



Assessment Literacy

**A review of key themes in assessment
literacy within Higher Education:
What we know**

**Duke Kunshan University Institute of Global Higher
Education (IGHE) Pedagogy and Research in
Practice Lab**

Prepared by: Laura Davies, Assistant Director for English for Academic Purposes Senior
Lecturer of English Language Laura.davies@dukekunshan.edu.cn

Contents

Executive Summary.....	3
Glossary	4
Introduction	6
Key Themes.....	7
1. Assessment Validity in the Agen of Generative AI (GenAI).....	7
1.1 Overall Effect on Student Learning Outcomes.....	7
1.2 Mechanisms.....	7
1.3. Confidence in the Evidence Base.....	8
1.4. Actionable Recommendations for (DKU) Faculty	8
2. Inclusive Assessment.....	9
2.1 Overall Effect on Student Learning Outcomes.....	9
2.2. Mechanisms.....	9
2.3. Confidence in the Evidence Base.....	10
2.4. Actionable Recommendations for (DKU) Faculty	11
3. Authentic Assessment	12
3.1 Overall Effect on Student Learning Outcomes.....	12
3.2 Mechanisms.....	13
3.3 Confidence in the Evidence Base.....	13
3.4 Actionable Recommendations for DKU Faculty	14
Conclusion.....	14
References.....	16

Executive Summary

Assessment literacy (AL) is widely recognised as central to effective practice in higher education, yet its application remains complex and context-dependent. This report synthesizes three key areas of contemporary AL research: assessment validity in the age of Generative AI (GenAI), inclusive assessment, and authentic assessment. It examines their effects on student learning outcomes, underlying mechanisms, the strength of the evidence base, and implications for practice.

The analysis suggests that the impact of assessment is contingent on design and context. GenAI challenges traditional product-based assessment, while process-oriented and programmatic approaches appear more effective in supporting valid interpretations of learning. Inclusive assessment is associated with increased engagement, motivation, and sense of belonging, though effects on measurable outcomes remain mixed. Similarly, authentic assessment can support deeper engagement when aligned with meaningful contexts, but its effectiveness depends on implementation.

Across all areas, alignment between assessment design, pedagogy, and learning outcomes is critical. While there is converging support for these approaches, the evidence base remains context-specific and methodologically constrained. Developing AL therefore requires the capacity to make informed, context-sensitive decisions about assessment design and implementation.

Glossary

Assessment literacy (AL)

The knowledge, skills, and contextual understanding required to design, interpret, and use assessment effectively to support learning and informed decision-making.

Assessment for learning (AFL)

Assessment practices that support ongoing learning through feedback, reflection, and iterative improvement during the learning process.

Assessment of learning (AoL)

Summative assessment practices used to evaluate and certify student achievement, typically at the end of a learning period.

Assessment validity

The extent to which assessment design and interpretation accurately capture intended learning outcomes within specific contexts, supported by appropriate evidence.

Authentic assessment

Assessment tasks requiring meaningful application of knowledge and skills in relevant or real-world contexts.

Authenticity (in assessment)

The degree to which assessment aligns with learners' contexts, experiences, and goals, enhancing relevance and meaningful engagement with learning.

Feedback literacy

The ability of students and educators to understand, interpret, and use feedback effectively to improve learning and performance.

Formative assessment

Low-stakes assessment that provides feedback and opportunities for improvement prior to summative evaluation.

Generative artificial intelligence (GenAI)

AI systems capable of producing human-like content, challenging traditional assessment by enabling task completion without demonstrating genuine capability.

Inclusive assessment

Assessment practices designed to provide equitable opportunities for diverse learners to demonstrate knowledge and skills.

Intrinsic motivation

Motivation driven by internal interest and engagement, supported by autonomy, competence, and relatedness.

Programmatic assessment

An approach where multiple assessment points across a program collectively inform judgments about student learning and development.

Social justice in assessment

Assessment practices that view assessment as a social practice to promote recognition, fairness in outcomes (not just procedures), and the development of learners who can contribute to a more just society.

Summative assessment

High-stakes assessment used to measure achievement of learning outcomes at a defined point in time.

Transnational higher education

Educational provision across national contexts, often involving hybrid or 'glocal' influences on curriculum, pedagogy, and assessment.

