

Strengthening Competency-Based Assessment in Vocational Training Institutions in the Oshana Region, Namibia

Kadonsi Kaziya^{1*}, Moses Lungameni²

¹George Benson Christian University College, Zambia

²Africa Research University, Zambia

*Corresponding Author

DOI: <https://doi.org/10.47772/IJRISS.2026.100700737>

Received: 25 July 2026; Accepted: 30 July 2026; Published: 11 August 2026

ABSTRACT

Competency-based vocational assessment requires practitioners to interpret occupational standards, design authentic tasks, gather sufficient evidence, exercise defensible judgement, provide useful feedback and maintain records that can withstand quality review. These demands are often treated as the responsibility of individually trained assessors, with insufficient attention to the institutional conditions that make sound practice possible. This article examines assessment literacy as a distributed institutional capacity. It reports a focused analysis of a qualitative multiple-case study conducted in three vocational training institutions in the Oshana Region of Namibia. Twenty purposively selected participants 14 educators and six quality-assurance officers—took part in semi-structured interviews, complemented by document analysis of assessment policies, schedules, tools, marking guides, moderation records, marksheets and assessment files. Data were analysed through reflexive thematic analysis using within-case and cross-case interpretation. Participants demonstrated strong reported understanding of assessment as evidence of practical and theoretical competence, a basis for learner feedback and a safeguard of qualification credibility. However, interview and documentary evidence also revealed uneven alignment between tasks and competencies, variable marking criteria, incomplete records, constrained practical authenticity and inconsistent uptake of quality-assurance feedback. These patterns were shaped by policy clarity, leadership follow-up, continuing professional development, workload, resources and collaborative culture. The article concludes that credible competency-based assessment depends on the interaction of individual knowledge and judgement with institutional support and professional learning. It proposes an Institutional Assessment Literacy Capacity Framework for sustained, practice-embedded development.

Keywords: *assessment literacy; competency-based assessment; vocational education and training; institutional capacity; professional learning; assessor judgement; Namibia*

INTRODUCTION

Competency-based education and training places unusually demanding responsibilities on those who assess learning. The assessor is not simply required to allocate marks to a written response. The assessor must interpret occupational and unit standards, decide what evidence would adequately represent competence, design or select tasks that elicit that evidence, observe both process and product, judge performance against transparent criteria, accommodate legitimate learner needs without weakening standards, provide useful feedback, maintain defensible records and participate in quality-assurance processes. In vocational education, these demands are intensified because competence frequently combines conceptual knowledge, practical performance, safety, judgement, work habits and the capacity to act in conditions resembling occupational practice. Weakness at any point in this chain can affect the credibility of the resulting competence decision.

Namibia's vocational education and training framework places assessment and certification within a standards-based system that requires credible evidence, moderation and traceable decision-making (Republic of Namibia, 2008, 2023). These expectations make assessment literacy central to qualification credibility rather

than a peripheral teaching skill. Formal assessor preparation remains important, but course completion alone cannot establish how consistently practitioners interpret competencies, design tasks, judge evidence or learn from quality-assurance feedback in their institutional settings.

Assessment literacy was historically associated with technical knowledge of measurement, testing and score interpretation. Contemporary scholarship has moved beyond that narrow conception. Popham (2009), Brookhart (2011), Willis et al. (2013), Xu and Brown (2016), Pastore and Andrade (2019) and Pastore (2023) describe assessment literacy as multidimensional, situated and enacted. It includes knowledge and skill, but also professional judgement, beliefs, identity, ethical responsibility, contextual responsiveness and the capacity to use assessment in support of learning. This shift is important because an assessor may understand principles in the abstract yet struggle to implement them where policies are unclear, equipment is unavailable, workloads are high or professional dialogue is weak. Conversely, a well-designed institutional system cannot compensate fully for an assessor who lacks the knowledge needed to recognise valid evidence or construct appropriate criteria.

Despite this conceptual development, much assessment-literacy research remains concentrated in school teaching, language education or initial teacher preparation. Studies commonly rely on self-report questionnaires, confidence ratings or knowledge tests, each of which captures only part of the construct. Self-reported confidence does not demonstrate competence; a test of declarative knowledge does not show whether that knowledge is enacted; and a review of artefacts does not necessarily explain the institutional conditions under which those artefacts were produced. The gap is particularly visible in vocational education, where assessment is inseparable from occupational standards, practical authenticity, equipment, evidence sufficiency and collective quality assurance.

African TVET studies have documented difficulties involving staff preparation, resources, policy enactment and internal continuous assessment, but assessment literacy has seldom been analysed as an interaction between individual capability and institutional capacity (Maja & Atukunda, 2022; Mutebi & Ferej, 2023). This article addresses that gap through a focused analysis of evidence from three vocational training institutions in the Oshana Region of Namibia. It asks not only what educators and quality-assurance officers say they know about assessment, but also how their accounts correspond with assessment artefacts and how institutional policies, leadership, resources, workload and collaborative culture enable or constrain practice.

The article is guided by four questions: (1) What dimensions of assessment literacy are reported or evidenced among educators and quality-assurance officers in the selected institutions? (2) Where do gaps appear in the planning, design, enactment, judgement, feedback and documentation of competency-based assessment? (3) How do policy, leadership, professional learning, workload, resources, quality-assurance support and institutional culture shape the enactment of assessment literacy? and (4) What institutional capacity-development framework is warranted by the findings? The article contributes a multidimensional account of vocational assessment literacy, an evidence-sensitive distinction among reported understanding, documentary enactment and unmeasured practice, and a context-responsive framework for sustained capacity development.

ASSESSMENT LITERACY IN COMPETENCY-BASED VOCATIONAL EDUCATION

2.1 From technical knowledge to situated professional capability

Early formulations of assessment literacy emphasised the knowledge and skills required to select, construct, administer, score and interpret assessments. These foundations remain necessary. Assessors must understand purposes, validity, reliability, bias, criteria, evidence and the consequences of decisions. Popham (2009) treats assessment literacy as fundamental because educators routinely make decisions that affect learners, while Brookhart (2011) argues that assessment knowledge must connect technical quality with instructional use, communication and ethical practice.

The limitation of a checklist conception is that assessment is not enacted in a vacuum. Willis et al. (2013) show that assessment literacies are shaped by disciplinary knowledge and the rules, relationships and discourses of the setting. Xu and Brown's (2016) Teacher Assessment Literacy in Practice framework places a knowledge base within layers of conceptions, institutional and sociocultural context, identity and compromise. Pastore and

Andrade (2019) similarly propose a model comprising knowledge, skills and dispositions in context, and Pastore's (2023) systematic review highlights the persistent difficulty of transferring assessment theory into practice. Assessment literacy is therefore better understood as professional capability in action than as possession of a fixed body of technical facts.

