Journal of Education and Educational Research
ISSN: 2957-9465 | Vol. 21, No. 2, 2026

Learning Evidence and Assessment Design for Traditional Chinese Medicine Cultural Education within an Inheritance-Integration-Innovation Framework: Digital Learning Tasks Using Central Plains Resources

Ran Zhang *, Tao Ruan

The First Affiliated Hospital of Henan University of Science and Technology, Luoyang, Henan, China
*** Corresponding author:** Ran Zhang

Abstract: Assessing a digital exhibit in traditional Chinese medicine (TCM) culture requires more than judging the finished slides. Teachers also need to know how students selected sources, interpreted them, and responded to criticism. This paper proposes an assessment design for university general education based on the inheritance-integration-innovation framework, evidence-centered design, and formative assessment. Six indicators specify the evidence to collect across three 90-minute units using Yaofang Cave materials from the Longmen Grottoes. The units move from source verification and annotation to exhibit design and revision. Individual explanations accompany group submissions, and an analytic rubric addresses cultural interpretation, digital resource use, and justified design decisions. A separate Funiu Mountain task examines performance with unfamiliar materials before feedback is provided. Subsequent revision is assessed separately. Cultural understanding and reasoned value reflection remain distinct from cultural identification throughout the design. The proposal sets out a procedure for classroom assessment; its feasibility and measurement properties still require testing.

Keywords: Traditional Chinese Medicine Cultural Education; Three-Dimensional Framework; Central Plains Resources; University General Education; Learning Evidence.

1. Introduction

A student explaining a historical medical inscription must decide what the source says and what can reasonably be inferred from it. Presenting that explanation digitally adds decisions about tools and the needs of an unfamiliar audience. In university TCM culture courses, these concerns connect with the inheritance-integration-innovation framework [1]. Applying the framework to assessment requires teachers to distinguish historical understanding from technical skill and originality of presentation. A polished exhibit may display all three, but its appearance alone cannot show how each developed.

Central Plains materials make this problem concrete. Yaofang Cave at Longmen preserves historical medical inscriptions, yet the museum account notes that a surviving rubbing is only partly legible [2]. Materials from Funiu Mountain place regional medicinal plant knowledge alongside accounts of contemporary research [3]. Students working across these sources must distinguish what can be observed from later interpretation and scientific investigation. When an exhibit presents every statement in the same voice, readers may miss those differences in evidential status.

The teacher's difficulty lies in deciding what a completed product warrants saying about learning. An accurate explanation might reflect careful reasoning, copied wording, or substantial help. Evidence-centered design asks assessors to specify how observable performance supports an intended judgment [4]. Formative assessment asks how that judgment will inform subsequent teaching [5]. For the tasks proposed here, both require access to selected decisions behind the finished exhibit: why a source was used and why an interpretation changed.

Existing research offers guidance on parts of this process. Digital competence frameworks address pedagogical choices, and social annotation research examines exchanges around online texts [6,7]. Work on standards and rubrics considers how learners and teachers judge the quality of a submission [8,9]. Applying these accounts to TCM culture raises more specific questions. Does the cited passage support the cultural interpretation? Does the chosen tool help communicate it, and does revision resolve an identified problem? Participation records and visual quality provide incomplete answers.

This paper develops a task-based assessment proposal with six indicators and an analytic rubric for university courses. Guided work with Longmen materials precedes an individual task using unfamiliar Funiu sources. The conceptual analysis identifies which learning claims these tasks can support, while the instructional design specifies submissions and feedback procedures. Particular attention is given to evidence that group products can obscure and to the difference between independent performance and revision with support.

2. Three-Dimensional Objectives and Assessment Boundaries

2.1. Inheritance: Cultural Understanding and Reasoned Value Reflection

For the inheritance dimension, students explain a resource in its historical setting and give reasons for the value they attribute to it. An interpretation of an inscription should distinguish the inscription itself from later commentary and the student's own inference. Assessment then considers the support for that interpretation and the reasons offered for preserving or communicating the resource. Agreement with a preferred value judgment is not a criterion; different

conclusions are acceptable when the material supports the reasoning.

Such work reveals cultural understanding and value reflection, but says less about a student's lasting attachment or sense of belonging. Cultural identification would require a separate investigation using suitable questionnaires or interviews, including examination of their content and measurement quality. Products could contribute evidence to that investigation alongside students' accounts. The two inheritance indicators proposed here concern classroom performances and should retain that narrower interpretation when results are reported.

