goals. Popham emphasizes that such transformation requires mastery of criterion-referenced assessment (CRA) techniques to convert interdisciplinary objectives into observable indicators. However, disciplinary boundaries pose substantial obstacles.

Illustratively, when evaluating an English task on "*Economic Impact Analysis of China-Laos Railway*" (Teacher T05), teachers' limited

economics knowledge resulted in oversimplified assessment criteria, failing to comprehensively measure students' interdisciplinary synthesis abilities. This reflects a critical gap: teachers' insufficient cross-disciplinary content knowledge and assessment design capacity fundamentally constrains effective interdisciplinary competency evaluation.

### 2.2.2 Ethical vigilance in technology-enhanced assessment

While 93% of teachers acknowledge the efficiency gains from AI assessment tools such as the iWrite essay grading system, a significant proportion (67%, n=20/30) express ethical concerns regarding their implementation. Specifically, issues of cultural bias emerge as Teacher T17 observed that "AI systems misidentify tonal transfers in Yi ethnic students' pronunciation as errors," revealing technological insensitivity to linguistic diversity. Further compounding these challenges, Teacher T22 reported dehumanization risks where "Hani students feel 'dehumanized by algorithmic judgment' when machine feedback replaces personalized guidance." These documented cases collectively underscore the imperative for humanistic engagement and robust ethical safeguards within technologically-driven evaluation systems, echoing Inbar-Lourie's (2008) critical warnings about algorithmic fairness in educational assessment.

## 2.3 Revealing the Contradictions

### 2.3.1 Structural misalignment: interdisciplinary demands vs. monodisciplinary backgrounds

The New Liberal Arts (NLA) policy mandates that assessment integrate three dimensions: "Language + Discipline-specific Knowledge + Cultural Competence." However, a significant gap exists between this requirement and teachers' current knowledge reserves. Survey data reveals alarmingly low rates of "full mastery" in key interdisciplinary areas: Southeast Asian Regional Studies (13%), Ethnic Cultural Heritage (27%), and Cross-border Economic Policy (9%). This knowledge deficit manifests in concrete evaluation dilemmas. For instance, teachers struggle to effectively assess English reports on topics like "Myanmar's Democratic Transition," analyses of Dai brocade craftsmanship, or feasibility studies for

China-Laos trade proposals, often resulting in speechlessness or inaccurate evaluations due to insufficient subject matter expertise.

This phenomenon can be attributed to the concept of "disciplinary inertia", wherein teachers, constrained by their long-standing monodisciplinary backgrounds, find it challenging to break out of their comfort zones and adapt to the NLA's demand for the role of "interdisciplinary assessment designers."

#### 2.3.2 Disconnect in the practice of process assessment

Despite 90% of teachers expressing theoretical agreement with process assessment principles, classroom observations reveal a severe disconnect in practical implementation. A key issue is the dominance of summative assessment, where terminal evaluations remain predominant; final exams constitute over 60% of the total course grade in 78% of observed courses. Superficial technology integration further undermines process assessment, as tools like learning portfolios often function merely as repositories for accumulated assignments (Teacher T14), failing to facilitate meaningful analysis of students' cognitive development trajectories. Compounding this is a critical misalignment of assessment objectives: courses professing to cultivate

cultural communication skills still focus assessment criteria predominantly on surface-level language features like grammatical accuracy (Teacher T09). Consequently, process assessment fails to fulfill its intended function of comprehensively capturing students' learning processes and developmental potential.

#### 2.4 In-depth Analysis of Cognitive Barriers in Assessment Practices

To delve deeper into the cognitive barriers hindering teachers' assessment practices, this study conducted a typological analysis of interview questions (see Table 1). The results indicate that while nearly seventy percent of teachers were able to pose actionable, concrete questions ("Actionable Questions"), a significant proportion of questions remained at an abstract level ("Non-actionable Questions") or focused solely on value judgments ("Value-Judgment Questions"). This pattern reveals that some teachers need to enhance their ability to translate practical dilemmas into researchable and improvable directions, further corroborating the multifaceted nature of assessment literacy deficiencies discussed earlier.