Introduction

Assessment – once seen primarily as an act of objective measurement – is now widely understood as a matter of inclusion, equity and social justice (Nieminen, 2025)

It is widely upheld that Assessment Literacy (AL) ought to be expanded for teachers, administrators, and policymakers across the HE sector, yet its definition is not unanimous. According to Stiggins (1991, p.535), “Those who are assessment literate, on the one hand, have a basic understanding of the meaning of high- and low-quality assessment and are able to apply that knowledge to various measures of student achievement”. Practitioners therefore require specific knowledge and skills (literacies) to effectively assess and interpret results, particularly if assessment is to inform both policy and instructional decisions. Furthermore, AL is context dependent and interwoven with a practitioner’s content knowledge, pedagogical skill, and understanding of testing, measurement, and data use (DePascale et al., 2018). Adding to this complexity are multiple forms of assessment (Tight, 2021), including assessment of and for learning, and formative and summative approaches. A fundamental starting point for understanding AL is therefore the key literature, including Stiggins (1991; 2002), Popham (2004; 2009; 2011), and Xu and Brown’s (2016, p.155) Teacher assessment literacy in practice (TALiP) framework, which marks a transition towards a more sophisticated understanding of assessment for learning.

The TALiP framework calls for a situated, dynamic, and evolving system “that depends on forming a virtuous circle of TALiP with joint efforts from appropriate stakeholders” (Xu and Brown, 2016, p.159). It conceptualizes AL development across levels—from initial “should do” knowledge, through internalized understanding, to self-directed awareness of assessment processes and identity. This reconceptualization positions practitioners along a continuum, rather than a literate/illiterate dichotomy, emphasizing AL as both a practitioner priority and a wider HE concern.

An understanding of AL is crucial as it fundamentally shapes learning (Yang and Zhang, 2024). Contextual features stem from institutional culture and are influenced by geo-political forces, global education reform (Sahlberg, 2006), and neoliberal versus global reform beliefs (Verger and Altinyelken, 2013). In Sino-JVUs, these influences often combine in ‘glocal’ forms, requiring assessment to serve as a rationale for learning. Ultimately, as assessment design always requires compromise (Ajjawi et al., 2024), and given the power faculty hold as assessment designers, developing practitioner AL is essential if assessment is to meaningfully shape learning and align with evolving pedagogical, technological, and contextual demands. Finally, educational research is often context-dependent, with limitations relating to sample size, methodological design, and variability across institutional settings. Findings should therefore be interpreted as indicative rather than universally generalizable.

This report synthesizes key themes in assessment literacy—validity in the age of Generative AI, inclusive assessment, and authentic assessment—through an evidence-informed lens. For each theme, it examines: (1) effects on student learning outcomes, (2) underlying mechanisms, (3) the strength and limitations of the evidence base, and (4) actionable

implications for faculty practice. In doing so, this report aims to support higher education practitioners in developing assessment approaches that are pedagogically robust, contextually appropriate, and responsive to emerging challenges.

Key Themes

1. Assessment Validity in the Age of Generative AI (GenAI)

“Generative AI (GenAI) challenges assessment validity by enabling students to complete tasks without demonstrating genuine capability” (Corbin et al., 2025, p.1).

1.1 Overall Effect on Student Learning Outcomes

The rise of Generative AI (GenAI) has revolutionized education practices, particularly within higher education. For over a decade, technology has played a significant role in shaping assessment design. This trajectory reached a pivotal turning point with the launch of OpenAI’s ChatGPT 3.5 in 2022, which accelerated the democratization of AI accessibility within the higher education sector. As one of the most powerful, transformative technologies available (Radanliev, 2025), AI’s role in assessment (Marshall et al., 2024) and its influence on validity are therefore paramount to AI development and assessment evolution.

Taking the features of validity outlined in the literature (Messick, 1989; Sireci, 2007, 2009, cited by Shaw and Crisp, 2011) into account, GenAI represents a significant disruption to HE assessment validity. Its impact on student learning outcomes is highly context-dependent. Where assessment relies on product-based outputs, GenAI may obscure whether students have independently demonstrated knowledge and skills, resulting in negative or indeterminate outcomes. However, when assessments emphasize process, multiple modes, and authentic demonstration of learning, GenAI can support deeper engagement. Designing assessments to capture “the coherent demonstration of learning across multiple appropriately designed touchpoints” (Corbin et al., 2025, p.9) may therefore enhance validity. Overall, the impact of GenAI remains mixed and contingent on assessment design, with stronger outcomes linked to structurally redesigned, programmatic, and multi-modal approaches.

1.2 Mechanisms

Initial responses to GenAI have involved discursive changes to policy and instruction (Corbin et al., 2025), including traffic light systems, declarative approaches, and the AI Assessment Scale (AIAS) (Perkins et al., 2023; Roe et al., 2026; Furze, 2025). The AIAS not only promotes assessment transparency, it is also easily adapted and translated contextually and used around the world (Furze, 2024). For example, Roe *et al.* (2026) have adapted the AIAS for an English for Academic Purposes context. Furthermore, it is a powerful pedagogical tool to enable

practitioners to provide non-binary instruction for AI use in assessment, and to negotiate with students. The authors seek ongoing feedback to revise the framework, and provide a suite of resources, including a chatbot to engage with the model available through their open-access website: <https://aiassessmentsscale.com/>. While these provide guidance and scaffolding, they create an “enforcement illusion” (Corbin et al., 2025, p,1) and do not ensure validity. Instead, validity must be built through structural redesign of assessment tasks.

