

## Digital and AI-Enabled Scaffolding in RPL Assessment: A Systematic Literature Review of Mechanisms, Evidence, and Design Principles

Fuad Abdillah<sup>1</sup>, Dhega Febiharsa<sup>2</sup>

<sup>1</sup>Ivet University, Semarang Indonesia

<sup>2</sup>Postgraduate Program, State University of Malang, Indonesia

### ABSTRACT

This study synthesizes mechanisms, evidence, and design principles of digital and AI-enabled scaffolding in Recognition of Prior Learning (RPL) assessment across higher education and vocational/professional certification contexts. Employing a systematic literature review following PRISMA 2020 guidelines, the study analyzed publications from January 2021 to September 2026, distinguishing empirical evidence from conceptual propositions. Four scaffolding mechanisms were identified: experiential learning articulation, portfolio development and evidence mapping, formative feedback, and interpretive support for assessors. The literature documented support for candidate readiness, narrative traceability, and document processing performance. However, it did not provide sufficient evidence that presentation quality or technological accuracy inherently improves the validity and fairness of recognition decisions. Classification accuracy and annotation consistency measure intermediate outputs differing fundamentally from competency evidence adequacy and final decision validity. The study contributes an integrated framework connecting support mechanisms to assessment tasks, evaluation objects, and implementation conditions, producing eight design principles regarding claim ownership, evidence traceability, assistance transparency, adjustment to self-reliance, assessor authority, uncertainty representation, accessibility, and institutional accountability. Because direct empirical evidence on AI-enabled RPL scaffolding remains emerging, several conclusions rely on hypotheses extrapolated from broader AI-in-education literature requiring RPL-specific validation. To guide practitioners, the study proposes an actionable agenda: restrict AI to initial evidence mapping, implement structured candidate templates, establish human-in-the-loop verification, separate presentation support from content generation, and monitor administrative workload shifts. This framework provides a basis for testing candidate and assessor support across diverse RPL contexts.

**Keywords:** Recognition of Prior Learning, digital scaffolding, artificial intelligence, assessment validity, design principles.

### Corresponding author

**Name:** Fuad Abdillah

**Email:** fuadabdillah88@gmail.com

## INTRODUCTION

Recognition of prior learning (RPL) allows institutions to evaluate learning gained through work, community participation, and other experiences against established criteria

for admission, academic credit, or qualifications. This review encompasses RPL in higher education as well as vocational and professional certification, where the credibility of assessment evidence is central to recognition decisions. Across these contexts, a key challenge is translating experience into credible evidence of learning. From the perspective of argument-based validation (Kane, 2013), this translation involves a series of inferences through which assessors connect candidates' claims about their learning to specified competency criteria using documented evidence. Weaknesses at any stage, including how experience is articulated, how supporting artifacts are mapped to criteria, and how narratives are interpreted, can undermine the validity and fairness of the resulting decisions. Comparisons across systems are further complicated by differences in the use of terms such as recognition, validation, assessment, and certification (Villalba-García, 2021). To maintain conceptual clarity, this review uses "RPL assessment" to refer to the evaluation of prior learning against specified criteria, while retaining alternative terms when discussing the contexts in which they originally appear.

Digital support offers ways to help candidates articulate prior learning and prepare assessable evidence. Rambharose (2026) and Brenner et al. (2022) document a hybrid RPL programme that combines reflective activities, portfolio templates, online workshops, and formative guidance. These practices connect experiential knowledge with institutional requirements while supporting academic and digital readiness. AI introduces additional possibilities. Mahfouz et al. (2026) and Krumm et al. (2024) describe a system that analyses learner narratives, maps evidence to competencies, generates explanations, and retains human review. Together, these studies illustrate distinct roles for technology in supporting evidence preparation and interpreting submitted evidence.

This review defines digital scaffolding as structured, technology-mediated support that helps candidates or assessors carry out specific tasks within RPL assessment. AI-enabled scaffolding is a form of this support that uses AI to interpret evidence or provide task-specific guidance. These operational definitions distinguish scaffolding from functions limited to document storage, platform access, or automated scoring. The analysis focuses on how such support helps users identify prior learning, substantiate their claims, align evidence with assessment criteria, and make informed judgments. It also considers how support responds to users' needs while maintaining candidates' responsibility for the evidence they submit and assessors' accountability for their decisions.

The selected literature leaves three connected gaps. First, existing synthesis provides a broad account of RPL without organizing evidence around digital and AI-enabled scaffolding. Raciti et al. (2024) reviewed 65 empirical studies published between 2013 and 2023 and identified themes concerning benefits, challenges, processes, institutions, and policy. This establishes a foundation for understanding RPL but leaves subsequent developments and task-specific technological mechanisms outside its stated synthesis. Second, the evidence addresses different outcomes. Rambharose (2026) examines pedagogical implementation and readiness within a qualitative institutional case study, whereas Mahfouz et al. (2026) and Krumm et al. (2024) report validation of a narrative annotation pipeline and acknowledge the absence of direct end-user usability testing.

Neither outcome alone establishes whether scaffolding improves authentic evidence production or the validity of final recognition decisions. Third, these differing units of analysis make it difficult to connect technological functions, assessment outcomes, and conditions for implementation across settings.