2.2 Interpreting standards and designing authentic evidence

Competency-based assessment begins with an explicit claim: a learner can perform specified work to an accepted standard. Assessment quality depends on the alignment between that claim, the task, the conditions of performance, the evidence collected and the criteria applied. Baartman et al. (2007) argue that competence assessment should be judged through an integrated set of criteria including fitness for purpose, comparability, authenticity, cognitive complexity, meaningfulness, fairness, transparency, educational consequences and reproducibility of decisions.

Authenticity is not achieved merely by labelling an activity practical. Gulikers et al. (2006) conceptualise authentic assessment through the task, physical context, social context, result or form and assessment criteria. A practical task can remain weak if it uses unrealistic conditions, omits critical process evidence or applies criteria that do not represent occupational quality. Conversely, resource constraints may compel educators to simplify or alter tasks, creating a tension between feasibility and validity. Assessment literacy requires the assessor to recognise this tension and determine whether the available evidence remains adequate for the competence claim.

Recent vocational evidence reinforces the need to compare professional perceptions with actual assessment artefacts. In Chilean vocational education, Villarroel et al. (2024) found that teachers often perceived assessment as work-related, while analysis of written examinations revealed limited realism and predominantly low cognitive complexity. Welsandt et al. (2024) similarly demonstrate that authentic competence assessment requires deliberate construction of the domain, scenario, evidence demands and test environment. These studies show why authenticity cannot be inferred from the practical label or from assessor confidence alone.

Reviews of vocational assessment indicate that competency-based and performance-oriented approaches are widely used, but their quality depends on the relation between assessment design and intended outcomes (Garraway, 2022; Yusop et al., 2022). Vocational assessment literacy consequently includes the ability to move from broad unit outcomes to observable evidence, select complementary methods, create judgement guides and anticipate the practical conditions necessary for a fair demonstration of competence.

2.3 Professional judgement, evidence sufficiency and fairness

Competence judgement cannot be reduced to mechanical scoring. Assessors must decide whether evidence is valid, authentic, current and sufficient; whether the performance meets the relevant criteria; and whether the decision would be defensible to learners, moderators, regulators and employers. Marking guides and rubrics can support consistency, but their value depends on the specificity and occupational relevance of their criteria. Vague descriptors merely relocate inconsistency from the assessor's mind into a document.

Professional judgement is therefore both indispensable and risky. It is indispensable because complex performance cannot always be captured through atomised indicators. It is risky because personal expectations, occupational traditions, prior experience and institutional culture can shape what an assessor notices and values. The multiple roles and identities enacted by vocational assessors reinforce the need for shared standards and professional dialogue (Leonardsen et al., 2022). Exemplars, collaborative task review and calibration make the grounds of judgement more visible without pretending to eliminate judgement.

Fairness extends beyond treating every learner identically. It requires clear expectations, reasonable access to the resources needed to demonstrate competence, appropriate assessment conditions and attention to sources of irrelevant difficulty. In practical assessment, shortage of equipment or insufficient practice opportunities may mean that performance partly reflects institutional deprivation rather than learner competence. Ethical assessment literacy therefore includes the capacity to identify when assessment conditions threaten the validity or equity of the inference being made.

2.4 Feedback, records and reflective improvement

Assessment literacy includes the capacity to communicate decisions and use evidence to support improvement. Feedback should help learners understand the expected standard, recognise the gap between current and desired performance and identify a credible next step. In vocational learning, feedback may address procedural sequence, tool handling, safety, accuracy, efficiency and professional conduct as well as conceptual understanding. A mark or competent/not-yet-competent label alone provides limited support for development.

Records are similarly more than administrative residue. Assessment tools, criteria, observation records, learner evidence, decisions and review comments create an audit trail through which an institution can defend its judgements. Yet record keeping can become compliance-oriented when files are completed without reflection or follow-through. This article therefore treats record literacy as a dimension of assessment capability while keeping the institutional use of aggregated results outside its analytical scope.

2.5 Institutional capacity and professional learning

An individual-deficit explanation assumes that weak assessment can be corrected by sending assessors to a course. The assessment-literacy literature suggests a more demanding process. Reviews of improvement programmes identify contextualised content, active and collaborative learning, continuity and attention to assessor identity as recurring features of effective development (Geng & Mohd Yusoff, 2023). Recent VET research likewise shows that teachers develop formative-assessment capability through collaborative, practice-embedded work on learning progressions and assessment activities rather than through decontextualised instruction alone (Baartman & Gulikers, 2025). Assessment knowledge becomes usable through repeated task design, feedback, mentoring, peer review, calibration and reflective examination of consequences.

Institutional capacity therefore includes more than the aggregate knowledge of employees. It comprises shared standards, role clarity, accessible policies, leadership expectations, mentoring, opportunities to practise and receive feedback, time, resources, quality-assurance support and a culture in which critique is interpreted as learning rather than punishment. Assessment literacy becomes institutional when knowledge is retained in routines and artefacts, distributed through collaboration, renewed through professional learning and protected through leadership and accountability. African TVET evidence further indicates that internal assessment is affected by staff preparation, time, infrastructure, equipment and governance arrangements (Maja & Atukunda, 2022; Mutebi & Ferej, 2023).

This beyond-classroom orientation is supported by recent research in Dutch higher vocational education, where teachers regarded assessment competence as involving not only classroom decisions but also alignment, collaboration, updating and improvement across the wider programme context (Meijer et al., 2025). The finding is important for the present study because assessment capability in vocational institutions must connect occupational standards, programme expectations, material conditions and collective quality responsibilities.

2.6 Evidence, measurement and the limits of individual-deficit explanations

Assessment-literacy research has used knowledge tests, self-report surveys, scenario tasks, interviews, observations and artefact analysis. Each method reveals a different aspect of capability. Knowledge tests can identify misconceptions about validity, reliability or scoring, but they may privilege declarative knowledge detached from occupational context. Self-report tools reveal beliefs and confidence, yet respondents may overestimate or underestimate their capability. Interviews provide access to professional reasoning but remain retrospective accounts. Observation offers stronger evidence of enactment, although the assessor's behaviour may change when being observed. Artefact analysis reveals what assessment designs and records make visible, but a good document may be used poorly and an incomplete document may be supported by tacit dialogue not captured in the file.

These methodological differences matter when institutions decide how to develop assessors. An individual-deficit interpretation can emerge when a single measure is treated as a complete diagnosis. A low knowledge score may lead to generic retraining even when the main barrier is lack of equipment, time or access to standards. Conversely, strong confidence or successful course completion may obscure weaknesses in actual instruments

and judgement. A defensible institutional diagnosis therefore requires triangulation: what the assessor explains, what the artefact shows, how the judgement is enacted, how learners experience the process and how the institution supports or constrains the work.