Redesigned assessment enhances learning through interconnected mechanisms. Pedagogically, designing assessments across multiple modalities and touchpoints (e.g., the ‘Swiss Cheese’ model; Dawson, 2024) reduces reliance on any single performance, ultimately reducing an assessment design’s vulnerability to AI invalidity. Cognitively, Self-Determination Theory (Ryan and Deci, 2000), highlights the role of autonomy, competence, and relatedness in fostering intrinsic motivation; flexible assessment formats and reflective components therefore promote deeper engagement. Furthermore, structurally, programmatic assessment (Baartman and Quinlan, 2024; Torre and Schuwirth, 2025) distributes and calibrates assessment across curricula, enabling more reliable judgments of capability. Finally, socially, trust-based approaches rely on shared norms, but without structural reinforcement may not safeguard validity. Together, these mechanisms demonstrate that validity is achieved through alignment between assessment design, pedagogy, and learner engagement.

1.3. Confidence in the Evidence Base

The evidence base regarding GenAI and assessment validity is rapidly emerging but not yet robust. Much of the literature (e.g., Corbin et al., 2025; Furze, 2024; Rudolph et al., 2023) is conceptual or theoretical, based on early-stage case studies, and focused on institutional responses rather than measured learning outcomes. There is limited longitudinal evidence examining the sustained impact of redesigned assessments, and little empirical work on enforcing secure environments, comparing open versus closed approaches, or understanding how students enact AI-related policies.

As is common within educational research, limitations include small and context-specific samples, difficulty in randomization, reliance on self-reported perceptions, variability across contexts, and limited replication. These constraints mean findings should be interpreted as indicative rather than definitive. Nevertheless, consistent emphasis on structural redesign and multi-modal assessment provides a moderate level of confidence as current recommendations.

1.4. Actionable Recommendations for (DKU) Faculty

Faculty should move beyond instruction-based control (e.g., AI declarations) and structurally redesign assessment tasks, incorporating multiple modes and combining open and secure approaches where appropriate. Pedagogically, process-oriented assessments such as drafts, reflections, and iterative tasks—should be embedded, alongside opportunities for choice, self-assessment, and reflection to foster autonomy and intrinsic motivation (Kearney, 2012). At the program level, assessments should be aligned across courses, ensuring learning outcomes are evaluated multiple times in varied formats.

Frameworks such as the AIAS should guide, rather than replace, design decisions, while dialogue with students regarding AI use, expectations, and ethics should be encouraged. Institutional support is also critical, including time for redesign, professional development in assessment literacy, and access to exemplars of innovative practice. These approaches are particularly relevant for the DKU context, including writing-intensive courses, take-home assessments, and the transnational, multilingual setting, where attention to linguistic and cultural diversity and scaffolding for both assessment completion and AI use is essential.

2. Inclusive Assessment

“Assessment that treats all students the same is inequitable, since it ignores differences in students’ past and present circumstances” (Tai et al., 2022, p.403).

2.1 Overall Effect on Student Learning Outcomes

The inclusive assessment literature has moved beyond technical accommodations, which have been shown to stigmatize disabled students by framing them as “unfit” and “abnormal” towards a more socially constructed understanding of accessible assessment design (Nieminen, 2023, p.65). Inclusive assessment is shaped significantly by practitioners’ conceptions of inclusivity and equity (DeLuca et al., 2019), with important implications for assessment literacy. The evidence suggests that inclusive assessment practices generally have a positive effect on student learning outcomes, particularly in relation to increased engagement, improved motivation, and an enhanced sense of belonging. Approaches such as co-design, self-assessment, and diversified assessment formats enable students to align tasks with their interests and strengths, thereby fostering deeper engagement. For example, co-designing assignments creates choice, allowing learners to develop a sense of ownership and autonomy, which in turn promotes intrinsic motivation (Ryan and Deci, 2000).

A balanced assessment diet that incorporates both formative and summative elements can further support learning by providing opportunities for growth and reducing anxiety associated with high-stakes assessment. However, the impact on measurable learning outcomes remains mixed and context-dependent, particularly where debates persist regarding the role of formative assessment in grading and whether it captures learning or the learning process, and whether summative assessment is inclusive (Kleinlein, 2026). Overall, inclusive assessment practices appear most effective when they reduce structural barriers, provide equitable opportunities to demonstrate learning, and support intrinsic motivation, although outcomes vary depending on implementation and practitioner assessment literacy.