2.2. Integration: Digital Resource Use and Its Rationale

Integration concerns the use of digital resources to produce an interpretation that another reader can check. Students retrieve and annotate materials, verify claims, and explain their choice of tools. They must identify both the source and the passage on which an explanation depends. DigCompEdu informs the teacher's instructional choices [6]; it is not adopted here as a student competence measure. A useful digital function makes the explanation easier to follow or its evidence easier to inspect.

These tasks need verifiable texts and images, an annotation tool, and storage for successive versions. A highlighted image region, for example, should be linked to a note explaining what is visible and where the interpretation comes from. Following that note should take the assessor back to the same region. Source cards and individual tool-use explanations complete the record. Common practice materials and records of assistance help teachers distinguish performance during the unit from differences in prior software experience.

2.3. Innovation: Problem Relevance and Supported Improvement

Innovation is assessed through the fit between a proposed explanation and an audience's difficulty. Students identify a likely misunderstanding, compare possible responses, and justify the one they choose. They subsequently consider whether it resolves the difficulty. Within this task, clarifying an ambiguous claim or connecting it more precisely to a source can count as improvement. An unusual visual treatment, on its own, gives the assessor little reason to judge the solution successful.

Six indicators translate the three dimensions into performances that the tasks can elicit. Evidence may be relevant to more than one indicator: an annotation can reveal both contextual understanding and competent tool use. The assessor must explain each judgment separately. Because a shared exhibit can obscure who made a decision, every student submits an interpretation and a tool-use explanation alongside the group product.

3. Learning Objectives, Digital Tasks, and Evidence Judgments

3.1. Establishing Traceable Task-Evidence Relationships

The evidence required determines what students need an opportunity to do. To assess verification, a task must elicit reasons for accepting or withholding a statement. To assess improvement, it must preserve the draft, the feedback, and the response. Table 1 maps these requirements to the project objectives. Source identifiers and annotation locations allow teachers to trace an explanation back to its evidence. Counts of clicks, sources, or edits describe activity; they do not establish the quality of the reasoning.

Table 1. Three-dimensional objectives, indicators, tasks, evidence, and boundaries

Dimension	Indicator	Learning task	Evidence	Boundary
Inheritance	I1 Cultural knowledge and context	Interpret inscriptions in context	Located annotations; individual explanation	History does not establish efficacy
	I2 Reflective value expression	Explain preservation and communication value	Supported value reflection	Not an identification measure
Integration	I3 Retrieval and annotation	Retrieve, locate, and annotate	Source cards; annotated files	Fluency is not understanding
	I4 Verification and tool rationale	Check claims; explain tool choice	Verification log; demonstration	Logs do not replace reasons
Innovation	I5 Problem relevance and design rationale	Compare audience-oriented explanations	Draft; design rationale	Novel appearance is insufficient
	I6 Response and supported improvement	Revise or retain with reasons	Version comparison; response	Edit counts do not show quality

Note: The mapping links course objectives to task evidence. Record individual/group contributions and assistance provided.

The review by Novak and colleagues provides a basis for considering exchanges around annotated digital materials [7]. A shared annotation platform and locally exchanged files differ in their sharing functions, but either can be used for this task. The record needs to preserve the source passage, the student's comment, and any response that led to a changed interpretation. If a statement cannot be traced to a source, the teacher requests clarification before assigning a performance level.

3.2. Interpreting Performance with an Analytic Rubric

Before using the rubric, students compare two teacher-written explanations of the same material. One repeats the source; the other explains its provenance and limits. Discussing the difference makes the criteria more concrete and gives students practice in judging work. This follows Sadler's emphasis on learners understanding quality and evaluating their own performance [8]. The examples are identified as instructional materials, with room for other interpretations when students can support them.

Table 2 describes four levels for each indicator, using explicit criteria and interpretable descriptors informed by rubric research [9]. Students receive it before drafting and first apply it to a worked example. Teachers report a profile of indicator judgments, without weighting them or calculating a total. When the task has not elicited the necessary evidence, the teacher records insufficient evidence and identifies a follow-up opportunity. Evidence that is present but weak is assessed against the descriptors. Each judgment also records any assistance provided.

