

Miller, W. (2026). Capacity and Disposition: A Framework for Understanding GenAI Engagement in Assessment Practice. *Intersection: A Journal at the Intersection of Assessment and Learning*, 7(5), 115–134. <https://doi.org/10.61669/001c.172677>

Capacity and Disposition: A Framework for Understanding GenAI Engagement in Assessment Practice

Will Miller, Ph.D.

Author Note

Will Miller, <https://orcid.org/0009-0005-7957-326X>

I have no conflicts of interest to disclose

Intersection: A Journal at the Intersection of Assessment and Learning

Vol 7, Issue 5 Special Issue: The Future of Assessment in an Era of Generative AI

Abstract: Generative artificial intelligence (GenAI) has made assessment work faster, messier, and more consequential. This conceptual article argues that assessment professionals' engagement with GenAI is shaped less by simple enthusiasm or resistance than by the relationship between two conditions: the time professionals have to learn, test, and implement GenAI responsibly, and the disposition they bring to GenAI-mediated assessment practice. Institutional support functions as the environment that expands or constrains both. The article proposes the Capacity-Disposition Framework, a model that identifies four engagement conditions: The Calendar Trap, The Margin Lab, The Evidence Gate, and The Practice Commons. The framework helps assessment leaders diagnose uneven GenAI adoption, protect time for meaningful experimentation, honor appropriate skepticism, and connect GenAI use to validity, equity, transparency, and continuous improvement rather than speed alone.

Keywords: *assessment practice, generative artificial intelligence, assessment professionals, institutional support, validity, continuous improvement*

Introduction

Generative artificial intelligence (GenAI) did not arrive in higher education at a calm moment. It arrived while assessment professionals were already balancing accreditation reporting, program review, general education assessment, survey analysis, faculty consultation, institutional effectiveness, and the daily work of trying to keep assessment connected to actual learning. Assessment has long lived inside a tension between accountability and improvement, between the need to produce evidence for external audiences and the need to use evidence to strengthen teaching, curriculum, and student success (Banta & Palomba, 2015; Ewell, 2009; Kuh et al., 2015; Suskie, 2018). GenAI intensifies that tension because it can produce the artifacts of assessment quickly while leaving unresolved the harder questions of validity, reliability, fairness, interpretation, and use.

Artificial intelligence (AI) is the broader category of systems used for tasks such as prediction, classification, and content generation. GenAI is the subset that generates text, images, code, or other content from learned patterns (National Institute of Standards and Technology, 2024). This framework addresses GenAI. References to AI literacy and broader human-AI research retain those sources' wider terminology.

The earliest campus conversations about GenAI often focused on students: whether students were using GenAI to complete assignments, whether faculty could detect that use, and whether academic integrity policies were adequate. That concern is understandable. In a survey of UK full-time undergraduates, the share reporting use of GenAI to help prepare assessed work rose from 53% in 2024 to 88% in 2025, while 92% reported using AI tools in some form (Freeman, 2025). Faculty are also grappling with the consequences. A 2026 Elon University and American Association of Colleges and Universities (AAC&U) survey of 1,057 faculty respondents reported deep concern about overreliance, critical thinking, academic integrity, and institutional preparation, while also showing that many respondents were already addressing AI literacy in their teaching (Watson & Rainie, 2026). The faculty sample was nonprobability-based; these findings describe the respondents rather than all faculty. Together, the surveys show how GenAI has entered the environment in which evidence of learning is produced, interpreted, and used.

Assessment professionals are therefore being pulled into a new role. Miller (2024) argued that assessment professionals should not simply react to GenAI as a threat to academic integrity; they should help institutions use GenAI in ways that preserve educational purpose, equity, and human judgment. Hathcoat et al. (2026) describe an even larger disruption, arguing that GenAI unsettles the boundaries among learner, test, and assessor and requires assessment professionals to become curators, auditors, translators, and stewards of GenAI-mediated evidence ecosystems. Those arguments are persuasive, but they leave a practical question in the middle of the room: under what conditions can assessment professionals actually do this work?