2.2. Mechanisms

Inclusive assessment operates through interconnected pedagogical, motivational, social, and structural mechanisms. Pedagogically, approaches such as self-assessment and co-design

involve students in identifying standards and evaluating their own work (Boud, 2000), thereby promoting metacognitive development and deeper engagement. Formative assessment provides scaffolding that supports knowledge development prior to summative evaluation, enabling students to refine their understanding. Motivationally, Self-Determination Theory (Ryan and Deci, 2000), offers a useful explanatory framework. Inclusive practices foster autonomy through choice and co-design, competence through scaffolded learning and feedback, and relatedness through collaborative and participatory assessment practices. These conditions enhance intrinsic motivation, which in turn supports deeper learning and meaningful engagement.

Moreover, socially, assessment functions as a form of power (Raaper, 2016), and inclusive approaches redistribute aspects of this power by involving students more actively in assessment processes. Co-design and collaborative assessment can therefore foster a sense of belonging and shared responsibility within learning communities, positively influencing engagement and participation. Finally, structurally, institutional policy, quality assurance systems, and cultural contexts shape how inclusivity is enacted in practice (Ajjawi et al., 2023).

Practitioners interpret these structures with varying degrees of flexibility, resulting in diverse approaches. Specifically, in transnational contexts such as Sino-JVUs, hybrid 'glocal' influences further shape how inclusivity is conceptualized and implemented. Socially just assessment (McArthur, 2016) integrates these dimensions by emphasizing sustained engagement with knowledge, recognition of individual student circumstances, reconsideration of grading practices, and alignment with broader social and educational purposes. Ultimately, Nieminen (2023) argues that the success of inclusive assessment practices depends largely on the level of teacher agency, expertise (AL), and support to adjust assessments. Inclusive assessment also needs to be inclusive for teachers and cannot be easily achieved when bound by rules, regulations, and rubrics.

2.3. Confidence in the Evidence Base

The evidence base for inclusive assessment is substantial but limited in scope. A growing body of research supports practices such as co-design, self-assessment, and formative assessment (Smith et al., 2024; Tai et al., 2022; McArthur, 2021). However, several limitations characteristic of educational research should be acknowledged. Firstly, much of the literature relies on self-reported student perceptions rather than direct measures of learning outcomes, and findings are often context-specific, influenced by discipline, institutional culture, and cohort characteristics. Additionally, sample sizes are frequently small or localized, limiting generalizability, and there is limited longitudinal research examining sustained impacts on learning. Furthermore, co-design studies often lack methodological transparency, making replication difficult (Smith et al., 2024).

Moreover, certain student populations, particularly international students and graduate learners, remain underrepresented in the literature, despite their increasing relevance within transnational higher education contexts. Furthermore, research on student self-assessment suggests that effects on measurable academic outcomes are variable and highly contingent on implementation, indicating that claims regarding direct impacts on learning achievement should

be interpreted with caution (e.g., Yan et al., 2023). Nevertheless, a growing body of research demonstrates that inclusive assessment practices can enhance student engagement, motivation, and perceptions of fairness, particularly through fostering agency and broadening the ways in which learners can demonstrate knowledge (e.g., Nieminen, 2024). There is, therefore, a moderate to high level of confidence in their pedagogical value.

Finally, Nieminen (2025) warns of the paradox of inclusive assessment. This paradox positions inclusive assessment as being in opposition to many common higher education norms and practices. For example, assessment standardization vs. student diversity, categories and hierarchies vs. fluidity and equality, and individualism vs. interdependence may suggest despite promising results, the problem of inclusive assessment may not be solvable, and remains an ongoing topic of negotiation and critical reflection.

2.4. Actionable Recommendations for (DKU) Faculty

Inclusive assessment requires intentional design, pedagogical alignment, faculty agency, expertise and institutional support (Nieminen 2023). At the level of assessment design, faculty should ensure that tasks allow multiple ways for students to demonstrate learning, for example through offering choice in assessment format or questions or developing balanced assessment diets that integrate formative and summative elements. Student involvement is also significant; integrating co-design practices, particularly in rubric development and assessment criteria discussions, alongside encouraging self- and peer-assessment, can strengthen student assessment literacy (Tai *et al.*, 2018). From a pedagogical perspective, clear scaffolding is essential, particularly for complex or unfamiliar tasks, while formative assessment should be used strategically to reduce anxiety and support progression. Transparency in expectations and criteria is also critical. Socially, fostering learning communities through collaborative assessment and maintaining awareness of power dynamics can help create inclusive, dialogic environments (Carless, 2012).