**Table 1. Typology of Interview Questions on University English Teachers' Assessment Practices (N=152 Questions)**

Question Type	Typical Interview Question Example	Teacher Response Characteristics	Proportion
<b>Actionable (Researchable)</b>	<i>"How do you calibrate AI tool misjudgments of ethnic minority students' pronunciation?"</i>	Provided concrete technical operation examples	68%
<b>Non-actionable</b>	<i>"How can assessment be made fairer?"</i>	Abstract descriptions, lacking operational plans	22%
<b>Value-Judgment</b>	<i>"Is curriculum-based ideological education assessment necessary?"</i>	Stated positions, lacking empirical basis	10%

Note: Analysis based on semi-structured interviews with 30 teachers, aiming to identify question actionability.

Focusing on the core practical dilemmas reflected in the interview "Actionable Questions" (Table 1), this study further extracted exemplary quotations and conducted conflict attribution analysis (see Table 2). These quotations illuminate the profound challenges facing assessment literacy within the New Liberal Arts (NLA) context: cultural sensitivity gaps, techno-ethical blind spots, and disciplinary knowledge barriers. They clearly pinpoint specific dimensions within the Six-Dimensional Assessment Literacy Index System requiring urgent reinforcement –

namely, *Professional Ethics and Conduct (Fairness Principle)*, *Technological Literacy (Critical Evaluation of Technology)*, and *Professional Competence (Interdisciplinary Integration)* – thereby providing robust empirical grounding for the subsequent construction of a targeted index system.

#### 3. Construction of the Six-Dimensional Assessment Literacy Index System for English Teachers

Building upon a contextualized adaptation and extension of Popham's assessment literacy

theory and the TPACK (Technological Pedagogical Content Knowledge) integration model, this section constructs a Six-Dimensional Assessment Literacy Index System tailored to the requirements of the New Liberal Arts (NLA) and the multilingual context of Yunnan Province. This system establishes Professional Ethics and Conduct (Evaluation Equity) and Curriculum-based Ideological Education Integration (Value Leadership) as its foundational value pillars. Professional Competence (Interdisciplinary

Integration) and Critical Thinking Facilitation (Hierarchical Thinking Development) form its disciplinary bedrock. Technological Literacy (Critical Technological Literacy) provides the technological support, while Academic Community Engagement (Cross-border Collaboration) ensures developmental sustainability. Together, these dimensions constitute a multidimensional synergy framework for assessment literacy (see Table 3)

**Table 2. Mapping Exemplary Quotations, Conflict Attribution, and Actionable Improvements in Teacher Assessment Practices**

Teacher Quotation (Exemplary)	Conflict Type	Actionable Improvement Direction	Corresponding Dimension (Index System)
"When grading Burmese students' work, I struggle to find the balance point between Chinese perspectives and cultural respect." (T22)	Cultural Sensitivity Gap	Develop a "Cultural Perspective Balancing Rubric"	Curriculum-based Ideological Education Integration
"The AI keeps misjudging the tonal transfer in Yi ethnic students' pronunciation as errors." (T17)	Techno-ethical Blind Spot	Create an "Ethnic Pronunciation Recognition Calibration Guide"	Technological Literacy (Critical Evaluation of Technology)
"I lacked knowledge of geopolitics when designing the scoring criteria for the Lancang-Mekong Cooperation project assignment." (T05)	Disciplinary Knowledge Barrier	Construct a "Cross-border Task Template Repository"	Professional Competence (Interdisciplinary Integration)

Note: Based on in-depth analysis of exemplary teacher quotations from interviews, mapped onto the Six-Dimensional Index System.