The present study could not include every evidence level, but its combination of interview accounts and documentary artefacts was sufficient to reveal discrepancies between aspiration and institutionalisation. Participants often articulated appropriate principles while documents varied in alignment, criterion detail and completeness. This pattern is not evidence of hypocrisy or carelessness. It demonstrates that assessment practice is mediated by workload, resources, policy enactment and professional interaction. The analytical task is therefore to identify where the translation from knowledge to practice is interrupted and what institutional response is most likely to restore it.

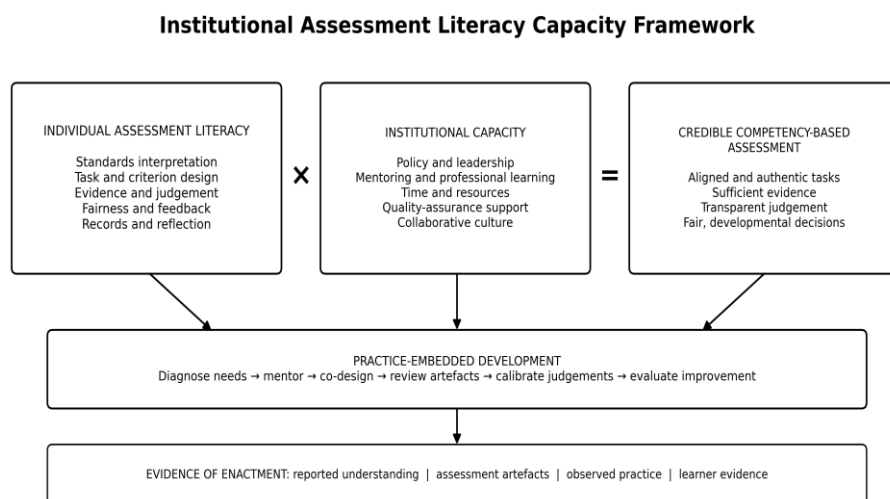
CONCEPTUAL FRAMEWORK: ASSESSMENT LITERACY AS INSTITUTIONAL CAPACITY

The conceptual framework distinguishes two interdependent domains. Individual assessment literacy comprises knowledge of assessment purposes and competency standards; planning and instrument design; method selection; evidence requirements and authenticity; validity, reliability and fairness; professional judgement; feedback; record keeping; ethical responsibility; and reflective improvement. Institutional assessment capacity comprises policy clarity, leadership, professional development, mentoring, peer collaboration, time, resources, access to standards, quality-assurance support and organisational culture.

The framework adopts a multiplicative rather than additive logic: individual assessment knowledge and judgement $\times$ institutional support and professional learning = effective assessment capacity. The multiplication sign communicates that serious weakness in either domain constrains the whole. High individual knowledge may not produce authentic, fair assessment where resources, time or policy support are absent. Strong systems and templates cannot produce defensible decisions when assessors do not understand evidence, criteria or standards. Effective capacity is realised through interaction in practice.

The framework also distinguishes evidence levels. Reported understanding refers to participants' explanations of purposes, standards or procedures. Perceived capability refers to their own sense of competence. Documentary enactment refers to qualities visible in assessment tools, criteria, schedules, files and review records. Observed enactment would refer to direct evidence of assessment behaviour and learner interaction, but the parent study did not include classroom or workshop observation. This distinction protects the analysis from treating interview confidence as demonstrated competence or treating a document as a complete account of practice.

Figure 1. Institutional Assessment Literacy Capacity Framework.



Assessment expertise becomes institutional when it is shared, supported, reviewed and renewed.

Source: Developed from the study findings and assessment-literacy literature.

METHOD

4.1 Research design and setting

This article reports a focused analysis of assessment-literacy evidence from a broader doctoral study on assessment practices as instruments of internal quality assurance. The parent study was located within an interpretivist paradigm and used a qualitative multiple-case design (Yin, 2018). Three vocational training institutions in the Oshana Region constituted the cases. The institutions operated under a common national vocational education and training environment but differed in programme scope, organisational arrangements, resources and professional routines. To protect confidentiality, they are reported as Institutions A, B and C.

4.2 Participants and sampling

Twenty participants were selected through criterion-based purposive sampling: 14 educators directly involved in planning, administering, marking, giving feedback and maintaining assessment records, and six quality-assurance officers responsible for coordinating, monitoring or supporting assessment-related quality processes. Participants were selected because they could provide experience-based accounts of how competency-based assessment was understood and enacted. The article does not infer differences by occupational field, qualification, certification status or years of experience because the available study materials did not provide a sufficiently complete profile for such analysis.

4.3 Data generation

Data were generated through semi-structured interviews and document analysis. Interviews explored how participants understood assessment purposes, developed tasks, combined written and practical methods, judged competence, used criteria, provided feedback, documented decisions and experienced policies, professional learning, workload, resources, leadership and collaboration. Interviews were audio-recorded with permission and transcribed for analysis.

Documentary material included available assessment policies and guidelines, assessment schedules, written and practical assessment tools, marking guides or criteria, marksheets, moderation records and assessment files. Documents were examined as assessment artefacts rather than as proof that the recorded procedure had been enacted exactly as written. The analysis considered whether tools made outcomes and competencies visible, whether criteria supported judgement, whether files were complete enough to trace decisions and whether records showed opportunities for professional review and learning.

4.4 Focused analytic procedure

The parent dataset was analysed through reflexive thematic analysis (Braun & Clarke, 2022), combining within-case and cross-case interpretation. For this article, all interview excerpts and documentary observations relating to assessment knowledge, task design, method selection, evidence, judgement, feedback, fairness, records, development needs and institutional support were extracted into an evidence matrix. Initial codes were organised around assessment-literacy dimensions but remained open to patterns generated by the data. The analysis then examined convergence, complementarity and tension among participant accounts, documentary artefacts and institutional conditions.

The article applies an evidence-status rule. Claims are described as reported understanding where supported by participant explanation; as documentary enactment where visible in artefacts; and as an evidence gap where neither interviews nor documents established the practice sufficiently. This rule was especially important for formal assessment blueprinting, accommodations for learner diversity, direct observation of feedback quality, assessor certification, measured knowledge and inter-assessor agreement.

4.5 Trustworthiness and ethics

Trustworthiness was supported through methodological triangulation, cross-case comparison, attention to discrepant evidence, maintenance of an analytic trail and use of quotations to show the basis of interpretation

(Lincoln & Guba, 1985). The analysis did not classify individuals or institutions as literate or illiterate. Instead, it examined differentiated strengths, gaps and enabling conditions. Ethical approval was obtained through the relevant university research ethics structures, formal permission was secured from the participating institutions and participants provided written informed consent. Role- and institution-based identifiers were used in reporting.