These gaps motivate a systematic review organized around mechanisms, evidence, and design principles. Specifically, this review is guided by the following research questions:

**RQ1:** How does digital and AI-enabled scaffolding operate across RPL assessment tasks?

**RQ2:** What evidence supports its reported effects?

**RQ3:** Which contextual conditions constrain those effects?

**RQ4:** What design principles can be justified by the available evidence?

To address these questions, the synthesis distinguishes between candidate and assessor support, pedagogical and assessment outcomes, and conceptual proposals and evaluated applications. These distinctions are crucial to prevent improvements in readiness or computational agreement from serving as substitutes for evidence of valid and equitable recognition.

To address these research questions, this review proposes an integrated framework that connects scaffolding mechanisms with the assessment tasks they support, their intended outcomes, the available evidence, and the conditions required for implementation. Its contribution lies in making these relationships explicit in RPL assessment, moving beyond a descriptive inventory of technologies or a general discussion of RPL. The framework identifies practices supported by evidence, highlights uncertainties about their applicability across contexts, and indicates priorities for further evaluation. It also provides a basis for formulating design principles that promote evidence traceability, context-sensitive interpretation, accessible support, and accountable human judgment. Throughout, it maintains a clear distinction between documenting prior learning and supporting the development of new learning.

## **METHOD**

This study adopts a systematic literature review (SLR) design, with reporting guided by PRISMA 2020 and PRISMA-S to support transparency and traceability. The review protocol, including the research questions, eligibility criteria, and search strategies, is archived on the Open Science Framework (OSF) to document methodological decisions and help reduce the risk of selective reporting. The review addresses three research questions: how digital and AI-enabled scaffolding mechanisms operate in Recognition of Prior Learning (RPL) assessment; what empirical evidence supports their use and what limits their applicability; and what design principles can be derived from the findings.

Study selection was guided by explicit operational definitions and eligibility criteria. RPL is defined as the recognition of learning acquired through prior experience against established institutional criteria. Digital scaffolding refers to structured, technology-mediated support that helps users articulate learning, map evidence to assessment criteria, or interpret evidence. AI-enabled scaffolding uses artificial intelligence to provide task-specific guidance, feedback, or evidence mapping. Functions limited to document storage

or automated classification, without support for candidates or assessors, fall outside this definition. Eligible publications were peer-reviewed journal articles indexed in Scopus, published between January 1, 2021, and September 30, 2026, and written in English or Indonesian. Articles also had to address RPL assessment directly and examine digital or AI-enabled scaffolding. Exclusion criteria covered studies of AI in general learning contexts without a direct connection to RPL, administrative recognition processes that did not evaluate learning, technologies without scaffolding functions, and conceptual articles lacking clearly articulated design arguments. Gray literature and publication types outside the scope of peer-reviewed journal articles, including editorials and conference abstracts, were also excluded.

Literature identification was carried out through four multidisciplinary databases: Scopus, Web of Science Core Collection, ERIC, and IEEE Xplore, to comprehensively capture educational and technological literature. The search strategy combined RPL-related terms (e.g., "recognition of prior learning" OR "prior learning assessment") with technological terms (e.g., digital OR "artificial intelligence" OR "eportfolio" OR "automated feedback"). In addition to database searches, supplementary searches were conducted through cross-referencing and citation tracking, referring to the TARCIS guidelines to ensure no relevant literature was missed.

Study selection proceeded through four stages, following the PRISMA 2020 reporting framework: identification, screening, eligibility assessment, and inclusion. Records were first identified through databases and other sources, imported, and deduplicated. Two reviewers then independently screened titles and abstracts against the eligibility criteria, resolving disagreements through discussion or consultation with a third reviewer. Full texts of potentially eligible reports were retrieved and assessed, and a primary reason was recorded for each exclusion at this stage, such as insufficient relevance to RPL or the absence of scaffolding functionality. Eligible reports were then linked to their underlying studies to prevent double-counting. The final corpus comprised 12 distinct studies reported across 16 publications, classified as empirical or conceptual.

Data were extracted using a standardized form covering publication details, study context, technology, assessment tasks, support mechanisms, outcome indicators, and limitations. For AI-based systems, extraction also captured technical details, including the models used and procedures for annotator calibration. Two reviewers independently completed methodological quality assessments using JBI appraisal instruments selected according to study design. Appraisal findings were reported by domain, without calculating an overall percentage score. The synthesis combined thematic analysis of scaffolding mechanisms and user experiences with a structured comparison of evaluation findings. Throughout the analysis, candidate support was distinguished from assessor support, pedagogical outcomes from assessment outcomes, and empirical findings from conceptual propositions. Design principles were subsequently developed using a matrix linking assessment tasks, scaffolding mechanisms, outcomes, and implementation conditions, with the evidential support for each principle explicitly indicated.

Figure 1 presents the study selection process using the PRISMA 2020 flow diagram, from initial identification and title and abstract screening to full-text assessment and final inclusion. It also reports the reasons for full-text exclusions. Records identified through databases and those located through other sources, including reference list screening and citation searching, are shown as separate pathways before deduplication. Counts at each stage reflect the actual search and selection results and were cross-checked against the selection flow table and the article list provided in the supplementary material. The diagram was prepared using an official PRISMA 2020 template to support transparent reporting.