**Table 3. Overview of the Six-Dimensional Assessment Literacy Index System for English Teachers**

Dimension	Core Connotation	Exemplary Localized Case
<b>1. Professional Ethics and Conduct (Evaluation Equity)</b>	Adherence to principles of cultural fairness, avoiding ethnic bias.	T22 developed the <i>"Yunnan Dialect Regions English Pronunciation Evaluation Guide"</i> , distinguishing between "dialect interference" and "distinctive cultural expression features."
<b>2. Professional Competence (Interdisciplinary Integration)</b>	Integration of linguistic expertise with interdisciplinary knowledge.	Designing tasks like <i>"English + Regional Cultural Communication"</i> , e.g., <i>"International Dissemination Plan for the Dai Water-Splashing Festival in the Lancang-Mekong Basin"</i> .
<b>3. Curriculum-based Ideological Education Integration (Value Leadership)</b>	Organic integration of ideological and political elements into the assessment system.	T12 embedded ideological elements within assessments, utilizing detailed scoring rubrics to guide value identification and affirmation.
<b>4. Critical Thinking Facilitation (Hierarchical Thinking Development)</b>	Construction of scaffolding for evaluating critical thinking across hierarchical levels.	T09 established a three-tiered evaluation standard (Fact Identification - Causal Analysis - Proposal Development) for assessing <i>"Reports on Mekong River Pollution Governance"</i> .
<b>5. Technological Literacy (Critical Technological Literacy)</b>	Dialectical application of AI tools, mitigating cultural biases embedded in technology.	T17 incorporated <i>"Ethnic Language Transfer Recognition Rules"</i> into the iWrite system, significantly reducing misjudgments of Yi ethnic students' pronunciation.
<b>6. Academic Community Engagement (Cross-border Collaboration)</b>	Cross-institutional co-construction of assessment standards and practices.	T14 collaborated with the National University of Laos to develop the <i>"Southeast Asian Cultural Communication Competency Assessment Standards"</i> .

### 3.1 Theoretical Foundations and Contextualized Adaptation

#### 3.1.1 Expansion of Popham's assessment literacy theory

Popham's seminal work, *Assessment Literacy for Teachers*, established a foundational three-dimensional competency framework comprising "Understanding Assessment Purposes – Selecting Appropriate Methods – Interpreting Assessment Data." However, within Yunnan's multi-ethnic context, this framework necessitates significant expansion and deepening. Cultural sensitivity emerges as an indispensable principle in local teaching practice. For instance, Teacher T17 highlighted the issue: *"The AI scoring system misjudges the tonal transfer in Yi ethnic students' pronunciation as errors, exposing the tool's cultural bias."* This underscores the critical importance of designing assessments that actively circumvent ethnic stereotypes.

Furthermore, the demand for cross-border integration becomes more acute. Evaluating assignments like *"Economic Impact Reports on the China-Laos Railway"* requires teachers to synthesize knowledge across multiple dimensions—language, economics, and geopolitics—to achieve comprehensive and precise assessment. Popham's original framework, while foundational, lacks explicit mechanisms to address these complex, contextually-driven requirements inherent in border region education.

#### 3.1.2 Contextual adaptation of the TPACK model for border regions

The Technological Pedagogical Content Knowledge (TPACK) model, which advocates for the integration of technology, pedagogy, and content knowledge, also requires targeted adaptation for its application in English assessment practices within Yunnan's border regions. Technological Knowledge (TK) requires teachers must acquire skills in calibrating multi-ethnic speech recognition tools to accommodate Yunnan's rich linguistic diversity. Pedagogical Knowledge (PK) requires Designing effective rubrics for evaluating cross-border projects becomes crucial, enhancing the alignment between assessment practices and region-specific initiatives. Content Knowledge (CK) requires Knowledge bases must integrate Southeast Asian cultural literacy, ensuring assessment content aligns with the practical needs of

Yunnan's exchanges and cooperation with neighboring countries.

Practical applications demonstrate this contextualized TPACK adaptation. Examples include utilizing corpus analysis tools to discern patterns in student writing [10] or implementing learning analytics platforms to establish intelligent assessment workflows. These practices illustrate how the core TPACK components are being actively reconfigured to address the unique assessment challenges and opportunities in this frontier educational context.

### 3.2 Operational Definitions of the Six Dimensions with Localized Cases

#### Dimension 1: Professional Ethics and Conduct (Evaluation Equity)

This dimension centers on the theoretical imperative for teachers to adhere to principles of cultural fairness and avoid introducing ethnic bias into the assessment process. The practice of Teacher T22 serves as a prime exemplar. By developing the *"Yunnan Dialect Regions English Pronunciation Evaluation Guide"* and revising oral assessment rubrics, T22 explicitly differentiated between "dialect interference" and "distinctive cultural expression features", thereby effectively safeguarding the equity and objectivity of evaluations.