Current survey evidence suggests that GenAI use in assessment is widespread alongside gaps in institutional support. The Spring 2026 GenAI Pulse Survey for Higher Education Assessment Professionals used a convenience sample of 278 assessment leaders and assessment-team members. Among respondents to the relevant items, 79% were regular or occasional users, 87% of users reported efficiency gains, and 87% described themselves as self-taught. At the same time, 71% reported privacy and ethical concerns, 43% reported lacking policy, and 37% reported lack of training. Influence was also uneven: 46% reported substantial influence on GenAI integration within their immediate team, compared with only 11% at the institutional level (Slotnick et al., 2026). Therefore, professionals already using GenAI while reporting uneven policy, access, training, and authority.

This article argues that uneven GenAI engagement among assessment professionals is best understood through the interaction of capacity and disposition. Capacity is defined primarily as time: protected time to explore, learn, test, validate, document, and implement GenAI-assisted assessment workflows. Disposition refers to the professional stance a person brings to GenAI-mediated work: the beliefs, motivations, affective responses, and values through which the person interprets GenAI as threat, shortcut, tool, partner, or catalyst. Institutional support functions as the enabling or constraining environment around both dimensions.

The purpose of this article is to propose the Capacity-Disposition Framework for GenAI engagement in assessment practice. The framework is conceptual and non-empirical, but it is designed to be useful: a language for assessment leaders, faculty partners, and institutional decision-makers who need to

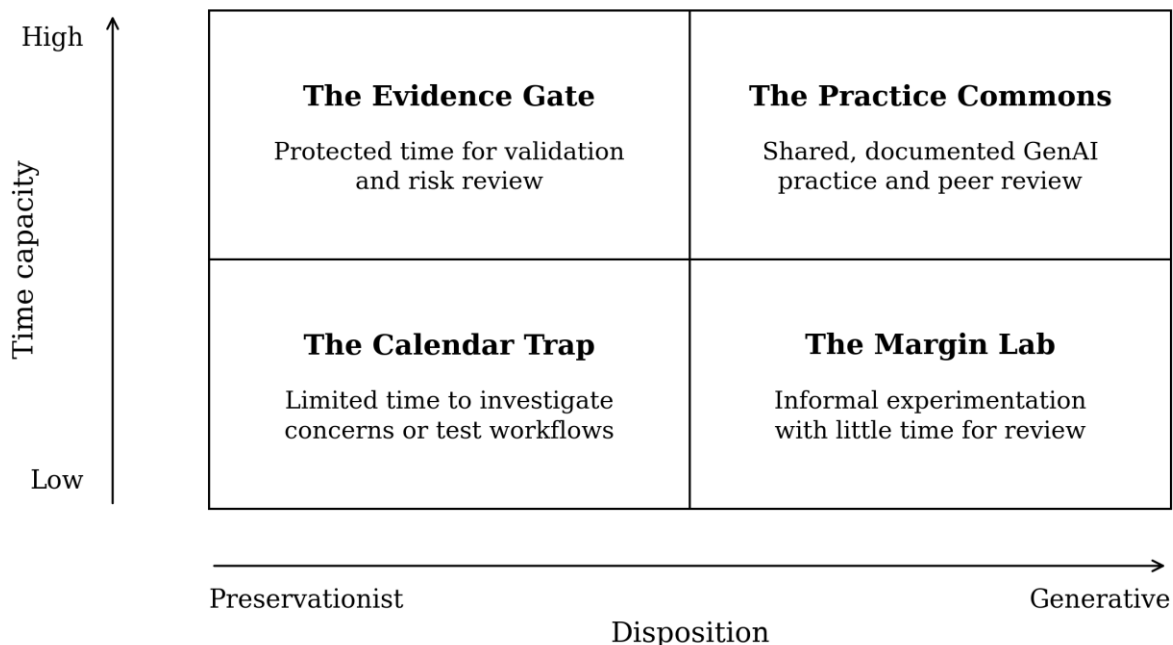
diagnose why GenAI engagement is uneven and what forms of support are most likely to move practice forward responsibly. The next section presents the framework first so the reader has a map. The sections that follow explain why the framework is needed now, how it relates to existing technology adoption and human-AI collaboration theories, how capacity and disposition are defined, and how institutions can use the framework in assessment work.

A Map of the Argument

Figure 1 presents the core framework. The two axes are intentionally simple. Time capacity ranges from low to high. Disposition ranges from preservationist to generative. The labels are not personality types, age categories, or fixed identities. They describe work conditions. A professional may occupy different conditions across a year, across projects, or even across tasks. During accreditation season, a willing experimenter may be pushed into The Margin Lab. After a serious data privacy concern, an enthusiast may move temporarily toward The Evidence Gate. With protected time and a trusted community of practice, someone who began in The Calendar Trap may move toward The Practice Commons.