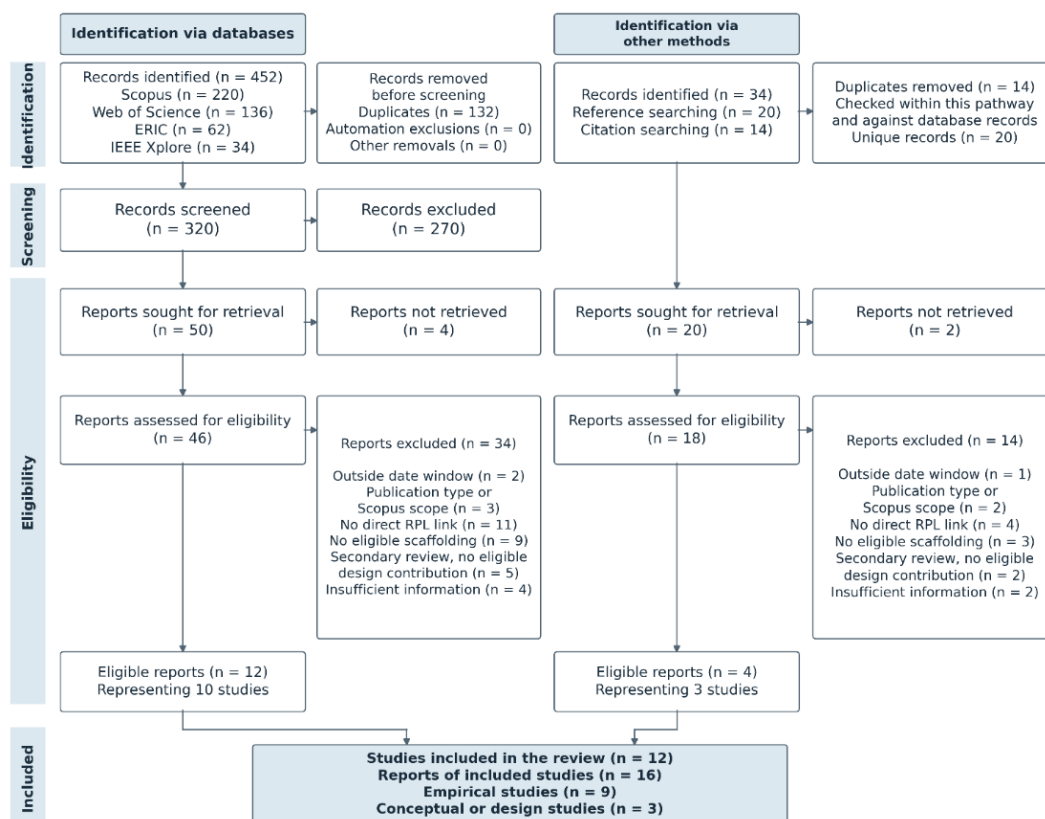


Figure 1. Prisma 2020 flow diagram

## FINDING AND DISCUSSION

### RESEARCH RESULT

The selected literature contains four support mechanisms: articulation of learning from experience, portfolio development and evidence mapping, formative feedback, and interpretation of evidence for assessors. These mechanisms operate across different assessment tasks, although some are applied within the same program or system. Crucially, the reviewed studies vary in their evidentiary basis: some report empirical findings from implemented programs or evaluated systems, while others propose conceptual frameworks

or design models without direct empirical testing. This distinction is maintained throughout the synthesis to avoid conflating demonstrated effects with theoretical propositions.

### **Documented Implementation and Evaluation**

(Rambharose, 2026a) reports an empirical case study of a hybrid RPL programme in South Africa that integrates reflective activities, online workshops, asynchronous assignments, and mentoring. This study documents how experiential articulation support is embedded within portfolio development, with templates and facilitator guidance helping candidates tailor their experiences to admissions requirements. The findings indicate that the programme supports academic literacy, digital skills, and self-confidence, though the case study design does not isolate the specific contribution of digital scaffolding from human facilitation and other program activities. Data sources included program documentation, design artifacts, facilitator reflections, and candidate portfolios.

Mahfouz et al. (2026) and Krumm et al. (2024) present empirical validation of the LivedX system, which analyses learner narratives, maps evidence to competency indicators, generates explanations, and retains human review. Their reported outcomes include annotation consistency after training and calibration, as well as traceable relationships between narratives and competency indicators. However, these studies explicitly acknowledge the absence of direct end-user usability testing, meaning that benefits to candidate abilities or assessor decision-making cannot yet be established from their evidence. The evaluation focuses on system performance under research conditions rather than authentic RPL recognition decisions.

Diyasa et al. (2025) provide empirical evidence on document classification, evaluating 525 non-formal learning certificates across seven computer science course categories using OCR, BERT, and text augmentation. Their results show validation accuracy above 88%, demonstrating the technical feasibility of automated certificate categorization. However, this measure indicates suitability of classification to document categories, not the percentage of correct RPL credit decisions or the certificate holder's level of competency mastery. The study evidences document processing performance but does not establish that candidates are better able to choose evidence or that assessors produce more informed decisions.