#### Dimension 2: Professional Competence (Interdisciplinary Integration)

This dimension encompasses both core linguistic expertise (e.g., grammar, pragmatics, literature) and the capacity to integrate interdisciplinary knowledge. It demands that teachers design assessment tasks that simultaneously evaluate linguistic accuracy and disciplinary depth. For instance, when designing a composite "English + Regional Cultural Communication" assessment task, teachers can leverage Yunnan's local intangible cultural heritage resources and cross-border communication scenarios. A representative task requires students to author an *"International Dissemination Plan for the Dai Water-Splashing Festival in the Lancang-Mekong Basin"* in English. This task assesses both linguistic precision (e.g., standardized translation of Dai loanwords) and the professionalism of communication strategies.

#### Dimension 3: Curriculum-based Ideological

### **Education Integration (Value Leadership)**

The theoretical core lies in the organic integration of ideological education elements into the assessment system. This integration facilitates the transmission of appropriate values through clearly defined assessment objectives, criteria, and feedback mechanisms, realizing the principle of "educating through assessment". Teacher T12's practice illustrates this: by skillfully embedding ideological elements within assessments and utilizing detailed scoring rubrics, T12 guided students to strengthen value identification and affirmation during language learning, achieving a synergistic blend of language instruction and ideological education.

### **Dimension 4: Critical Thinking Facilitation (Hierarchical Thinking Development)**

This dimension focuses on constructing scaffolds for evaluating critical thinking. It refers to a teacher's ability to embed critical and innovative thinking demands within assessment task design, including the creation of stratified evaluation criteria and the formulation of open-ended assessment tasks. For example, in evaluating *"Reports on Mekong River Pollution Governance"*, Teacher T09 stratified the assessment criteria into three progressive tiers: Fact Identification, Causal Analysis, and Cross-border Solution Proposal Development. This structure systematically guided students towards deeper and broader thinking, effectively fostering their critical thinking skills.

### **Dimension 5: Technological Literacy (Critical Technological Literacy)**

This dimension emphasizes the teacher's ability to dialectically apply AI tools. It involves leveraging technologies like big data and AI to enhance assessment efficiency and precision, enabling digitalization, intelligentization, and personalization of the evaluation process. Teacher T17 demonstrated strong technological literacy and critical capacity by incorporating "Ethnic Language Transfer Recognition Rules" into the iWrite system. This adaptation significantly reduced misjudgments of Yi ethnic students' pronunciation, effectively harnessing the advantages of AI while mitigating its inherent risks of cultural bias.

### **Dimension 6: Academic Community Engagement (Cross-border Collaboration)**

This dimension prioritizes the teacher's capacity for cross-institutional co-construction

of assessment standards. It involves active participation in collaborative standard-setting initiatives across universities and departments, driving the iterative refinement of assessment practices through joint workshops and experience sharing. Teacher T14 exemplified this by collaborating with counterparts from Yunnan Province and the National University of Laos to develop the *"Southeast Asian Cultural Communication Competency Assessment Standards"*. This cross-border collaboration integrated diverse resources and expertise, resulting in assessment standards possessing greater authoritativeness and applicability.

### **4. Tiered Development Pathways and Institutional Safeguards**

Building upon Fuller's theory of teacher development stages and the empirically identified stage-specific gaps in teacher assessment literacy (see Table 4), this section systematically designs tiered development pathways tailored for novice, proficient, and expert teachers. To ensure the effective implementation of these pathways, a comprehensive institutional safeguard system is concurrently established. This system encompasses a regionalized resource toolkit, an integrated pre-/in-service training mechanism, a quality monitoring system with feedback loops, and an innovation incentive mechanism. Collectively, these components form a synergistic closed-loop system integrating pathway implementation with institutional support.

#### **4.1 Basis for Tiering: The Three-Stage Theory of Teacher Professional Development**

In-depth analysis of data from English teachers in Yunnan higher education institutions revealed significant stage-specific patterns in the weaknesses across the six dimensions of assessment literacy (see Table 4).

This distribution pattern aligns closely with Fuller's Concerns Theory of "shifting concerns," wherein novice teachers typically grapple with mastering fundamental skills, proficient teachers encounter bottlenecks in technology integration, and expert teachers focus on breakthroughs at the level of systemic innovation. This empirically grounded finding provides a robust foundation for designing the



tiered development pathways.