Figure 1

Capacity-Disposition Framework for GenAI Engagement in Assessment Practice



Note. The framework describes conditions rather than fixed personality types. Professionals may move across conditions as workload, support, policy clarity, access, and experience change.

Why This Framework, and Why Now?

The framework responds to a gap in the current GenAI adoption conversation. Higher education has no shortage of GenAI guidance, tool demonstrations, policy templates, or lists of promising uses. What is

less developed is a practitioner-centered account of why assessment professionals engage GenAI unevenly when many of them face the same general technology shock. Existing technology adoption theories are useful but incomplete for this problem.

The technology acceptance model suggests that perceived usefulness and perceived ease of use shape whether people accept a technology (Davis, 1989). The unified theory of acceptance and use of technology expands this logic by emphasizing performance expectancy, effort expectancy, social influence, and facilitating conditions (Venkatesh et al., 2003). Diffusion of innovations theory helps explain how innovations spread through social systems and why adoption is shaped by perceived advantage, compatibility, complexity, trialability, and observability (Rogers, 2003). These theories help explain whether people see value in GenAI, whether the tool feels usable, whether peers and leaders encourage its use, and whether the institutional environment makes use possible.

But assessment work adds a specific complication: using GenAI is not simply adopting a productivity tool. It can change how evidence is generated, summarized, scored, interpreted, and reported. Assessment professionals are not only users of GenAI; they are guardians of evidence claims. They are asked to support faculty, protect student data, interpret learning evidence, manage accreditation narratives, and help institutions avoid mistaking documentation for improvement (Ewell, 2009; Hutchings, 2010; Kuh et al., 2015). A generic adoption model can explain willingness to try GenAI, but it does not fully explain the professional burden of deciding whether a GenAI-assisted output is valid, fair, transparent, and appropriate for use.

Human-AI collaboration frameworks also help, but they enter the story at a different point. Randazzo et al. (2025), building on work associated with Mollick and Mollick (2024), identify modes of GenAI knowledge work such as fused, directed, and abdicated co-creation. That literature is valuable because it shifts attention from whether people use GenAI to how they use it and what forms of expertise are developed or eroded through different workflows. The Capacity-Disposition Framework is complementary. It asks a prior question: What conditions allow assessment professionals to move toward responsible human-in-the-loop work at all? An assessment professional cannot become a disciplined GenAI collaborator if there is no time to learn the tool, no approved environment for student data, no policy clarity, and no institutional recognition for the labor of validation.

The framework also responds to professional learning literature. Effective technology learning rarely comes from one exposure. Sustained, job-embedded, collaborative professional development is more likely to change practice than isolated training (Desimone, 2009; Guskey, 2002). For assessment professionals, this is especially important because GenAI learning involves not only tool familiarity but assessment literacy: the ability to understand evidence, measurement quality, interpretation, fairness, and use (Popham, 2009; Suskie, 2018). GenAI readiness is therefore not achieved by attending a webinar. It requires time to practice, compare outputs, ask methodological questions, and integrate the tool into local assessment routines.

The two dimensions in the framework were selected through a conceptual synthesis of four bodies of literature: assessment and continuous improvement, technology adoption, professional learning, and

human-AI collaboration. Other dimensions could have been selected, including policy clarity, tool access, AI literacy, data governance, role authority, leadership support, risk tolerance, prior technical skill, and local assessment culture. Those factors matter, but they operate as supports, constraints, or moderators of the two central conditions. They shape whether professionals have usable capacity and how their disposition develops. The framework is therefore intentionally parsimonious. It does not explain everything, but it names two conditions that repeatedly surface in practice: people need time, and they need a professional reason to engage.

Assessment Practice at the Edge of GenAI Disruption

Assessment in higher education has never been merely a technical exercise. It includes measures, rubrics, sampling plans, scoring procedures, reports, dashboards, and improvement actions, but it is also a set of conversations about what matters, what counts as evidence, and what should happen because of what was learned. The late American Association for Higher Education (AAHE) principles of good practice framed assessment as most useful when it begins with educational values, involves the people closest to learning, and leads to improvement (AAHE, 1992). Good assessment's complexity becomes more important, not less, when GenAI can make evidence-like products appear polished and complete.