Table 2. Analytic four-level rubric for six indicators

Indicator	1 Needs clarification	2 Emerging	3 Adequate	4 Elaborated
I1 Cultural understanding	Key descriptions inaccurate	Mostly accurate; weak context	Accurate, contextual explanation	Also qualifies limits and differences
I2 Value reflection	Judgment lacks reasons	Reasons weakly linked to resource	Resource-based value judgment	Compares perspectives; qualifies judgment
I3 Retrieval/annotation	Source/location untraceable	Source located; annotation incomplete	Accurate source-annotation correspondence	Distinguishes source, note, and inference
I4 Verification/tools	No checks or tool rationale	Limited checks; vague rationale	Checks key claims; explains tool	Compares source and tool limitations
I5 Solution design	Audience problem unaddressed	Problem identified; weak rationale	Relevant solution with rationale	Compares alternatives and trade-offs
I6 Feedback and revision	Key issue unanswered	Changes insufficiently explained	Supported revision or retention	Checks whether explanation improved

Note: Report indicators separately; no weights or total. Record insufficient evidence separately. Level 4 requires level 3. For transfer, assess I6 after feedback. Value reflection concerns reasoning, not agreement

3.3. Connecting Feedback with Product and Instructional Changes

Feedback directs attention to a specific problem that the student can examine. For example, an exhibit might use an ancient inscription to support a claim about present-day therapeutic effectiveness. The teacher asks which passage warrants that inference and whether the wording needs qualification. The student keeps the original statement, the revision, and an explanation of the change. Figure 1 locates these records within the teaching sequence and shows the separate transfer task.

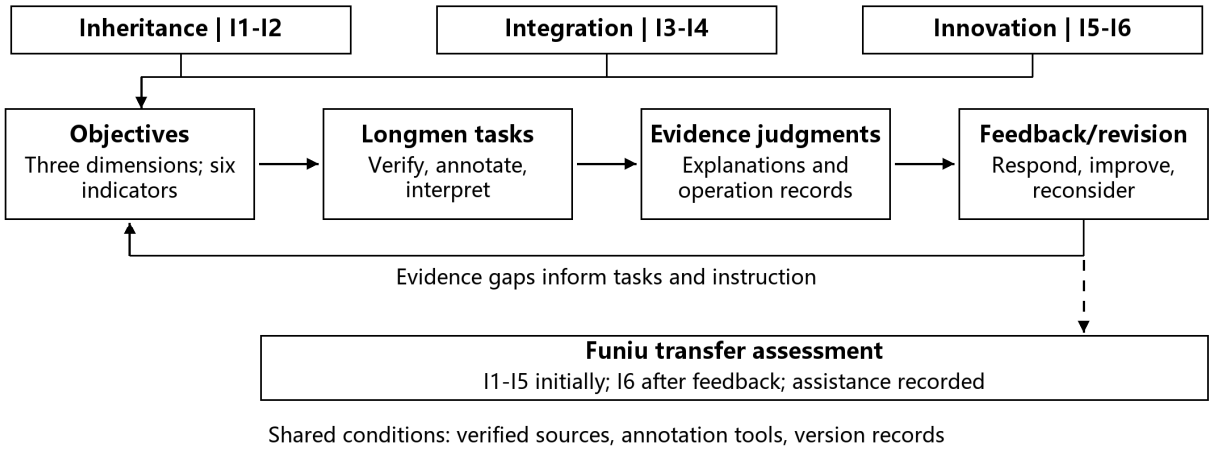


Figure 1. Assessment sequence and the separate Funiu transfer task. The dashed connection identifies the transfer assessment.

A difficulty shared by several groups may also originate in the task. Poor image legibility, for instance, limits what students can establish regardless of their cultural knowledge. The teacher can narrow the question or supply a verifiable explanatory source, recording the adjustment. These instructional changes are documented separately from revisions to student products so that later judgments take account of the conditions under which the work was produced.

4. Digital Instructional Design Using Longmen Resources

4.1. Unit One: Verification and Annotation

The proposed course sequence comprises three 90-minute units in university general education. Students address one question: how can Yaofang Cave inscriptions be explained to non-specialist classmates without overstating what they establish? The teacher supplies a common source pack containing a site introduction with an identified publisher, verifiable explanations, and images cleared for instructional use. If no legible image is available, students use the published explanatory text and state that limitation. No project-built digital archive is required.