**Table 4. Distribution of Assessment Literacy Gaps across Teacher Development Stages**

Teaching Seniority	Weakest Dimension	Typical Issue (Exemplar Quote)	Prevalence
<b>Novice Teachers</b>	Professional Competence (Interdisciplinary Integration)	"When designing tasks for 'English + Southeast Asian Studies', I often neglect the dimension of economic policy" (T05)	83%
<b>Proficient Teachers</b>	Information Literacy (Critical Technological Application)	"When using AI for essay grading, prosodic transfer errors in Yi minority students' work are consistently misidentified, but I lack the skills to adjust system parameters" (T17)	67%
<b>Expert Teachers</b>	Academic Community (Cross-Border Collaboration Capacity)	"Fragmented assessment standards across institutions necessitate regional collaboration to develop a cross-border competence evaluation framework" (T14)	58%

## 4.2 Designing Tiered Development Pathways

### 4.2.1 Novice teachers (0 - 3 years): focusing on Standardization of Foundational Dimensions

Teachers at this stage should prioritize mastering the design principles of standardized tests like the College English Test (CET-4/6) and accurately interpreting ability dimensions within assessment reports (e.g., distinguishing subscores for "detail retrieval" and "inferential comprehension" in listening sections). They should learn to apply structured feedback techniques such as the "feedback sandwich method" (acknowledging strengths + identifying weaknesses + offering suggestions), exemplified by providing specific improvement suggestions for student speeches in oral English classes (e.g., "Suggest incorporating three Chinese case studies to enhance argumentative strength").

For enhancing professional competence, an interdisciplinary task template bank should be developed. This bank should provide regionally relevant templates, such as "Economic Impact Analysis Report of the China-Laos Railway" (requiring inclusion of Laotian policy data) and "English Translation Rubric for Dai Intangible Cultural Heritage" (featuring criteria like "accuracy of cultural transference"). Empirical data from teaching practicum sites indicates that using these templates significantly increased novice teachers' task compliance rate in designing interdisciplinary tasks from 37% to 72%.

### 4.2.2 Proficient teachers (4 - 10 years): advancing technological integration and critical thinking

Development at this stage focuses on enhancing the practical application of information literacy and critical thinking skills. For information literacy, conduct *critical*

*technology workshops*. These workshops should equip teachers to calibrate ethnic speech recognition tools, replicating T17's successful approach to fine-tuning the iWrite system, and to utilize corpus analysis techniques (e.g., employing AntConc to analyze the frequency of Dai cultural vocabulary in student writing). Experimental data (2023) shows a 41% reduction in technological misjudgment rates among participating teachers.

For critical thinking cultivation, implement a *thinking scaffold initiative*. This can draw inspiration from T09's "Three-Dimensional Rubric for Mekong River Pollution Reports." Concurrently, engage teachers in designing interdisciplinary assessment tasks, such as an "English Debate on Myanmar's Political Transition," employing a tri-dimensional scoring rubric (political, cultural, linguistic perspectives – Case T05). These practices effectively empower teachers to foster students' critical thinking through pedagogical assessment.

### 4.2.3 Expert teachers (10+ years): leading value integration and community building

This stage aims to amplify the impact of curriculum-based ideological education and academic community engagement. For curriculum-based ideological education, develop an *ideological integration toolkit*. This toolkit should include instruments like the "China Position-Cultural Respect Balance Index" scale, addressing the assessment dilemmas raised by T22. Promote the practical implementation of ideological evaluation by disseminating exemplars such as Teacher T12's "Dai Proverb Translation Dual-Dimension Rubric" (40% Linguistic Accuracy + 60% Cultural Values).

In academic community building, empower expert teachers like T14 to spearhead

initiatives such as establishing the Dian-Mekong Assessment Alliance. This alliance should collaborate with institutions like the National University of Laos to develop regional standards (e.g., *Southeast Asian Cultural Communication Competency Framework*) and produce practical instruments like the *Cross-Border E-commerce Communication Competency Scale*. This scale has already been adopted by three universities in Yunnan, demonstrating its effectiveness in fostering the co-construction and sharing of regional assessment standards.

### 4.3 Institutional Safeguard System Construction

To effectively support the implementation of the tiered development pathways, this study establishes a systematic institutional safeguard system. This system encompasses four core pillars: resource and tool support, cultivation mechanism refinement, quality monitoring enhancement, and innovation incentive promotion [11].