FINDINGS

5.1 Evidence profile: literacy was present but unevenly enacted

The evidence did not support a simple conclusion that participants were either assessment literate or assessment illiterate. Instead, it showed a differentiated profile. Participants generally articulated purposes associated with competence, learner improvement, fairness and credibility. They described practices central to competency-based assessment, including combining theory and practice, observing procedures, using criteria and providing feedback. Documentary evidence nevertheless showed variation in the alignment of tasks, the specificity of criteria and the completeness of records. Institutional conditions helped explain why stated understanding did not always become consistent practice.

The strongest empirical claims arose where accounts and artefacts converged. For example, participants emphasised criteria, and document analysis identified some assessment tools with specific marking guides. The same review also identified general or weakly linked criteria, showing that the capacity was present but not institutionalised uniformly. The most limited claims concerned formal blueprinting, systematic evidence-sufficiency protocols, comprehensive inclusive-assessment practice and direct observation of assessor behaviour. These areas are reported as evidence gaps rather than presumed deficiencies.

5.2 Understanding assessment purposes and competency standards

Educators commonly understood assessment as a process for establishing whether learners could demonstrate expected knowledge and practical skill. Their accounts extended beyond the production of marks. Assessment was described as evidence of readiness, a means of identifying weakness and a basis for feedback. One educator explained: “When I prepare an assessment, I first look at what the learners were supposed to achieve in that unit. The task must come from the outcomes, otherwise you may assess something that was not part of the training” (Educator 2, Institution A).

Participants also understood that vocational competence required more than theoretical recall. “For practical assessment, we look at the competence,” one educator noted. “The question is not only whether the learner knows the answer, but whether the learner can perform the task according to what is expected” (Educator 4, Institution B). Another emphasised the need to combine theoretical and practical evidence because assessing theory alone could miss the learner’s practical ability. These accounts show a strong reported orientation towards competency alignment and the integration of knowledge and performance.

Documentary evidence partly corroborated this understanding. Some tools showed visible links to module content, outcomes or practical competencies and combined written questions with practical demonstrations or assignments. Other tools listed tasks without clearly identifying the outcome or competence to which each task was linked. The study therefore showed conceptual awareness of alignment, but uneven documentary visibility of the reasoning through which outcomes were translated into tasks.

5.3 Assessment planning, task design and authenticity

Educators reported that task design was guided by learning outcomes, content covered, practical activities and the competence learners were expected to demonstrate. The use of both continuous and final assessment was also described as important. Continuous assessment allowed educators to identify problems before a final decision, while final assessment confirmed whether the expected level had been attained. Assessment schedules supported the presence of tests, assignments, practical activities and final assessments across the training period.

The use of written and practical methods represented an important strength. In practical assessment, educators described observing both process and product. “For written assessment, learners sit and answer questions. But for practical assessment, we observe what they are doing. We look at the process and the final product” (Educator 2, Institution B). This account indicates awareness that practical competence cannot be inferred from a final answer or product alone.

Authenticity, however, was constrained by available resources. Educators explained that they sometimes adjusted tasks according to tools, materials, time and learner numbers. “Sometimes you may want to design a very practical task, but you also must look at the equipment and materials available. You design what is possible within the situation” (Educator 5, Institution B). Such adjustment may be necessary, but it creates a validity question: whether the revised task still produces sufficient evidence for the competence claimed. Documents showed practical requirements but did not consistently record whether the necessary conditions were available at the time of assessment.

5.4 Competence judgement, criteria and fairness

Educators judged competence through marking guides, assessment criteria, practical observation and professional judgement. In written assessments, guides structured the allocation of marks. In practical assessments, educators reported attending to procedure, safety, accuracy, tool handling, time management and quality of output. One educator stated: “In practical work, we do not only look at the final product. We also look at the process, because the learner may get something right by chance but fail to follow the correct procedure” (Educator 4, Institution B).

Criteria were valued for reducing guesswork and supporting explanation to learners. “The marking guide helps us to know what to look for. It reduces guessing because you are guided by the criteria” (Educator 3, Institution C). Another participant noted that clear criteria made it easier to explain why a particular mark or decision had been reached. This represents a strong reported understanding of transparency and defensibility.

Documentary evidence revealed variation. Some marking guides and criteria sheets were specific and offered clear direction; others were general and left substantial space for individual interpretation. The issue was not the presence or absence of professional judgement, which remained necessary, but whether shared criteria sufficiently disciplined that judgement. Where criteria were broad, competence decisions became more dependent on the assessor’s personal standards and were harder to explain or compare.

Participants also understood fairness as giving learners a genuine opportunity to demonstrate competence. Resource constraints and limited practice time were recognised as possible threats. “The learner may know what to do, but if there was not enough time to practise, the performance may be affected,” an educator explained (Educator 7, Institution B). The available evidence did not establish systematic knowledge or implementation of accommodations, bias review or appeals. Inclusive-assessment literacy therefore remains an important but under-evidenced dimension.

5.5 Feedback and record literacy

Educators described feedback as part of the learning function of assessment. Continuous assessment was valued because it could identify weaknesses early and allow support before final decisions. Participants explained that feedback helped learners correct errors and enabled educators to adjust teaching strategies. These accounts indicate a strong developmental orientation, but the parent study did not directly observe the timing, specificity or learner uptake of feedback. The finding therefore concerns reported purpose and practice rather than measured feedback quality.

Record keeping was widely understood as necessary for accountability and defence of decisions. Marksheets, tools, guides, assessment files and moderation records provided evidence that assessment had occurred. “If there is a query later, the record helps us to explain how the learner was assessed and what mark was awarded” (Educator 4, Institution C). Quality-assurance officers similarly expected a file to show the assessment tool, marking guide, results and review evidence.

Document analysis showed that assessment records existed, but completeness differed. Some files brought together the principal components of the assessment process; others lacked dates, comments, signatures or evidence of follow-up. This variation demonstrates that record literacy is not simply the ability to complete a form. It includes understanding which evidence is necessary for a decision to be traceable and organising that evidence so that it can support professional review. Detailed institutional use of aggregated assessment results lies outside this article's analytical focus.

5.6 Development needs and the distinction between subject expertise and assessment expertise

Participants explicitly distinguished occupational knowledge from assessment expertise. A quality-assurance officer observed: "Some trainers know their subject very well, but assessment is another skill. You can know the trade, but still need support in setting a good task, preparing criteria and showing how competence will be judged" (Quality Assurance Officer 2, Institution B). This statement captures the central assessment-literacy problem: occupational credibility is necessary but does not automatically provide knowledge of validity, evidence, task design or transparent judgement.