Inclusive assessment requires intentional design, pedagogical alignment, faculty agency, expertise and institutional support (Nieminen, 2023). At the level of assessment design, faculty should ensure that tasks allow multiple ways for students to demonstrate learning. For example, through offering choice in assessment format or questions, or developing balanced assessment diets that integrate formative and summative elements. Student involvement is also significant; integrating co-design practices, particularly in rubric development and assessment criteria discussions, alongside encouraging self- and peer-assessment, can strengthen student AL. From a pedagogical perspective, clear scaffolding is essential, particularly for complex or unfamiliar tasks, while formative assessment should be used strategically to reduce anxiety and support progression. Transparency in expectations and criteria is also critical. Socially, fostering learning communities through collaborative assessment and maintaining awareness of power dynamics can help create inclusive, dialogic environments.

In the DKU context, these practices are particularly relevant for multilingual learners, and transnational student populations. Faculty should therefore consider linguistic and cultural diversity in assessment design, provide additional support where needed, and align practices with institutional values of inclusivity and equity. Finally, institutions play a key role in

supporting faculty through professional development in inclusive assessment design, sharing of best practices, and opportunities for collaborative curriculum development. To summarize, Ajjawi et al. (2023, p.232) when synthesizing major converging themes from a recent edited collection of inclusive assessment in higher education propose four major recommendations: 1) students must be active agents in assessment design; 2) regulatory frameworks governing assessment must prioritize inclusion; 3) faculty must adopt ethical reflexivity; and 4) positivist and ableist discourses of assessment need to be disrupted.

3. Authentic Assessment

Assessment should be about “the outcome of engagement with knowledge that lasts” (McArthur 2023, p.21).

3.1 Overall Effect on Student Learning Outcomes

Authentic assessment has been widely discussed in the literature for over 30 years (Ajjawi et al., 2023), although its conceptualization continues to evolve. Traditionally, assessment has been framed as either assessment of learning or assessment for learning. The former reflects summative approaches focused on evaluation, while the latter emphasizes the role of assessment in supporting ongoing learning and development (Sambell et al., 2013). Assessment for learning (AfL) reframes assessment as a process that promotes engagement, recognizes emotional and affective dimensions of learning, and supports learners in navigating risk through development rather than control (Boud and Falchikov, 2007).

Within this framing, authentic assessment is associated with positive effects on student learning outcomes, particularly in relation to engagement, application of knowledge, and the development of transferable skills. Learning communities and collaborative practices foster increased engagement, defined as the time, effort, and interest students invest in learning (Trowler and Trowler, 2010 cited in Sambell et al., 2013). Furthermore, authentic assessment supports intrinsic motivation through the development of competence, autonomy, and relatedness (Sambell et al., 2013), encouraging students to engage more meaningfully with knowledge. Such insights highlight a recurring theme in this report, pointing towards the strong correlation between assessment literacy, expertise in assessment design and learner motivation. Authentic assessment has also been proposed as a mechanism for enhancing inclusive assessment (see section 2) (Nieminen, 2023).

However, the impact of authentic assessment remains contingent on its design and implementation. Early conceptualizations focused on replicating real-world or workplace tasks (Ajjawi et al., 2024; McArthur, 2023), yet these approaches can be limiting. More recent perspectives emphasize authenticity in assessment, where alignment with learners’ lived experiences, values, and future goals enhances relevance and engagement (Ajjawi et al., 2024).

Moreover, Quinlan et al. (2025) surveyed students about their most engaging assessments and found assessment choice and real-world connection promoted meaningful engagement. They argue for a move away from professional, societal, disciplinary, and developmental authentic and toward designing for educational authenticity. As such, authentic assessment can be operationalized in different ways, but appears most effective when it moves beyond replication towards meaningful, context-sensitive engagement with knowledge.

3.2 Mechanisms

Authentic assessment operates through interconnected pedagogical, social, and motivational mechanisms. Pedagogically, it emphasizes application-based learning, where students engage with complex, often ambiguous tasks that require the integration and transfer of knowledge. This aligns with calls to move beyond individual attainment towards understanding what learners are able to do upon completion of assessment (Boud, 2009 p.42). Collaborative tasks, such as group research projects or presentations, position others as integral to the learning process and contribute to authentic learning experiences.

Motivationally, authentic assessment supports intrinsic motivation by enabling autonomy through choice and personalization, and competence through engagement with meaningful and challenging tasks. This aligns with earlier discussion of Self Determination Theory (Ryan and Deci, 2000), where autonomy, competence, and relatedness underpin sustained engagement. By connecting learning to meaningful contexts, authentic assessment encourages deeper engagement and ownership of learning.