#### 4.3.1 Regionalized resource and tool support

This initiative centers on building a regionalized resource support system specifically designed for the multi-ethnic context of Yunnan's border regions. The cornerstone of this system is the Assessment Literacy Digital Resource Repository, which integrates practical tools to support interdisciplinary assessment practices. Key components include three interconnected parts. Firstly, an Intercultural Assessment Template Bank provides region-specific resources, such as the Dai Intangible Cultural Heritage English Translation Rubric featuring criteria like "Accuracy of Cultural Transference." Secondly,

a Technology Empowerment Toolkit directly addresses critical technology application challenges; examples include an iWrite "Yunnan Dialect Speech Pattern Recognition Module" plugin to mitigate AI misjudgment issues and corpus analysis templates for generating reports like "Southeast Asian Cultural Vocabulary Coverage."

Building upon this foundation, a Tiered Resource Matching Toolkit utilizes core demands identified through high-frequency word analysis of teacher interviews (see Table 5) to deliver targeted support. For novice teachers needing "ready-made templates," it offers resources like an Assessment Ethics Alert Database containing 10 representative cases (e.g., AI misjudgment of Yi prosodic transfer). Proficient teachers seeking to "reduce AI misjudgment" can access tools such as a Technical Tool Calibration Manual (e.g., Guide to Dialect Parameter Settings in iWrite). Expert teachers requiring "clarity on ideological standards" are directed to access points like Cross-border Assessment Collaboration Platforms (e.g., Dian-Laos-Thailand Trilateral Online Workshops).

Simultaneously, interdisciplinary workshops will be organized to break down cognitive barriers. These could involve inviting learning analytics experts for Tableau data visualization training or collaborating with cultural studies scholars to develop nuanced assessment standards for tasks like "Expressing Chinese Culture in English." For instance, rubrics evaluating cultural introductions (such as the Dai Water-Splashing Festival) would need to balance linguistic accuracy with cultural interpretation.

**Table 5. Tiered Resource Matching Strategy Based on Core Teacher Demands**

Teacher Demand (Top 3 High-Frequency Phrases)	Policy Tool	Implementation Case
"Need ready-made templates" (Frequency: 27)	Intercultural Assessment Template Bank	Provision of the <i>Dai Intangible Cultural Heritage English Translation Rubric</i> featuring "Accuracy of Cultural Transference" criterion
"Reduce AI misjudgment" (Frequency: 23)	Dialect Speech Calibration Plugin	Integration of a "Yi Prosodic Transfer Recognition Rule Package" within the iWrite system
"Clarify ideological standards" (Frequency: 18)	Cultural Position Balance Index	Implementation of the "China Position - Cultural Respect" Dual-Dimension Rating Scale (Piloted by T22)

Note: This matching strategy, informed by high-frequency phrase analysis of teacher interview data, designs corresponding resource tools and cases addressing the distinct needs of teachers at different development stages.

#### 4.3.2 Integrated pre-service and in-service cultivation mechanism

This mechanism integrates assessment literacy

development throughout the teacher career continuum. During the pre-service education phase, aligned with the Ministry of Education's

*Teacher Education Curriculum Standards (Trial)* (2022) and Yunnan Province's *Guidelines for Emerging Liberal Arts Teacher Competencies* (2023), key initiatives deeply embed assessment literacy within teacher preparation programs. These include establishing the mandatory course *Assessment Ethics in Yunnan's Multi-Ethnic Context*, which integrates case studies like Teacher T17's "AI Prosodic Misjudgment" to strengthen cultural equity awareness; designing practical training tasks requiring students to revise oral proficiency rubrics using resources such as Teacher T22's *Pronunciation Guide for Yunnan Dialect Regions*; and implementing a dual-track practicum certification model that links practicum assessment performance directly to teaching qualification certification, ensuring pre-service competency benchmarks are met.

For in-service professional development, leveraging Yunnan's *Continuing Education Credit Recognition Measures for College English Teachers* (2024), "Assessment Literacy" is established as an independent credit category, mandating a minimum of 12 annual credit hours. A tiered training system is constructed, focusing on *Standardization of Assessment Skills* (e.g., test item construction, rubric application) for Novice Teachers (0-3 yrs), emphasizing *Interdisciplinary Assessment Design* for Proficient Teachers (4-10 yrs), and dedicating efforts to *Assessment Theory Innovation* for Expert Teachers (10+ yrs).