Educators also identified continuing development needs. "We need more training on how to prepare competency-based assessments and how to use rubrics properly. Sometimes people assess, but they are not very sure whether the tool is measuring the competence correctly" (Educator 6, Institution B). Another educator stated that assessment was not merely setting questions; it required knowing whether the task measured what it was intended to measure. These accounts indicate awareness of limitations rather than a lack of concern for quality.

The evidence did not include a standardised assessment-literacy test, an inventory of completed assessor qualifications or direct observation of post-training practice. It therefore cannot rank participants' actual knowledge or claim that formal training had or had not been effective. It supports a more precise conclusion: participants recognised specialised assessment demands, artefacts showed uneven enactment and staff development needed to be ongoing and connected to real assessment work.

5.7 Policy, leadership and quality-assurance support

Participants viewed clear policies and guidelines as resources for shared practice. Procedures could indicate when tools should be submitted, how marking guides were prepared, how moderation was organised and how records were filed. "When there is a clear assessment policy, everyone knows what is expected. It guides how assessments are prepared, checked, marked and recorded" (Quality Assurance Officer 3, Institution C). Educators similarly argued that guidelines reduced the risk that every trainer would work according to an individual procedure.

The existence of policies was not considered sufficient. "Having a policy is not enough. The important thing is whether people are using it. If the policy is only in the file, it will not improve assessment" (Quality Assurance Officer 4, Institution A). Document analysis found templates, moderation forms, schedules and filing procedures but variable evidence of comprehensive policy clarity and consistent follow-up. Policy functioned as institutional capacity only when understood, enacted and monitored.

Leadership shaped whether assessment quality was treated as an institutional priority. Quality-assurance officers linked management support to time for review, submission of documents, staff development and follow-up. "If leadership takes assessment seriously, trainers also take it seriously. But if there is no follow-up, people may treat it casually" (Quality Assurance Officer 2, Institution C). Signatures from heads of department, moderators or quality personnel showed that support structures existed, yet records did not always demonstrate developmental comments or action. Leadership presence was therefore visible, but its effect depended on whether it enabled learning and accountability rather than only authorisation.

5.8 Workload, resources and professional collaboration

Assessment literacy was enacted under material and temporal constraints. Practical assessment required tools, equipment, consumables, workshops and time. "For practical assessment, resources are very important.

You cannot assess a learner properly if the tools or materials are not enough. The learner needs a fair chance to show the skill” (Educator 3, Institution B). A quality-assurance officer added that inadequate resources could weaken the evidence because assessment might not fully represent the expected competence.

Time was a hidden assessment resource. Participants described preparation, administration, observation, marking, feedback, documentation and review as competing demands. “Assessment takes time if it is done properly,” an educator explained; “when time is limited, some of these things become rushed” (Educator 1, Institution A). Workload did not merely inconvenience staff. It influenced whether professional knowledge could be applied with sufficient depth.

Collaboration helped distribute assessment knowledge. “When we discuss assessment with other trainers, it helps. You may think your task is clear, but a colleague can see something you missed” (Educator 7, Institution A). Participants also linked collaboration to the use of similar standards. Feedback culture mattered: where comments were interpreted as professional support, educators were more likely to revise tools and learn; where they were experienced as personal criticism, review became defensive. Assessment literacy therefore emerged as a property of relationships and routines as well as individuals.

Table 1. Evidence matrix for dimensions of assessment literacy and institutional capacity.

Assessment-literacy dimension	Empirical evidence	Interpretation	Evidence status
Purposes and standards	Participants linked assessment to outcomes, practical competence, feedback and readiness.	Strong conceptual orientation to competency-based purposes.	Reported understanding; partial documentary convergence.
Planning and task design	Tools included written and practical tasks; some showed outcome links, others did not.	Alignment capability was present but unevenly institutionalised.	Interview and documentary evidence.
Method selection and authenticity	Participants valued theory plus practical demonstration; tasks were adjusted to available resources.	Multiple methods were understood, but authenticity was institutionally constrained.	Reported practice; conditions not directly observed.
Evidence and judgement	Assessors described process, product, safety, tool handling and procedures; criteria varied in detail.	Professional judgement was informed but not uniformly disciplined by explicit criteria.	Interview and documentary evidence.
Fairness and inclusion	Participants linked fairness to clear expectations, resources and opportunity to demonstrate skill.	Fairness awareness was evident; accommodations and bias review were under-evidenced.	Reported understanding; major evidence gap.
Feedback	Educators described feedback as identifying weakness and supporting correction.	Strong developmental orientation.	Reported practice; quality and timing not observed.
Records	Files, marksheets, tools and guides existed but completeness varied.	Record literacy was important but inconsistently enacted.	Documentary evidence.
Professional learning	Participants requested continuing development and distinguished trade expertise from assessment expertise.	One-off training is unlikely to be sufficient.	Interview evidence.

Assessment-literacy dimension	Empirical evidence	Interpretation	Evidence status
Institutional support	Policies, leadership, QA support, resources, workload and collaboration shaped practice.	Assessment literacy functioned as distributed institutional capacity.	Strong cross-case theme.

DISCUSSION

6.1 Strong conceptual orientation, uneven institutionalisation

The findings answer the first research question by showing that the strongest dimensions of reported assessment literacy concerned purposes, the integration of theory and practice, observation of practical performance, the value of criteria, feedback and traceability. Participants understood assessment as evidence of competence rather than as marks alone. This orientation is consistent with the wider movement from technical assessment knowledge towards situated professional capability (Willis et al., 2013; Xu & Brown, 2016; Pastore & Andrade, 2019).

The weakness was not a complete absence of relevant knowledge. It was uneven institutionalisation. Some artefacts demonstrated alignment and specific criteria; others obscured the competency link or left judgement largely individual. Some files supported traceability; others were incomplete. The evidence therefore resists deficit language that labels trainers as incapable. It indicates that assessment capability was present in pockets but not reliably retained and reproduced across institutional practice. This pattern resembles Villarroel et al.'s (2024) finding that favourable professional perceptions of authenticity may coexist with assessment artefacts that make limited demands for realistic application. The implication is not that practitioners' accounts are false; rather, perceptions and artefacts represent different evidence levels and should be interpreted together.

6.2 The knowledge–practice gap must be analysed through evidence levels

The article's methodological contribution is the distinction among reported understanding, perceived capability, documentary enactment and observed practice. The study found thoughtful accounts of standards, authenticity and fairness, but it did not administer a knowledge test or observe assessment events. Conversely, artefact analysis revealed features of enacted capacity but could not establish how every task was administered or how assessors reasoned in real time. Treating any one evidence source as complete would overstate the conclusion.