Unit one begins with 15 minutes on the question and source types, then allocates 20 minutes to verification and 35 minutes to annotation. Individual explanations and group reports take the remaining 20 minutes. In four-person groups, students divide retrieval, checking, annotation, and interpretation roles; these rotate in the next unit. A source card records the publisher, date, resource location, relevant statement, and limits of use. Annotations distinguish visible features from institutional explanations and student inferences, leaving unreadable characters unresolved.

Each group submits at least three annotations linked to

specific source passages, including one annotation about a limitation. The teacher checks the supporting material and whether the interpretation goes beyond it. Work remains focused on medical culture and history: students neither prepare formulas nor convert dosages or offer treatment advice. Neither the age of a source nor its inscription format establishes contemporary clinical efficacy.

4.2. Unit Two: Audience-Oriented Digital Interpretation

In unit two, groups turn their source work into a three-slide digital exhibit for classmates unfamiliar with the material. The slides introduce the resource, explain its cultural context, and identify questions that still need checking. Each contains approximately 80-120 Chinese characters; source identifiers and necessary image captions are recorded separately. Presentation software or annotated documents are sufficient for the task. Students explain unfamiliar terms and show where observation ends and interpretation begins, since shortening a site description may leave its difficulties unresolved.

Groups spend 10 minutes considering the audience, 15 comparing presentation options, 40 drafting, and 25 explaining their decisions. For one item likely to be misunderstood, they compare at least two explanations and give reasons for their choice. Each student identifies a judgment they contributed and describes what the tool allowed them to do. These accounts let the teacher assess content, digital operations, and design reasoning even when the finished slides differ in visual polish.

4.3. Unit Three: Assessment, Response, and Revision

Unit three allocates 15 minutes to modeling rubric use and review comments, followed by 20 minutes of cross-group review. Revision takes 35 minutes and individual reflection the final 20 minutes. Peer assessment research informs the inclusion of this activity [10]. A review comment must locate a statement or image region, describe the concern, and connect it to a criterion. The teacher checks the comments for accuracy and makes the final assessment; peer agreement does not determine a score.

The final submission contains both versions and a response identifying comments accepted, partly accepted, or declined. The response records how students understood and used feedback, as emphasized by Carless and Boud [11]. Retaining an explanation can be justified, provided the group addresses the concern raised. Teachers examine what substantive revisions achieve rather than how many edits appear. Each student receives a six-indicator profile with a brief explanation of the evidence and one task directed at the next learning need.

4.4. A Funiu Mountain Transfer Task

After the 270-minute sequence, students undertake a separate 45-minute individual task with unfamiliar *Cornus officinalis* materials. The pack contains the Funiu station research account [3] and regional and cultural passages republished on the National Forestry and Grassland Administration website [12]. Students check provenance, write a one-page cultural interpretation, and explain their tool use. Comparing these accounts requires them to decide which source-evaluation strategies carry over and which need adjustment. Legends must remain identified as such rather than being presented as established historical or medical facts.

The first submission is assessed against I1-I5, with attention to regional claims, traditional knowledge, and accounts of modern research. During the 45 minutes, students receive no content feedback; common technical assistance remains available and is recorded. I6 is marked not yet assessed because students have not had an opportunity to respond to feedback. A subsequent feedback-and-revision phase supplies that evidence, with its time and work recorded separately. Keeping both submissions makes it possible to distinguish initial transfer from improvement achieved with support.

5. Implementation Conditions and Validation Pathways

5.1. Resource Review, Scoring Consistency, and Learner Differences

Before teaching, instructors check the provenance, image-use conditions, and interpretive scope of every resource. Where accounts differ, the source pack should preserve the disagreement. Scoring procedures also need examination, as research on rubric reliability and validity makes clear [13]. Two teachers can begin by independently assessing the same instructional examples and then discussing the evidence for their judgments. A pilot should retain both sets of initial ratings as well as any negotiated decisions. Agreement reached through discussion is insufficient evidence of stable reliability.

Shared sources and tool demonstrations give students a common starting point, with classroom practice intended to reduce disadvantages from limited prior experience. A brief oral explanation may supplement writing when fluency obscures the reasoning. Assessing digital competence still requires an observable operation or a record of one. Portfolios contain sources, annotations, drafts, key comments, revisions, and individual explanations. Recording assistance and time spent also allows teachers to judge whether the work fits the proposed schedule.