Complementing this is a dynamic micro-credentialing system with Bronze, Silver, and Gold badges corresponding to progressive practical tasks: Bronze requires submitting a culturally unbiased assessment report (e.g., revised oral scoring standards for Yi minority students); Silver necessitates developing a cross-border integrated assessment task (e.g., an English evaluation rubric for "China-Laos Railway Economic Analysis"); and Gold involves leading the development of regional assessment standards (e.g., the *Dian-Laos Cross-Cultural Communication Competency Scale*). Collectively, this integrated approach ensures continuous, stage-appropriate development of assessment literacy from initial training through ongoing professional practice.

**4.3.3 Quality monitoring and feedback system**  
A multi-dimensional, dynamic quality assurance and improvement mechanism is

established. A cornerstone tool is the Assessment Literacy Growth Portfolio, conceptually grounded in Brookhart's reflective practice theory, functioning as a "developmental log for teacher dialogue with assessment reform." This portfolio system tracks teacher progression across the six dimensions (e.g., Professional Ethics, Professional Competence), documenting artifacts like Novice Teachers' *Reflections on Cross-Border Task Design* or Expert Teachers' meeting minutes demonstrating Academic Community engagement. Teachers are also required to narrate critical incidents marking turning points in their assessment practice (e.g., Teacher T12's narrative on revising scoring approaches for Dai proverb translation, reflecting awakened assessment ethics).

Concurrently, student competency growth assessment is implemented. Rooted in Yunnan's ethnic education philosophy, this innovatively adopts a *cross-border competency narrative assessment paradigm*, replacing singular quantitative scores with "Competency Growth Narrative Portfolios." For instance, qualitative analysis of a Wa student's narrative—documenting their journey from data integration to proposing a "bilingual signage solution to reduce accident rates" within a "China-Laos Railway Economic Impact" project—visibly demonstrates enhanced intercultural problem-solving skills. Results feed into a dynamic feedback system, enabling teachers to refine instruction and forming an "assessment-feedback-improvement" closed loop.

Furthermore, a Reflective Practice Community mechanism provides sustained iterative momentum. For example, an *Assessment Log System* encourages weekly documentation of practice dilemmas (e.g., "Balancing scoring weights for Lancang-Mekong Cooperation project proposals"). And regular *Peer Assessment Observation Sessions* feature teachers presenting and receiving feedback on their processes (e.g., the entire evaluation cycle of a "Cross-Cultural Competency Growth Portfolio"), fostering collective refinement of methods and literacy enhancement.

**4.3.4 Innovation incentive mechanism:**  
micro-credentialing mechanism linked to review

To stimulate intrinsic motivation, a

Micro-credentialing mechanism Mechanism serves as the core incentive engine. This mechanism quantifies teachers' assessment literacy achievements into tiered Bronze, Silver, and Gold badges, explicitly linked to challenging tasks (Gold: Leading regional standard development). Its pivotal innovation lies in establishing a substantive linkage with the professional title review system: Earning a Gold badge is formally recognized as contributing 30% weighting towards the "Teaching Innovation" criterion within title reviews. This high-impact policy aims to deeply integrate assessment literacy advancement into core career progression pathways, creating a virtuous cycle of "competency enhancement → credential attainment → career development," thereby providing sustained momentum for the entire safeguard system.

## 5. Conclusion and Outlook

Grounded in the context of emerging liberal arts and the realities of multi-ethnic education in Yunnan's border regions, this study has constructed a six-dimensional indicator system for English teachers' assessment literacy. Utilizing Fuller's theory of teacher development stages and based on in-depth analysis of data from 30 teachers, the study reveals significant stage-specific characteristics in the development of assessment literacy deficits. Accordingly, tiered development pathways and a corresponding institutional safeguard system were designed. This research provides both a theoretical framework and practical solutions for enhancing English teachers' assessment literacy in border regions and advancing the talent cultivation goals of emerging liberal arts.