### **Proposed Frameworks and Design Models**

Rambharose (2026b) proposes a multidisciplinary theoretical framework for RPL research that integrates qualitative, quantitative, and mixed-methods inquiry approaches. This conceptual work theorizes how recognition processes can be studied but does not report empirical evaluation of scaffolding mechanisms or their effects on assessment outcomes.

Parenzee & Rambharose (2026) present conceptual analysis of human-centred approaches to RPL policy implementation in higher education, identifying tensions between policy objectives and institutional practices that limit professional autonomy and

organizational support. Their work highlights institutional conditions and governance design but does not measure the influence of specific digital features or AI functions.

Brenner et al. (2022) document a hybrid RPL programme that combines reflective activities, portfolio templates, online workshops, and formative guidance, connecting experiential knowledge with institutional requirements while supporting academic and digital readiness. While this study reports on program implementation, its primary contribution is pedagogical design rather than empirical evaluation of scaffolding effectiveness.

### Synthesis of Mechanisms Across Study Types

Table 1 consolidates the four support mechanisms identified across both empirical and conceptual studies, distinguishing between documented features and proposed functions.

**Table 1. Support mechanisms in selected literature**

| <b>Mechanism</b>                                | <b>Users and tasks</b>                                                                       | <b>Reported forms of support</b>                                                                 | <b>Sources and limits of evidence</b>                                                                                                                     |
|-------------------------------------------------|----------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Articulation of learning</b>                 | The candidate explains the lessons learned from the experience.                              | Reflective questions, online workshops, asynchronous tasks, and mentoring.                       | Rambharose (2026a) [empirical]. The program case study does not separate the contribution of technology from facilitation.                                |
| <b>Portfolio development and mapping</b>        | The candidate compiles evidence; the assessor links the evidence to the criteria.            | Portfolio template and the relationship between the narrative section and competency indicators. | Rambharose (2026a) [empirical]; Mahfouz et al. (2026)[empirical]. Admission requirements and competency taxonomies are distinct targets for mapping.      |
| <b>Formative feedback</b>                       | The candidate reviews and revises the evidence.                                              | Facilitator guide, peer interaction, system explanation, and follow-up suggestions               | Rambharose (2026a) [empirical]; Mahfouz et al. (2026) [empirical]. The use of feedback and the fading of assistance have not yet been directly evaluated. |
| <b>Interpretation of evidence for assessors</b> | The examiner evaluates the proposed relationship between the narrative and the competencies. | Text annotation, confidence scores, explanations, and human corrections.                         | Mahfouz et al. (2026) [empirical]. System evaluation has not yet established benefits regarding the recognition decision.                                 |

## Evaluation Outcomes and Their Interpretive Boundaries

Table 2 distinguishes between reported evidence and the limits of interpretation, clarifying that technical performance metrics (such as classification accuracy) differ fundamentally from assessment validity measures.

**Table 2. Evaluation results and scope of interpretation**

| Dimensions                        | Reported evidence                                                                                               | Source                                                             | Interpretation limits                                                                    |
|-----------------------------------|-----------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------|------------------------------------------------------------------------------------------|
| <b>Candidate readiness</b>        | Academic literacy, digital skills, and confidence in a hybrid RPL program                                       | Rambharose (2026a) [empirical]                                     | The influence of technology is inseparable from mentoring and academic activities        |
| <b>Representation of evidence</b> | The relationship between narrative parts and competency indicators                                              | Mahfouz et al. (2026) [empirical]; Krumm et al. (2024) [empirical] | Traceability does not establish the authenticity or adequacy of evidence                 |
| <b>Annotation consistency</b>     | Assessor agreement after training and calibration                                                               | Mahfouz et al. (2026) [empirical]; Krumm et al. (2024) [empirical] | Annotation consistency is different from the validity of the recognition decision        |
| <b>Document classification</b>    | Validation accuracy above 88% on certificate classification                                                     | Diyasa et al. (2025) [empirical]                                   | Results limited to research categories and data; does not establish RPL decision quality |
| <b>Access and representation</b>  | Flexibility of online activities, device and connectivity barriers, and limitations of narrative representation | Rambharose (2026a) [empirical]; Mahfouz et al. (2026) [empirical]  | Platform access and model audits don't cover the entire candidate journey                |
| <b>Institutional conditions</b>   | Practitioner autonomy, accountability, empathy, and sustainability of implementation                            | Parenzee & Rambharose (2026) [conceptual]                          | Institutional findings do not measure the influence of digital features or AI            |

The quality of evidence and the validity of decisions are distinct objects of evaluation. Completeness indicates the availability of requested documents, alignment indicates the relationship of evidence to criteria, while sufficiency indicates the overall support of evidence for the claim. The selected literature evaluates readiness, annotation,

and classification more than the relationship between these three characteristics and the recognition decision.

Accessibility is discussed through the flexibility of participation and barriers to technology use. Rambharose (2026a) documents device constraints, connectivity, and digital readiness. Mahfouz et al. (2026) discuss experience representation and algorithmic fairness audits. Both sources examine different parts of the process; results on participants who have submitted evidence do not explain the barriers faced by candidates who discontinue prior to submission.