This distinction responds to a recurring problem in assessment-literacy research. Self-report instruments can identify beliefs and confidence, while knowledge tests can identify declarative understanding, but neither necessarily reveals practice. Artefacts provide stronger evidence of design and documentation but may not capture dialogue, judgement or adaptation. Future studies of vocational assessment literacy should combine these evidence levels instead of treating a single score or training certificate as proof of competence.

6.3 Assessment literacy is distributed institutional capacity

The third research question is answered most directly by the interaction among policy, leadership, resources, workload, quality-assurance support and professional culture. Participants could understand authentic assessment but remain unable to implement it where equipment and materials were insufficient. They could value feedback but have insufficient time to provide it. They could possess marking guides but lack opportunities to discuss what criteria meant in borderline cases. Assessment quality was consequently produced through a system of people, standards, artefacts, resources and relationships.

This distributed view extends assessment-literacy models into vocational institutional analysis. Individual knowledge remains indispensable, but it becomes durable organisational capacity only when it is retained in common procedures, shared exemplars, mentoring relationships, peer review routines and traceable assessment artefacts. Policies without enactment, templates without reasoning and training without supported practice

cannot substitute for this infrastructure. The multiplicative logic of the framework therefore fits the empirical pattern better than an additive checklist of staff competencies.

6.4 Why one-off training is insufficient

Participants requested more training, but the findings indicate that professional development must be designed carefully. A workshop may introduce principles of validity, rubrics or competency standards, yet complex capabilities develop through repeated application and feedback. Assessors need opportunities to design tasks, receive critique, observe or discuss performance evidence, compare judgements and revise artefacts. Without these opportunities, new knowledge may remain detached from the conditions under which assessment occurs. This interpretation is consistent with Geng and Mohd Yusoff (2023) and with Baartman and Gulikers (2025), whose work emphasises contextualised, collaborative and sustained development connected to real assessment activity.

The study also shows why development should begin diagnostically. Staff may have different strengths: occupational interpretation, task design, criteria development, feedback, record keeping or fairness. A generic course risks repeating what some participants already know while leaving specific practice gaps unresolved. Diagnostic review of real assessment tools and decisions provides a more credible basis for mentoring and continuing professional development than attendance records alone.

6.5 Confidence, expertise and professional judgement

The findings did not include a direct measure of confidence, but they show why confidence should not be equated with competence. Educators who recognised uncertainty about rubrics or alignment demonstrated reflective awareness rather than failure. Conversely, a confident assessor may still rely on broad criteria or adapt tasks in ways that weaken the intended inference. Continuing competence should therefore be evidenced through the quality of artefacts, reasoning, feedback and participation in professional review.

Vocational assessors also enact multiple identities: occupational expert, teacher, judge, supporter of learners and representative of institutional or regulatory standards (Leonardsen et al., 2022). These roles can generate tension. An assessor may wish to support a learner while protecting a non-negotiable safety standard; may value authentic practice while facing resource constraints; or may use occupational tacit knowledge that is difficult to express in criteria. Assessment literacy includes the capacity to make these tensions visible and resolve them ethically rather than concealing them behind a score.

6.6 Fairness is institutional as well as interpersonal

The Oshana evidence broadens fairness beyond individual bias. Access to tools, practice time, materials and appropriate conditions shaped whether learners had a legitimate opportunity to demonstrate competence. This means that fairness cannot be secured only through assessor goodwill. Institutional leaders must ensure that the conditions of assessment do not systematically introduce construct-irrelevant disadvantage. Where resources are inadequate, institutions should document the risk, reschedule, provide equivalent resources or redesign the assessment without weakening the competence standard.

The limited evidence on accommodations, appeals and bias review is itself significant. It identifies a capacity domain requiring further investigation and development. A mature assessment-literacy framework should include ethical and inclusive decision-making, not only task construction and scoring. Learners' perspectives are also necessary to determine whether expectations and feedback are experienced as clear, respectful and fair.

6.7 Boundaries with moderation and evidence-use research

Moderation and institutional evidence use are analytically adjacent to assessment literacy but are not the central outcomes examined here. Moderation is considered only where peer review, feedback and shared standards contribute to professional learning. Record keeping is considered as a capability needed to make decisions traceable, rather than as a full analysis of how aggregated results are converted into programme review.

This scope preserves the article's distinctive contribution: explaining the professional capabilities and institutional conditions required for credible competency-based assessment.

6.8 Institutional memory, professionalisation and regulatory responsibility

Treating assessment literacy as institutional capacity also raises the question of professional continuity. Vocational institutions are vulnerable when assessment expertise resides only in a small number of experienced staff. Retirement, transfer, turnover or programme expansion can remove tacit knowledge about standards, recurring task-design problems and acceptable evidence. Institutional memory must therefore be created deliberately through reviewed assessment banks, annotated exemplars, records of common judgement problems, mentoring arrangements and induction processes. Such resources should not freeze practice into inflexible templates; they should provide a defensible starting point from which assessors can adapt to new technologies, occupational practices and learner needs.

Professionalisation also requires clearer recognition of assessment work. Task design, observation, judgement, feedback and record keeping are often treated as activities that can be absorbed within ordinary teaching time. The findings show that this assumption encourages compression and uneven enactment. Institutions should recognise assessment responsibilities in workload models, role descriptions and performance-development systems. Continuing competence should be reviewed periodically through a portfolio of assessment work and participation in professional dialogue rather than inferred indefinitely from an earlier qualification.

Regulators and qualification authorities have a complementary responsibility. They should set clear expectations for validity, evidence, fairness and documentation while avoiding systems that reward paperwork volume over judgement quality. Guidance should include practical exemplars from different occupational areas, minimum expectations for evidence and criteria, and developmental support for providers with limited capacity. External quality assurance is most useful when it strengthens internal capability: identifying systemic risks, supporting institutions to analyse their own artefacts and confirming that development has improved practice. A punitive inspection culture may produce defensive compliance, whereas a developmental regulatory relationship can strengthen both public accountability and professional learning.

This wider professionalisation agenda is consistent with the study's central conclusion. Assessment literacy is not a personal possession acquired once and then carried unchanged through a career. It is a capability that must be refreshed as occupational standards, technologies, learners and institutional conditions change. The institution and the individual assessor therefore share responsibility for maintaining competence, while the regulator has responsibility for creating standards and support that make credible practice possible.

INSTITUTIONAL ASSESSMENT LITERACY CAPACITY FRAMEWORK

The fourth research question asked what framework was warranted by the findings. The proposed Institutional Assessment Literacy Capacity Framework is cyclical rather than course-based. It begins with diagnosis of actual practice, develops capability through supported enactment and evaluates whether learning is visible in improved tools, judgements, feedback and records. The framework does not assume that all assessors require identical support or that institutional conditions can be separated from professional knowledge.