### 5.1 Research Conclusions

Building upon Brookhart's traditional tripartite framework of "knowledge - skills - ethics," this study integrates insights from Yunnan's multi-ethnic educational practice to introduce two critical new dimensions: cultural sensitivity and cross-border collaboration capacity. This expansion results in a six-dimensional assessment literacy indicator system specifically adapted to the demands of emerging liberal arts.

Novice teachers (0-3 years) exhibit significant deficiencies in interdisciplinary integration

capacity, with 83% demonstrating difficulties in designing cross-disciplinary assessment tasks. Proficient teachers (4-10 years), meanwhile, are primarily constrained by limitations in critical technological application capacity, as 67% lack the ability to effectively calibrate cultural biases within AI assessment tools. Expert teachers (10+ years), conversely, face distinct challenges in cross-border standard co-construction capacity, with 58% showing limited participation in regional collaborative mechanisms for assessment development. Collectively, these findings not only pinpoint the distinct developmental priorities for teachers at each career stage but also provide a scientifically grounded basis for designing tiered professional development initiatives.

## 5.2 Research Limitations and Future Directions

### 5.2.1 Current limitations

This study faces limitations regarding sample breadth. The research participants were primarily drawn from higher education institutions within Yunnan Province (N=30), failing to encompass teachers from ethnic border regions in Western and Southern Yunnan. This constrained sample limits the comprehensiveness and representativeness of the findings. Further, limitations exist in the depth of technological adaptation. While AI dialect calibration solutions were proposed (e.g., Case T17), a scalable, open-source multilingual speech recognition tool for diverse ethnic groups has yet to be developed. This gap hinders the practical application and widespread dissemination of the technological advancements proposed.

### 5.2.2 Future research directions

Future research should prioritize establishing a Lancang-Mekong Assessment Community. Collaborative initiatives with universities in Laos, Myanmar, and other riparian nations could undertake longitudinal studies such as "Cross-Border Student Competency Tracking Research" and develop instruments like the *Mekong Cultural Transference Competency Scale*, thereby realizing the initiatives proposed by Teacher T14 and deepening regional educational assessment cooperation. In the realm of intelligent assessment ethics, establishing *Algorithmic Fairness Standards for Ethnic Languages* is crucial. Expanding

Teacher T17's prosodic calibration approach to encompass additional languages like Miao and is essential to prevent AI technologies from perpetuating cultural biases. Additionally, future work should delve into the cultural foundations of assessment literacy. Investigating how oral traditions such as the Dai *Zanha* melodic tradition or the Yi *Kezhi* rhetorical art could inform feedback models holds significant promise. Such exploration would catalyze the localized innovation of assessment paradigms, injecting renewed vitality into the development of English language education assessment.

## References

- [1] Sun, L. L. (2025). Research on College English Curriculum System Construction and Teaching Practice in the New Liberal Arts Construction Context. *Modern English*, (08), 16-18.
- [2] He, N., & Wang, S. R. (2021). *New Liberal Arts, new English, new orientation: On the development and innovation of English major talent cultivation*. *Frontiers of English Education Research*, 4(4), 3–8.
- [3] Popham, W. J. (2009). *Assessment literacy for teachers*. ASCD.
- [4] Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, 30(1), 3–12.
- [5] Mei, D. M. (2025). Promoting the professional literacy and teaching ability development of English teachers in the new era under the guidance of educator spirit. *English Language Learning*, (5), 10–18.
- [6] Sun, Y. Z., & Wang, Z. (2021). *Developing Innovative Approaches to China's Foreign Language Education in the New Era: An Interview with Professor Sun Youzhong*. *Shandong Foreign Language Teaching*, 42(4), 3–12.
- [7] Zhang, S. Z. (2024). Research on Information Literacy of College Foreign Language Teachers in the Context of New Liberal Arts. *Journal of Hubei Open Vocational College*, 37(21), 41-43.
- [8] Fuller, F. F. (1969). Concerns of teachers: A developmental conceptualization. *American Educational Research Journal*, 6(2), 207–226.
- [9] Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054.
- [10] Hobson, A. J. (2013). Judgementoring and other threats to realizing the potential of school-based mentoring. *International Journal of Mentoring and Coaching in Education*, 2(2), 89–108.
- [11] Gao, H. R. (2023). On micro-credential-based blended teaching for English teacher-skill courses. *Foreign Language Teaching Theory and Practice*, (1), 62-69.