Implementation conditions include facilitation capacity, assessor authority, organizational support, and system maintenance. Parenzee & Rambharose (2026) and Brenner et al. (2022) discuss the tension between RPL policy objectives and institutional practice conditions. In AI systems, taxonomic maintenance and human checks are part of the interpretation flow. The selected literature has not established the success of transferring support across changes in language, discipline, population, and recognition objectives. Proposed system expansion is therefore noted separately from implementation that has been evaluated.

#### Design Principles and Their Evidentiary Support

Design principles are grouped based on the relationship between support, evidence, and user responsibility functions. Table 3 differentiates their support status between documented features (from empirical studies), contextual findings (from case studies), and synthesis propositions (derived from conceptual analysis or cross-study synthesis).

**Table 3. Design principles and support status in the literature**

| Principle                                    | Operational representation                                                                         | Source policy                                                      | Support status                                                                           |
|----------------------------------------------|----------------------------------------------------------------------------------------------------|--------------------------------------------------------------------|------------------------------------------------------------------------------------------|
| <b>Ownership of the candidate's claim</b>    | Confirmation of experiences, personal contributions, and substantive changes to the narrative      | Rambharose (2026a) [empirical]; Mahfouz et al. (2026) [empirical]  | Derived from the practice of mentoring and system limitations                            |
| <b>Traceability of evidence and criteria</b> | The relationship between artifacts or pieces of text, competency indicators, and mapping reasons   | Mahfouz et al. (2026) [empirical]; Krumm et al. (2024) [empirical] | Documented in the system design; adequacy for RPL decisions has not yet been established |
| <b>Transparency of aid contributions</b>     | Differentiation of language correction, organization, mapping suggestions, and addition of content | Synthesis of mechanisms and limits of proof                        | Design proposition                                                                       |

|                                             |                                                                                                         |                                                                              |                                                                                               |
|---------------------------------------------|---------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------|
| <b>Adjustment to self-reliance</b>          | Changes in the intensity of assistance based on the candidate's ability to explain and examine evidence | Rambharose (2026a) [empirical]                                               | Guided support reported; aid reductions have not been evaluated                               |
| <b>Assessor's authority</b>                 | Acceptance, change, or rejection of recommendations for reasons listed                                  | Mahfouz et al. (2026) [empirical]                                            | Human examination is documented; benefits to the final decision have not yet been established |
| <b>Representation of uncertainty</b>        | The status has not been mapped, needs clarification, or there is not enough evidence                    | Mahfouz et al. (2026) [empirical]; Synthesis of Interpretation Limits        | Operational proposition                                                                       |
| <b>Access and context of the experience</b> | Variety of presentation media and support according to candidate barriers                               | Rambharose (2026a) [empirical]                                               | Supported contextual findings; the impact of alternative formats has not yet been established |
| <b>Institutional accountability</b>         | Person in charge of corrections, re-examinations, criteria changes, and system maintenance              | Parenzee & Rambharose (2026) [conceptual]; Mahfouz et al. (2026) [empirical] | Supported by institutional findings and governance design                                     |

Authenticity in the synthesis includes the origin of the artifact, the truth of the experience, and the attribution of the candidate's contribution. Presentation aids can change the narrative structure without altering the experience on which the claim is based. The selected literature has not provided thorough testing linking the type of aid, the candidate's ability to explain the evidence, and the validity of the final decision. The principles in Table 3 are recorded according to their support status and are not treated as a single design device whose components have been validated.

## DISCUSSION

Findings Findings regarding experiential articulation, portfolio development, feedback, and assessor interpretation position digital and AI-assisted scaffolding at the critical intersection of candidate experience and institutional evidentiary requirements. This

aligns with the core problem raised in the introduction: translating situated experience into credible evidence through a valid inferential process. However, because the selected RPL-specific literature evaluates disparate parts of this process, success at one stage cannot be assumed as the overall success of the assessment. Furthermore, because direct empirical evidence on AI-enabled RPL scaffolding is still emerging, this discussion explicitly differentiates between demonstrated mechanisms from RPL-specific empirical studies and hypotheses extrapolated from the broader AI-in-education literature. Therefore, this discussion utilizes the argument-based validation framework established earlier (Kane, 2013) to contextualize each finding, structured explicitly around the four research questions guiding this review.

### **Operational Mechanisms of Scaffolding (Addressing RQ1)**

The findings reveal that scaffolding operates not merely through the existence of a technological platform, but through the interaction between the candidate, the device, and the institutional environment. Regarding experiential articulation, Rambharose (2026a) empirically demonstrates that portfolio development programs integrate reflection, digital activities, and facilitation. In this context, templates provide structural presentation, while facilitator dialogue clarifies the relationship between experience and requirements. Consequently, the analytical implication is that scaffolding mechanisms must be evaluated through assisted actions and conditions of use.

This operational mechanism also involves complex knowledge representation. Integrating adult learning and decolonial perspectives, Rambharose (2026b) and Brenner et al. (2022) conceptually highlight the diversity of knowledge bases in RPL. The difficulty candidates face in writing academic narratives should not be conflated with an absence of learning. While digital support aids presentation, the formats and categories used dictate which experiences are easily recognizable. A mismatch between a candidate's narrative and the system's taxonomy remains an operational friction point, indicating that the benefits of alternative formats (text, audio, demonstration) are strictly dependent on their suitability for the assessed competencies.