Socially, authenticity can be understood as a complex social practice (Ajjawei et al., 2024). Dialogue between practitioners and learners regarding the purpose and meaning of assessment tasks can enhance understanding of real-world complexities. Co-design and collaborative approaches further strengthen authenticity by enabling shared ownership of assessment processes, while also supporting reflexive pedagogical practice (Ajjawei et al., 2024 p.506).

At a structural level, authentic assessment requires broader curricular and pedagogical transformation. This includes creating space for problem-based learning, projects, and community-based tasks (Barnett, 2007), as well as recognizing the influence of prior assessment experiences on student engagement. Assessment literate practitioners must therefore design tasks with appropriate scaffolding and support to enable students to engage effectively with authentic forms of assessment.

3.3 Confidence in the Evidence Base

The evidence base for authentic assessment is substantial but remains conceptually driven and uneven in empirical depth. Much of the literature (Ajjawi et al., 2024; McArthur, 2023) focuses on theoretical development and conceptual reframing, particularly in relation to the shift from authentic assessment towards authenticity in assessment (Ajjawi et al., 2024). While collections such as Sambell and Brown (2021) provide practical examples across disciplines, there remains limited large-scale or longitudinal evidence examining sustained impacts on learning outcomes.

With broader educational research, findings are often context-specific and influenced by disciplinary and institutional variation. There is also a need for greater clarity in the operationalization of authenticity, as differing definitions limit comparability across studies (Ajjawi et al., 2024). Nevertheless, there is consistent support for the role of authentic assessment in enhancing engagement, motivation, and application of knowledge. As such, there is a moderate level of confidence in its pedagogical value, although claims regarding direct impact on measurable outcomes should be interpreted with caution.

3.4 Actionable Recommendations for DKU Faculty

Authentic assessment requires intentional alignment between assessment design, pedagogy, and student context. Faculty should prioritise the development of tasks that require meaningful application of knowledge, rather than reproduction, and incorporate opportunities for reflection and iterative engagement. Providing flexibility in how students demonstrate learning can enhance relevance and support diverse learner needs enhancing both the authenticity and inclusivity of assessment (Nieminen, 2023).

Pedagogically, authentic assessment should be supported through appropriate scaffolding, particularly for students transitioning from more traditional, exam-oriented systems such as the Chinese college entrance examination (Gaokao) (Barnett, 2007). Embedding collaborative and community-based tasks can foster engagement and support the development of learning communities, while co-design approaches can enhance both authenticity and student ownership of assessment.

At a program level, curricula should create space for problem-based and interdisciplinary learning, enabling students to engage with complex, real-world challenges. In transnational contexts such as DKU, attention to cultural and educational diversity is essential, alongside alignment with broader institutional goals. Finally, institutions should support faculty through professional development, opportunities for innovation, and access to exemplars of authentic assessment practice.

Conclusion

This report has examined assessment literacy through three key areas: assessment validity in the age of Generative AI, inclusive assessment, and authentic assessment. Across these areas, a consistent pattern emerges. The effectiveness of assessment is not determined by individual tools, policies, or formats, but by how assessment is designed, implemented, and aligned with learners and the intended learning outcomes in mind.

In relation to validity, Generative AI challenges traditional product-based assessment approaches, particularly where student work can be completed without demonstrating capability. The evidence suggests that validity is better supported through process-oriented,

multi-modal, and programmatic approaches that capture learning across multiple touchpoints. However, this remains an emerging and rapidly changing area, and consequently, assessment practices should be regularly evaluated. Inclusive assessment is fundamental if assessment is to be used to support social justice. It demonstrates clear benefits in terms of engagement, motivation, and creating a sense of belonging, particularly where students are involved in co-design and supported through formative assessment processes. At the same time, its impact on measurable outcomes remains mixed and dependent on context, implementation, and practitioner expertise. Similarly, authentic assessment supports deeper engagement and application of knowledge when tasks are meaningful and aligned with student contexts, although its effectiveness is also contingent on design and scaffolding.

Across all three areas, the importance of alignment is clear. Assessment is most effective when design, pedagogy, and learning outcomes are coherent, and when students are supported to engage meaningfully with tasks. Mechanisms such as intrinsic motivation, participation, and the use of multiple assessment points play an important role in shaping outcomes, but these require deliberate design rather than reliance on policy or guidance alone.

The report also highlights the limitations of the current evidence base. Much research is conceptual in nature and findings are context-specific, rely on self-reported data, and lack longitudinal depth. As such, while there is consistent support for many of the approaches discussed, their application requires careful consideration of local contexts and constraints.

Overall, developing assessment literacy requires practitioners to make informed, context-sensitive decisions about assessment design and implementation. This includes adapting practices in response to Generative AI, embedding inclusive approaches, and designing assessments that support meaningful engagement with knowledge. For institutions, this requires ongoing support for faculty, including time, resources, and professional development to enable effective and sustainable change.