7.1 Diagnostic review

Institutions should begin by reviewing a sample of assessment tools, marking guides, observation records, learner evidence and feedback. The purpose is developmental diagnosis, not fault-finding. Review should identify strengths and needs across standards interpretation, authenticity, evidence sufficiency, criteria, fairness, feedback and documentation. Staff self-assessment can inform the process but should be triangulated with artefacts and professional dialogue.

7.2 Modelling, mentoring and supervised practice

New or developing assessors should work with experienced colleagues to translate unit standards into assessment claims, tasks, conditions and evidence requirements. Mentoring should include co-design,

observation or review of practical assessment, discussion of borderline evidence and revision of tools. A mentor's role is to make tacit judgement visible rather than simply approve completed forms.

7.3 Peer review and calibration

Programme teams should periodically review tools before use and discuss anonymised examples of learner evidence. Assessors should first judge independently, then compare reasoning and identify the standards underlying agreement or disagreement. These activities distribute knowledge, reduce isolation and create common interpretive resources. Moderation feedback should feed into this learning cycle, while the present analysis remains focused on its capacity-development function rather than the full implementation of moderation.

7.4 Enabling conditions

Leadership must allocate time, clarify roles, provide access to standards, ensure practical resources and authorise quality-assurance staff to follow up development needs. Policies should specify expected evidence while leaving room for contextually justified professional judgement. A constructive feedback culture is essential because assessors learn when critique is legitimate, specific and connected to improvement rather than punishment.

7.5 Evaluation and institutional memory

Capacity development should be evaluated through change in practice rather than attendance. Institutions should examine whether revised tasks are better aligned, criteria are more explicit, learner evidence is more sufficient, feedback is more useful and records are more traceable. Strong artefacts and resolved examples should be retained in an assessment-resource repository to support induction, mentoring and future standardisation. In this way, expertise becomes an institutional resource that can survive staff turnover.

Table 2. Capacity gaps and institutional interventions.

Capacity gap	Institutional intervention	Responsible actor	Evidence of improvement
Weak outcome–task visibility	Use a design template linking competence, task, conditions, evidence and criteria; review before use.	Programme team and QA officer	Revised tools show explicit alignment and evidence requirements.
Variable judgement criteria	Co-develop detailed guides and review anonymised learner evidence.	Assessors, mentors and moderators	Greater criterion specificity and explainable decisions.
Constrained practical authenticity	Audit equipment, materials, time and practice opportunity before assessment.	Institutional leadership and programme head	Assessment conditions support the intended competence claim.
Training detached from practice	Use diagnostic artefact review, mentoring, co-design and supervised assessment.	Assessor-training provider and institution	Improved tools, feedback and records after development.
Individualised judgement	Hold standardisation and calibration sessions using common evidence.	Programme team and QA unit	Shared reasoning and reduced unexplained variation.
Incomplete records	Define a minimum defensible assessment file and provide administrative support.	QA unit and records staff	Complete, accessible and traceable files.

Capacity gap	Institutional intervention	Responsible actor	Evidence of improvement
Weak feedback culture	Train staff in evidence-based collegial critique and protect developmental review.	Institutional leaders	Feedback leads to revision rather than defensive compliance.
Workload pressure	Allocate explicit time for design, observation, feedback, review and documentation.	Management	Fewer rushed or incomplete assessment processes.
Capacity loss through turnover	Create an assessment-resource repository of standards, exemplars and reviewed tools.	QA unit and programme teams	Institutional memory supports induction and consistency.

IMPLICATIONS FOR POLICY AND PRACTICE

Institutional leaders should establish assessment capability as a strategic quality priority. They should allocate workload for task development, practical observation, feedback, peer review and documentation; ensure that resource shortages do not invalidate assessment; and require evidence that professional-development needs have been addressed. Leadership monitoring should focus on quality and learning, not only timely submission of forms.

Trainers and assessors should engage in continuing practice-based development. They should explicitly map tasks to competencies, define the evidence needed, combine methods appropriately, use clear judgement guides, document important reasoning and reflect on feedback. Occupational expertise should be treated as the foundation for assessment practice, not as a substitute for assessment expertise.

Quality-assurance officers should act as capacity developers as well as compliance monitors. Their responsibilities should include diagnostic review of assessment artefacts, coordination of mentoring and calibration, identification of recurring capability needs and evaluation of whether development leads to improved practice. Regulators and qualification authorities should provide accessible standards, exemplars and guidance on evidence sufficiency, fairness and authentic assessment.

Assessor-training providers should move beyond short generic courses. Programmes should include authentic vocational cases, co-design of tools, review of learner evidence, supervised practice and post-course mentoring. Researchers should develop multidimensional studies that combine knowledge tasks, professional accounts, artefact analysis, observation and learner perspectives, and should test whether institutional interventions improve assessment quality over time.

CONTRIBUTION TO THEORY, EVIDENCE AND PRACTICE

Theoretically, the article reframes vocational assessment literacy as a multiplicative interaction between individual knowledge and institutional capacity. It extends situated assessment-literacy models by identifying the organisational conditions through which knowledge becomes durable and shareable practice. It also distinguishes evidence levels so that reported understanding is not confused with demonstrated competence.

Empirically, the article provides evidence from an under-researched Namibian setting. The three cases showed substantive professional awareness alongside uneven assessment artefacts, demonstrating why simple individual-deficit or individual-competence narratives are inadequate. The evidence identifies task alignment, criterion specificity, practical authenticity, documentation and reflective follow-through as important sites where literacy becomes visible.

Practically, the Institutional Assessment Literacy Capacity Framework connects diagnostic review, mentoring, co-design, peer calibration, enabling conditions and evaluation. It offers institutions a pathway for

strengthening competence without reducing development to one-off workshops or reducing quality assurance to document completion.

LIMITATIONS AND FUTURE RESEARCH

The qualitative multiple-case design provides contextual depth but does not support statistical generalisation. The study relied on interviews and documentary artefacts and did not include a standardised assessment-literacy test, direct observation of assessment events, learner perspectives or formal analysis of inter-assessor decision consistency. Interview accounts may have been influenced by recall or social desirability, and the documentary corpus was limited to materials made available by the institutions.

The findings therefore distinguish reported understanding from documentary enactment and avoid treating either as a complete measure of competence. Future research should combine knowledge tasks, artefact analysis, observed practice and longitudinal evaluation of capacity-development interventions. It should also examine learner experiences of clarity, feedback and fairness; compare occupational fields and institution types; and test whether strengthened capacity improves consistency, authentic performance and stakeholder confidence in competence decisions.