For assessors, scaffolding operates through the explanatory function of AI, linking system recommendations to human examination. Adapting Khosravi et al. (2022) work on explainable AI to the RPL context, the mechanism requires that explanations provide relevant information to examine the evidence-criteria relationship. A mere competency category statement is insufficient; the system must indicate the location of evidence, specific indicators, and missing information. Thus, the operational mechanism for assessors relies on the readability of explanations, their alignment with evidence, and their utility in identifying errors.

### **Evidence of Effects and Their Limitations (Addressing RQ2)**

The review indicates that current RPL-specific evidence often measures intermediate outputs rather than final assessment validity. Regarding portfolio development, findings clarify the distinction between evidence completeness (document

availability), alignment (relationship to criteria), and adequacy (supporting claims at the required level). Viewed through the argument-based validation framework (Kane, 2013), documented mapping provides a basis for examination but does not inherently justify the entire interpretive argument regarding inferred competencies and recognition decisions.

Similarly, in the context of automated classification, Diyasa et al. (2025) provide empirical RPL evidence of validation accuracy above 88% in grouping certificates. While this evidences the success of document organization, it does not measure the certificate holder's ability to apply competencies. Within the proposed framework, classification output only identifies evidence relevance; moving toward recognition decisions still requires substantive reasoning regarding learning mastery and individual contribution.

Evidence regarding feedback further illustrates the boundary between RPL-specific findings and broader educational theories. Drawing on Carless & Boud (2018) feedback literacy framework, the ability to understand and act on information is central to the feedback process. In RPL, this means candidates must recognize unsupported claims and justify revisions. Therefore, the frequency of AI comments is insufficient evidence of formative success; the true mechanism lies in the relationship between comment content, candidate understanding, revision actions, and subsequent evidence adequacy.

Crucially, where RPL-specific empirical evidence is absent, we must extrapolate hypotheses from the broader AI-in-education literature. For instance, evidence of portfolio improvement must be distinguished from candidate independence. Darvishi et al. (2024) and Kobaba et al. (2025) demonstrate in general educational contexts that work quality with AI assistance differs from ability without it. *Extrapolating this to RPL forms the hypothesis that* an improved portfolio does not prove the candidate can independently account for the content. Independence here refers to the ability to explain experiences, evidence selection, and claim limits. Similarly, regarding authenticity, Mahfouz et al. (2026) and Krumm et al. (2024) warn of the risk of candidates presenting strategic narratives to fit competency categories. Thus, ownership of the claim must be tied to the veracity of experience and attribution of contribution, not merely to who compiled the final text.

Finally, evidence regarding human-AI collaboration presents mixed outcomes. While Mahfouz et al. (2026) document examiner involvement in validating system interpretations, Vaccaro et al. (2022) meta-analysis in general contexts indicates that human-AI combinations do not generally outperform their best individual components. *Applying this to RPL suggests the hypothesis that* human involvement alone does not guarantee quality; the assessor's contribution must be demonstrated through recognized errors, considered evidence, and documented reasons for altering recommendations.

### **Contextual Conditions and Constraints (Addressing RQ3)**

The effectiveness of scaffolding is heavily constrained by both individual accessibility barriers and broader institutional conditions. At the candidate level, Baker & Hawn (2022) highlight sources of bias in educational algorithms, demonstrating that model audits fail to capture all RPL fairness issues. Candidates often face device, language, or instructional comprehension barriers before their evidence even enters the system.

Consequently, analyses relying solely on completed submission data obscure the experiences of marginalized groups. Evaluation must therefore distinguish between the opportunity to participate, the opportunity to represent learning, and the opportunity to obtain decision scrutiny.

At the institutional level, Parenzee & Rambharose (2026) identify conceptual tensions between RPL policy objectives and practices that limit professional autonomy and organizational support. Technology recommendations require the institutional authority and resources to be acted upon. For instance, flagging evidence shortages yields no improvement if assessors lack the time or procedural pathways to request clarification. This reveals a critical constraint: support design must encompass the division of responsibilities, correction pathways, and re-examination protocols. The effectiveness of these procedures is not guaranteed by technical features but is contingent upon institutional implementation and the preservation of assessor autonomy.

### **Justified Design Principles and Contributions (Addressing RQ4)**

Bringing together the scaffolding mechanisms, limitations of the evidence, and contextual constraints provides a basis for formulating design principles grounded in the review findings. Technology should support clearly defined assessment tasks, with its use guided by the strength and scope of the available evidence. For candidates, scaffolding should help them articulate their experience and explain why selected artifacts substantiate their learning claims. For assessors, it should facilitate evidence traceability, scrutiny of how evidence is mapped to criteria, and documentation of the reasoning behind recognition decisions. Institutions can use these principles to guide development and evaluation while maintaining a clear distinction between documenting prior learning and developing new learning. Any time savings associated with AI should also be considered alongside the work required for clarification, correction, and mentoring, so that efficiency gains reflect a reduction in overall workload rather than a redistribution of effort across the assessment process.