References

- Ajjawi, R., Tai, J., Boud, D., & Jorre de St Jorre, T. (2023). Moving forward: Mainstreaming assessment for inclusion in curricula. In R. Ajjawi, J. Tai, D. Boud, & T. Jorre de St Jorre (Eds.), *Assessment for inclusion in higher education: Promoting equity and social justice in assessment*. Routledge.
- Ajjawi, R., Tai, J., Dollinger, M., Dawson, P., Boud, D., & Bearman, M. (2024). From authentic assessment to authenticity in assessment: Broadening perspectives. *Assessment & Evaluation in Higher Education*, 49(4), 499–510.
<https://doi.org/10.1080/02602938.2023.2271193>
- Baartman, L. K. J., & Quinlan, K. M. (2024). Assessment and feedback in higher education reimaged: Using programmatic assessment to transform higher education. *Perspectives: Policy and Practice in Higher Education*, 28(2), 57–67.
<https://doi.org/10.1080/13603108.2023.2283118>
- Barnett, R. (2007). Assessment in higher education: An impossible mission? In D. Boud & N. Falchikov (Eds.), *Rethinking assessment in higher education: Learning for the longer term* (pp. 29–40). Routledge.
- Boud, D. 2000. “Sustainable Assessment: Rethinking Assessment for the Learning Society.” *Studies in Continuing Education* 22 (2): 151–167. doi: 10.1080/713695728
- Boud, D., & Soler, R. (2015). Sustainable assessment revisited. *Assessment & Evaluation in Higher Education*, 41(3), 400–413. <https://doi.org/10.1080/02602938.2015.1018133>
- Boud, D., & Falchikov, N. (Eds.). (2007). *Rethinking Assessment in Higher Education: Learning for the Longer Term* (1st ed.). Routledge. <https://doi.org/10.4324/9780203964309>
- Carless, D. (2012). Trust and its role in facilitating dialogic feedback. In *Feedback in higher and professional education* (pp. 90–103). Routledge.
- Corbin, T., Dawson, P., Nicola-Richmond, K., & Partridge, H. (2025). “Where’s the line? It’s an absurd line”: Towards a framework for acceptable uses of AI in assessment. *Assessment & Evaluation in Higher Education*. <https://doi.org/10.1080/02602938.2025.2456207>
- Dawson, P. (2024). AI in education podcast: Assessment and Swiss cheese – Phill Dawson (Episode 9, Series 9). *Wonkhe podcast*.
- DeLuca, C., Coombs, A., Macgregor, S., and Rasooli, A. (2019). Toward a Differential and Situated View of Assessment Literacy: Studying Teachers' Responses to Classroom Assessment Scenarios. *Frontiers in Education*. 4. 10.3389/feduc.2019.00094.

- DePascale, C., Sharp, A., Ryan, K., & Betebenner, D. (2018, May). *Assessment literacy framework: Building a conceptual framework for assessment literacy*. PDF report.
- Kearney, S. (2012). Improving engagement: The use of authentic self- and peer-assessment for learning. *Assessment & Evaluation in Higher Education*, 38(7), 875–891.
<https://doi.org/10.1080/02602938.2012.751963>
- Kleinlein, E. (2026). Conceptualising inclusive assessment. A critical literature review at the nexus of education, inclusion, and assessment. *European Journal of Special Needs Education*, 41(1), 101–120. <https://doi.org/10.1080/08856257.2025.2500137>
- Marshall, S., Blaj-Ward, L., Dreamson, N., Nyanjom, J., & Bertuol, M. T. (2024). The reshaping of higher education: technological impacts, pedagogical change, and future projections. *Higher Education Research & Development*, 43(3), 521–541.
<https://doi.org/10.1080/07294360.2024.2329393>
- McArthur, J. (2016). Assessment for social justice: the role of assessment in achieving social justice. *Assessment and Evaluation in Higher Education*, 41(7), 967–981.
<https://doi.org/10.1080/02602938.2015.1053429>
- McArthur, J. (2021, March 3). For assessment to count as authentic it must mean something to students. *Wonkhe*.
- McArthur, J. (2023). Rethinking authentic assessment: Work, well-being, and society. *Higher Education*, 85, 85–101. <https://doi.org/10.1007/s10734-022-00822-y>
- Messick, S. (1989). Meaning and values in test validation: the science and ethics of assessment. *Educational Researcher*, 18.
- Nieminen, J. H. (2023). Inclusive assessment, exclusive academy. In R. Ajjawi, J. Tai, D. Boud, & T. Jorre de St Jorre (Eds.), *Assessment for inclusion in higher education: Promoting equity and social justice in assessment*. Routledge.
- Nieminen, J. H. (2024). Assessment for inclusion: Rethinking inclusive assessment in higher education. *Teaching in Higher Education*, 29(4), 841–859.
<https://doi.org/10.1080/13562517.2021.2021395>
- Nieminen, J. H. (2025). The paradox of inclusive assessment. *Assessment & Evaluation in Higher Education*, 50(4), 564–576. <https://doi.org/10.1080/02602938.2024.2419604>
- Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2025). AI assessment scale (AIAS).
<https://aiassessmentscale.com/>