5.2. Connection with the Project's Technology Pathway

Digital documents and annotations are sufficient for the baseline task. The project's technology pathway also allows possible extensions using virtual reality (VR), augmented reality (AR), or artificial intelligence (AI). Mortara and colleagues discuss the relationship between cultural heritage learning, technology choices, and learning settings [14]. Here, VR could support spatial observation, while AR could link an object to an explanation. AI could supply candidate statements for students to verify. These options would need suitable resources and task-specific assessment before classroom adoption.

With any extension, the record must connect the observation target and tool output to the student's judgment and verification. AI-generated statements need external support because the generated text is not a historical source. Immersion and automated activity logs likewise provide no direct evidence of increased cultural identification. Changes in equipment or software would require teachers to reassess accessibility, task difficulty, and the meaning of ratings. Otherwise, differences in technical provision could be misread as differences in learning.

5.3. Application Conditions and Further Research

The proposal has not undergone classroom trials or psychometric validation. An initial content review should examine whether the indicators adequately represent the intended objectives. Cognitive interviews with students can then test how they understand the instructions and performance descriptors. Classroom trials would establish what evidence can be collected, where raters disagree, and how much work the procedures require. Any investigation of cultural identification needs its own measures and reporting, separate from cultural understanding and reflective expression.

Research on learning effects could compare practiced tasks with independent interpretations of unfamiliar materials. Such comparisons need to account for prior knowledge, tool proficiency, assistance, and task difficulty. Scoring studies should examine ratings made before discussion to identify indicators or evidence types that generate disagreement. Learning performance, scoring consistency, and classroom feasibility require separate analyses because a difficulty in one does not establish a failure in the others.

6. Conclusion

This proposal turns the inheritance-integration-innovation objectives into six assessable indicators for university TCM culture courses. Its main contribution is to specify which records teachers need when interpreting an exhibit: sources, individual explanations, design decisions, and responses to feedback. Guided Longmen tasks and independent work with Funiu materials address different stages of learning, with later revision kept distinct from initial performance. The design offers a procedure for planning assessment and subsequent teaching; classroom feasibility and the quality of the resulting judgments remain to be tested.

Acknowledgments

This work was supported by the Research on the Construction of a Three-Dimensional Practice System for Traditional Chinese Medicine Cultural Education under Grant 2026JYQS311.

References

[1] Zhang, R., & Ruan, T. (2026). Construction of a traditional Chinese medicine cultural education practice system driven by "inheritance-integration-innovation": An exploration of integrating distinctive traditional Chinese medicine resources with modern educational technologies. Academic Journal of Education, 2(3), 5–9.

[2] Museum of Beijing University of Chinese Medicine. (2010). Yaofang Cave at the Longmen Grottoes. Beijing University of Chinese Medicine.

[3] Funiu Mountain Field Scientific Observation and Research Station for Biological Resources and Ecological Environment, Zhengzhou University. (2024). New progress by the Funiu Mountain Station in evaluating Cornus officinalis germplasm resources from major production areas. Zhengzhou University.

[4] Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2003). A brief introduction to evidence-centered design (Research Report RR-03-16). Educational Testing Service.

[5] Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. Educational Assessment, Evaluation and Accountability, 21(1), 5–31.

[6] Redecker, C. (2017). European framework for the digital competence of educators: DigCompEdu (Y. Punie, Ed.; EUR 28775 EN). Publications Office of the European Union.

[7] Novak, E., Razzouk, R., & Johnson, T. E. (2012). The educational use of social annotation tools in higher education: A literature review. The Internet and Higher Education, 15(1), 39–49.

[8] Sadler, D. R. (1989). Formative assessment and the design of instructional systems. Instructional Science, 18(2), 119–144.

[9] Panadero, E., Jonsson, A., Pinedo, L., et al. (2023). Effects of rubrics on academic performance, self-regulated learning, and self-efficacy: A meta-analytic review. Educational Psychology Review, 35(4), 113.

[10] Double, K. S., McGrane, J. A., & Hopfenbeck, T. N. (2020). The impact of peer assessment on academic performance: A meta-analysis of control group studies. Educational Psychology Review, 32(2), 481–509.

[11] Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. Assessment & Evaluation in Higher Education, 43(8), 1315–1325.

[12] Hu, C. Y. (2021). Cornus officinalis: An extraordinary plant and a time-honored medicine. National Forestry and Grassland Administration.

[13] Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. Educational Research Review, 2(2), 130–144.

[14] Mortara, M., Catalano, C. E., Bellotti, F., et al. (2014). Learning cultural heritage by serious games. Journal of Cultural Heritage, 15(3), 318–325.