Conceptually, this study contributes a framework that connects support mechanisms directly to the evidentiary relationships they mediate, complementing broader RPL scopes (e.g., Raciti et al., 2024) by focusing on specific candidate/assessor tasks and outcomes. Methodologically, it enforces the separation of candidate ability, artifact quality, system performance, and decision validity. Design-wise, it ensures the traceability of each principle to its evidentiary source and support status.

### **Limitations and Future Research Directions**

The scope of this discussion is constrained by variation in what the included studies evaluated, limited end-user testing within the primary corpus, and uncertainty about the transferability of findings across contexts. Direct empirical evidence on AI-enabled scaffolding in RPL assessment remains limited. Consequently, some conclusions draw on hypotheses informed by the broader AI-in-education literature and require validation in RPL settings. The absence of evidence for a particular outcome should therefore be understood

as a gap requiring targeted evaluation, rather than as proof that the support offers no benefit.

Future research should examine how scaffolding influences changes to assessment evidence and how these changes affect recognition decisions. To guide research and institutional implementation, this review proposes five priorities:

1. Identify the contribution of technology. Comparative studies should examine conventional mentoring, digital scaffolding without AI, and AI-enabled scaffolding. Holding assessment tasks, criteria, and the intensity of human support constant would help distinguish the contribution of each approach.
2. Assess candidate independence. Evaluations of portfolio quality should be complemented by assessments of candidates' ability to explain and substantiate their evidence after scaffolding has been withdrawn, without AI assistance. These assessments should distinguish prior learning from newly acquired presentation skills and competencies developed through mentoring.
3. Examine assessor decision-making. Studies should compare assessment decisions made with and without AI support, using outputs generated independently by AI as an additional comparator. Relevant indicators include the types of errors made, the identification of relevant evidence, the quality of assessment reasoning, and the time required to review submissions.
4. Establish proportionate administrative safeguards. Institutions should evaluate human oversight protocols that limit AI to preliminary evidence mapping and the identification of inconsistencies, while reserving final credit decisions and judgments about narrative authenticity for human assessors. Implementation studies should also examine whether these safeguards can be maintained without imposing excessive administrative demands.
5. Test robustness across contexts. Research across institutions, languages, and disciplines should assess whether scaffolding functions remain effective under different conditions. Accessibility evaluations should include candidates who discontinue the process, so that successful technical implementation is assessed separately from candidates' ability to participate and complete the assessment.

Figure 2 brings these priorities together in an integrative framework linking digital and AI-enabled scaffolding mechanisms to the flow of evidence through RPL assessment, the aspects being evaluated, and the resulting design principles. The framework underscores that improvements in candidate capability, evidence quality, or system performance do not, on their own, establish the validity or fairness of recognition decisions.

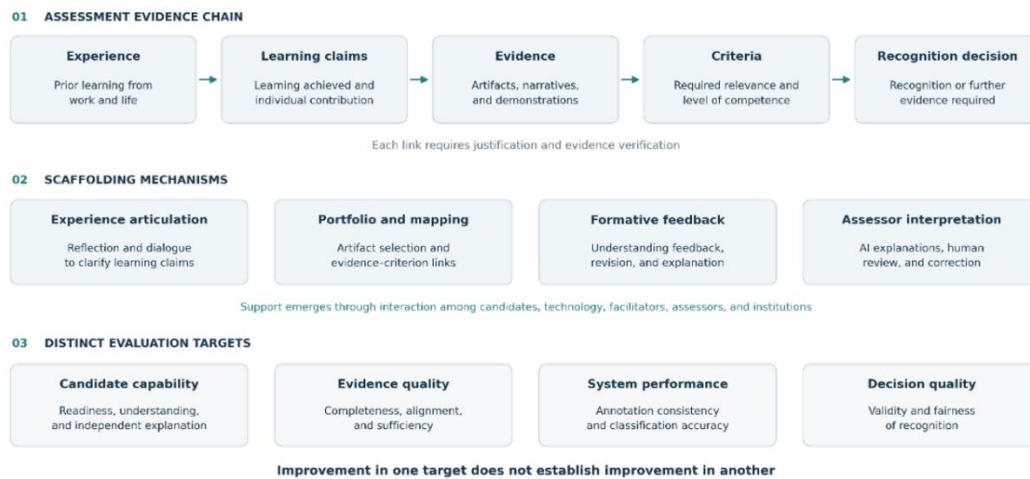


**Figure 2. An Integrated Framework of Digital and AI-Enabled Scaffolding in RPL Assessment**

## CONCLUSION

This study identifies four digital and AI-assisted scaffolding mechanisms in Recognition of Prior Learning (RPL) assessment across higher education and vocational/professional certification contexts: experiential learning articulation, portfolio development and evidence mapping, formative feedback, and interpretive support for assessors. The reviewed literature documents support for candidate readiness, narrative traceability, and document processing performance. However, these results do not provide sufficient evidence that presentation quality or technological accuracy inherently improves the validity and fairness of recognition decisions. Classification accuracy and annotation consistency measure intermediate outputs that differ fundamentally from competency evidence adequacy and final decision validity.