- Popham, W. J. (2004). Why assessment illiteracy is professional suicide. *Educational Leadership*, 62(1), 82.
- Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory Into Practice*, 48(1), 4–11.
- Popham, W. J. (2011). Assessment literacy overlooked: A teacher educator's confession. *The Teacher Educator*, 46(4), 265–273.
- Quinlan, K. M., Sellei, G., & Fiorucci, W. (2025). Educationally authentic assessment: Reframing authentic assessment in relation to students' meaningful engagement. *Teaching in Higher Education*, 30(3), 717–734. <https://doi.org/10.1080/13562517.2024.2394042>
- Raaper, R. (2016). Tracing power and student subjectivity in assessment: A Foucauldian perspective. *British Educational Research Association Conference*.
- Radanliev, P. (2025). Artificial intelligence: reflecting on the past and looking towards the next paradigm shift. *Journal of Experimental & Theoretical Artificial Intelligence*, 37(7), 1045–1062. <https://doi.org/10.1080/0952813X.2024.2323042>
- Roe, J., Perkins, M., & Tregubova, Y. (2026). The EAP-AIAS: Adapting the AI assessment scale for English for academic purposes. *TESOL Journal*, 17(2), e70122. <https://doi.org/10.1002/tesj.70122>
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessment in higher education? *Journal of Applied Learning and Teaching*, 6(1), 342–363.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Sahlberg, P. (2006). Education reform for raising economic competitiveness. *Journal of Educational Change*, 7(4), 259–287.
- Sambell, K., & Brown, S. (2021). Covid-19 assessment collection. <https://sally-brown.net/kay-sambell-and-sally-brown-covid-19-assessment-collection/>
- Sambell, K., McDowell, L., & Montgomery, C. (2013). *Assessment for learning in higher education*. Routledge.
- Shaw, S., & Crisp, V. (2011). Tracing the evolution of validity in educational measurement. *Cambridge Assessment Research Matters*.

- Smith, A., McConnell, L., Iyer, P., Allman-Farinelli, M., & Chen, J. (2024). Co-designing assessment tasks with students in tertiary education: A scoping review. *Assessment & Evaluation in Higher Education*, 50(2), 199–218. <https://doi.org/10.1080/02602938.2024.2376648>
- Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappan*, 72(7), 534–539.
- Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, 83(10), 758.
- Tai, J. H. M., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education*, 76, 467–481. <https://doi.org/10.1007/s10734-017-0220-3>
- Tai, J. H. M., Dollinger, M., Ajjawi, R., Jorre de St Jorre, T., Krattli, S., McCarthy, D., & Prezioso, D. (2022). Designing assessment for inclusion: An exploration of diverse students' assessment experiences. *Assessment & Evaluation in Higher Education*, 48(3), 403–417. <https://doi.org/10.1080/02602938.2022.2082373>
- Tight, M. (2021). *Synthesis of higher education research: What we know*. Bloomsbury.
- Torre, D., & Schuwirth, L. (2025). Programmatic assessment for learning: AMEE Guide No. 174. *Medical Teacher*, 47(6), 918–933. <https://doi.org/10.1080/0142159X.2024.2409936>
- Trowler, V., & Trowler, P. (2010). Student Engagement Evidence Summary. Higher Education Academy.
- Verger, A., & Altinyelken, H. K. (2013). Global education reforms and the new management of teachers: A critical introduction. In A. Verger, H. Altinyelken, & M. de Koning (Eds.), *Global managerial education reforms and teachers: emerging policies, controversies and issues in developing contexts* (pp. 1-18).
- Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149–162.
- Yan, Z., Wang, X., Boud, D., & Lao, H. (2023). The effect of self-assessment on academic performance and the role of explicitness: A meta-analysis. *Assessment & Evaluation in Higher Education*, 48(1), 1–15. <https://doi.org/10.1080/02602938.2021.2012644>
- Yang, M., Yan, Z., Yang, L., & Zhan, Y. (2025). Insights into student assessment literacy: A scoping analysis of SSCI publications. In *Understanding and developing student assessment literacy* (Springer Briefs in Education). Springer. https://doi.org/10.1007/978-981-97-9484-3_2