CONCLUSION

Assessment literacy in competency-based vocational education is not adequately represented by attendance at an assessor course, possession of occupational expertise or confidence in one's own practice. The Oshana cases showed meaningful professional understanding of competence, practical evidence, criteria and feedback, but also uneven task alignment, criterion quality, documentation and reflective follow-through. These variations were shaped by policies, leadership, quality-assurance support, workload, resources and professional culture.

The article therefore concludes that assessment literacy is an organisational resource produced through interaction between individual judgement and institutional conditions. Institutions strengthen credibility when they make standards interpretable, provide authentic assessment conditions, support assessors through mentored practice and peer calibration, and convert review into professional learning. The proposed framework offers a pathway for doing so while recognising existing institutional strengths.

The central implication is that strengthening competency-based assessment requires more than correcting individual educators. It requires institutions to create the conditions in which assessment knowledge can be enacted, scrutinised, shared and renewed. When that occurs, assessment literacy becomes part of institutional capacity to generate fair, transparent and defensible evidence of vocational competence.

DECLARATIONS

Ethical approval: Ethical approval was obtained through the relevant university research ethics structures, and the participating vocational training institutions granted formal permission for the study.

Informed consent: All participants received information about the study and provided written informed consent. Participation was voluntary, and anonymised quotations are reported using role and case identifiers.

Funding: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Conflict of interest: The authors declare no competing interests.

Data availability: The interview transcripts and institutional documents contain potentially identifiable professional and institutional information. De-identified data may be made available by the corresponding author on reasonable request, subject to ethical approval, participant confidentiality and institutional permission.

Author contributions: Kadonsi Kaziya: conceptualisation, supervision, methodology, interpretation, writing—original draft, and writing—review and editing. Moses Moses Lungameni: investigation, data curation, formal analysis, and writing—review and editing. Both authors approved the final manuscript.

Acknowledgements: The authors are grateful to the participating educators, quality-assurance officers and vocational training institutions for their contribution to the study.

REFERENCES

1. Baartman, L. K. J., Bastiaens, T. J., Kirschner, P. A., & van der Vleuten, C. P. M. (2007). Evaluating assessment quality in competence-based education: A qualitative comparison of two frameworks. *Educational Research Review*, 2(2), 114–129. <https://doi.org/10.1016/j.edurev.2007.06.001>
2. Baartman, L. K. J., & Gulikers, J. T. M. (2025). VET teacher professional development in formative assessment using learning progressions. *Journal of Formative Design in Learning*, 9(2), 136–148. <https://doi.org/10.1007/s41686-025-00107-4>
3. Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide*. SAGE.
4. Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, 30(1), 3–12. <https://doi.org/10.1111/j.1745-3992.2010.00195.x>
5. Garraway, J. (2022). Designing complex, challenging and creative assessments for work preparedness: A review of competency-based assessment. *Journal of Vocational, Adult and Continuing Education and Training*, 5(1), Article 20. <https://doi.org/10.14426/jovacet.v5i1.249>
6. Geng, J., & Mohd Yusoff, N. (2023). The shared features of effective improvement programmes for teachers' assessment literacy. *Pedagogies: An International Journal*, 18(4), 710–727. <https://doi.org/10.1080/1554480X.2022.2084398>
7. Gulikers, J. T. M., Bastiaens, T. J., & Kirschner, P. A. (2006). Authentic assessment, student and teacher perceptions: The practical value of the five-dimensional framework. *Journal of Vocational Education & Training*, 58(3), 337–357. <https://doi.org/10.1080/13636820600955443>
8. Leonardsen, J. K., Utvær, B. K. S., & Fjørtoft, H. (2022). The five faces of an assessor: Conceptualizing the enactment of teacher assessment identity in vocational education and training. *Educational Assessment*, 27(4), 339–355. <https://doi.org/10.1080/10627197.2022.2106967>
9. Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE.
10. Maja, M. M., & Atukunda, J. (2022). Computer practice module lecturers' experiences of internal continuous assessment at technical vocational education and training colleges. *International Journal of Educational Methodology*, 8(1), 151–162. <https://doi.org/10.12973/ijem.8.1.151>
11. Meijer, K., Baartman, L., Vermeulen, M., & de Bruijn, E. (2025). Mapping teachers' perceived importance of assessment competence: A quantitative exploration in the context of Dutch higher vocational education. *Empirical Research in Vocational Education and Training*, 17, Article 10. <https://doi.org/10.1186/s40461-025-00185-5>
12. Mutebi, R., & Ferej, A. (2023). A review of TVET quality assurance practice in Uganda. *East African Journal of Interdisciplinary Studies*, 6(1), 182–196. <https://doi.org/10.37284/eajis.6.1.1327>
13. Pastore, S., & Andrade, H. L. (2019). Teacher assessment literacy: A three-dimensional model. *Teaching and Teacher Education*, 84, 128–138. <https://doi.org/10.1016/j.tate.2019.05.003>
14. Pastore, S. (2023). Teacher assessment literacy: A systematic review. *Frontiers in Education*, 8, Article 1217167. <https://doi.org/10.3389/educ.2023.1217167>
15. Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory Into Practice*, 48(1), 4–11. <https://doi.org/10.1080/00405840802577536>
16. Republic of Namibia. (2023). Regulations relating to assessment and certification of Namibia Qualifications Framework awards within the vocational education and training system of Namibia. *Government Gazette of the Republic of Namibia*.
17. Republic of Namibia. (2008). *Vocational Education and Training Act, 2008 (Act No. 1 of 2008)*. Government of the Republic of Namibia.
18. Villarroel, V., Melipillán, R., Santana, J., & Aguirre, D. (2024). How authentic are assessments in vocational education? An analysis from Chilean teachers, students, and examinations. *Frontiers in Education*, 9, Article 1308688. <https://doi.org/10.3389/educ.2024.1308688>

19. Welsandt, N. C. J., Fortunati, F., Winther, E., & Abs, H. J. (2024). Constructing and validating authentic assessments: The case of a new technology-based assessment of economic literacy. *Empirical Research in Vocational Education and Training*, 16, Article 4. <https://doi.org/10.1186/s40461-024-00158-0>
20. Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualising teachers' assessment literacies in an era of curriculum and assessment reform. *The Australian Educational Researcher*, 40(2), 241–256. <https://doi.org/10.1007/s13384-013-0089-9>
21. Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149–162. <https://doi.org/10.1016/j.tate.2016.05.010>
22. Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE.
23. Yusop, S. R. M., Rasul, M. S., Mohamad Yasin, R., Hashim, H. U., & Jalaludin, N. A. (2022). An assessment approaches and learning outcomes in technical and vocational education: A systematic review using PRISMA. *Sustainability*, 14(9), Article 5225. <https://doi.org/10.3390/su14095225>