The study contributes an integrated framework connecting support mechanisms to assessment tasks, evaluation objects, and implementation conditions. This framework enforces separation of candidate ability, artifact quality, system performance, and decision validity, producing eight design principles regarding claim ownership, evidence traceability, transparency of assistance, adjustment to self-reliance, assessor authority, uncertainty representation, accessibility, and institutional accountability.

This discussion is bounded by heterogeneity of evaluation objects, limitations of end-user testing, and constraints of cross-context transfer. Because direct empirical evidence on AI-enabled RPL scaffolding remains emerging, several conclusions rely on hypotheses extrapolated from broader AI-in-education literature requiring RPL-specific validation.

To provide actionable guidance for RPL administrators implementing AI-assisted scaffolding without increasing administrative burden, we propose: (1) restrict AI to initial evidence mapping and flagging inconsistencies, keeping final credit decisions with human assessors; (2) implement structured candidate templates separating experience description

from competency claims; (3) establish human-in-the-loop verification with documented reasons for accepting or modifying AI recommendations; (4) separate presentation support from content generation to preserve claim ownership; and (5) monitor workload shifts to ensure burdens are not merely relocated across assessment stages.

Future research must examine relationships between support, evidence modification, and recognition decisions through comparative studies, post-scaffolding independence assessments, and cross-context testing across institutions, languages, and disciplines.

## REFERENCES

- Baker, R. S., & Hawn, A. (2022). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education*, 32(4), 1052–1092. <https://doi.org/10.1007/s40593-021-00285-9>
- Brenner, A., Goodman, S., Meadows, A., & Cooper, L. (2022). From prior learning assessment to specialised pedagogy: facilitating student transition through RPL assessment and selection. *Studies in Continuing Education*, 44(3), 425–440. <https://doi.org/10.1080/0158037X.2021.1874333>
- Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315–1325. <https://doi.org/10.1080/02602938.2018.1463354>
- Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D., & Siemens, G. (2024). Impact of AI assistance on student agency. *Computers and Education*, 210(July 2023). <https://doi.org/10.1016/j.compedu.2023.104967>
- Diyasa, I. G. S. M., Puspaningrum, E. Y., Saputra, D., & Wan Awang, W. S. (2025). Enhancing non-formal learning certificate classification with text augmentation: A comparison of character, token, and semantic approaches. *Interdisciplinary Journal of Information, Knowledge, and Management*, 20, 18. <https://doi.org/10.28945/5525>
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1–73. <https://doi.org/10.1111/jedm.12000>
- Khosravi, H., Buckingham Shum, S., Chen, G., Conati, C., Tsai, Y. S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 3, 100074. <https://doi.org/10.1016/j.caeai.2022.100074>
- Kobaba, B. M., Lihega, J. K., & Khalfan, K. A. (2025). AI in Higher Education: Effect on Student Competency and Independent Learning. *Asian Journal of Education and Social Studies*, 51(7), 1100–1106. <https://doi.org/10.9734/ajess/2025/v51i72194>
- Krumm, A. E., Chahine, S., Schuh, A. M., Schumacher, D. J., Zabar, S., George, B. C., Marcotte, K., Sebok-Syer, S. S., Barone, M. A., & Smirnova, A. (2024). Digital Evidence: Revisiting Assumptions at the Intersection of Technology and Assessment. *Perspectives on Medical Education*, 13(1), 553–560. <https://doi.org/10.5334/pme.1270>

- Mahfouz, J., Verma, G., Douglass, H., Davis, S. J., Bassaj, R., Galvan, A., Jhong, E., & Alam, M. A. (2026). From experience to evidence: A human–AI conceptual model for assessing and credentialing lived learning. *Frontiers in Education*, 11, 1802418. <https://doi.org/10.3389/feduc.2026.1802418>
- Parenzee, R., & Rambharose, R. (2026). Human-centred approaches to recognition of prior learning: Policy implementation in higher education. *International Journal of Lifelong Education. Advance Online Publication*. <https://doi.org/10.1080/02601370.2026.2702060>
- Raciti, M., Tham, A., & Dale, J. (2024). Recognition of prior learning in higher education: A systematic literature review. *Journal of University Teaching and Learning Practice*, 21(9). <https://doi.org/10.53761/bys3aj56>
- Rambharose, R. (2026a). Facilitating access to higher education and promoting academic readiness through digital pedagogical implementation for recognition of prior learning: A South African case study. *Journal of Adult and Continuing Education*. <https://doi.org/10.1177/14779714261429421>
- Rambharose, R. (2026b). Theorising Recognition of Prior Learning Research in Higher Education: A Multidisciplinary Framework for Qualitative, Quantitative, and Mixed-Methods Inquiry. *International Journal of Qualitative Methods*, 25, 1–14. <https://doi.org/10.1177/16094069251404325>
- Vaccaro, M., Almaatouq, A., & Malone, T. (2022). When combinations of humans and AI are useful: A systematic review and meta-analysis. *Nature Human Behaviour*, 8, 2293–2303. <https://doi.org/10.1038/s41562-024-02024-1>
- Villalba-García, E. (2021). Validation of non-formal and informal learning: The hero with a thousand faces? *European Journal of Education*, 56(3), 351–364. <https://doi.org/10.1111/ejed.12468>