The disruption can be considered through the assessment cycle described in Horst and Prendergast's (2020) Assessment Skills Framework. Its first seven skill areas move from specifying outcomes and mapping programming through instrument design, implementation fidelity, evidence collection, analysis and reporting, and use of results for improvement. For this discussion, those activities are grouped into planning, evidence collection, analysis, reporting, and improvement; implementation fidelity remains a concern across them. During planning, GenAI can draft outcomes, suggest measures, align outcomes with taxonomies, generate assignment prompts, or propose curriculum maps. During evidence collection, it can help design surveys, create item drafts, organize artifacts, or build feedback prompts. During analysis, it can summarize open-ended comments, suggest themes, compare rubric language, produce tables, or draft narrative findings. During reporting, it can write executive summaries, accreditation language, action plans, or stakeholder-specific communications. During improvement, it can generate possible interventions, workshop activities, or faculty development materials.

These capabilities address work that requires sustained faculty and staff participation. Hutchings (2010) identifies time and competing responsibilities as barriers to faculty involvement in assessment, while Ewell (2009) describes the demands of external accountability. GenAI may reduce some of that burden, but speed can obscure the questions assessment professionals are trained to ask. A rubric that reads well may not represent the intended construct. A summary of student comments may flatten important differences by modality, program, or student group. A set of GenAI-generated findings may exaggerate patterns or miss discrepant cases. An accreditation narrative may sound convincing while overstating evidence. Fluent text still requires review before it can support an educational claim.

Validity theory is central here. The Standards for Educational and Psychological Testing define validity in terms of the evidence and theory supporting proposed interpretations of test scores for specified

uses (American Educational Research Association et al., 2014). Messick (1989) framed validity as an integrated evaluative judgment about the degree to which evidence and theory support interpretations and uses of assessment results. Kane (2013) similarly emphasized validation as an argument connecting observations, interpretations, and decisions. If GenAI helps generate, score, summarize, or narrate evidence, then the validation argument must expand to include the role of the system, the prompt, the input data, the human review process, and the intended use. The question is how GenAI shaped the evidence and whether the resulting evidence supports the proposed interpretation and use.

Fairness and reliability also become more complicated. The *Standards for Educational and Psychological Testing* identify fairness as fundamental to assessment design, interpretation, and use (American Educational Research Association et al., 2014). If GenAI is used in item drafting, rubric design, qualitative coding, or scoring, assessment professionals must ask whether the output represents diverse student language, avoids construct-irrelevant variance, and can be reviewed consistently. If GenAI contributes to scoring or coding, institutions need evidence of inter-rater consistency, human review, and procedures for resolving disagreements (Stemler, 2004). GenAI does not remove the need for item review, rubric norming, or bias review. It makes those practices more urgent.

This is why the assessment professional's role becomes more important rather than less important. If GenAI can quickly produce plausible evidence products, someone must decide whether those products are useful, defensible, ethical, and connected to improvement. That work requires assessment knowledge, AI literacy, institutional context, and time.

Defining Capacity: Time as Assessment Infrastructure

Capacity is often used broadly to describe staffing, expertise, tools, money, or organizational readiness. Those elements matter. In this framework, however, capacity is defined primarily as time because time is the most immediate condition that determines whether GenAI engagement becomes responsible practice. Capacity means time on the calendar. Not theoretical time. Not 'explore this when you can.' Real time: a protected afternoon to test prompts, a pilot cycle to compare outputs, a meeting where assessment staff can ask messy questions without being expected to have the policy already written.

This emphasis is necessary because GenAI is often marketed as a time-saving technology. In mature workflows, some GenAI uses may save time. But responsible adoption initially requires more time, not less. Assessment professionals must learn tool capabilities, test prompts, compare outputs, check for hallucinations, evaluate bias, consult with information technology and legal colleagues, document workflows, train faculty, and decide which tasks should remain human-led. That front-end work is not inefficiency. It is the cost of responsible implementation.

Professional learning research supports this point. Learning new professional practices requires sustained engagement, feedback, coherence with work demands, and opportunities to apply learning in context (Desimone, 2009; Guskey, 2002). Job demands-resources theory also helps explain why time matters. When demands increase without corresponding resources, professionals experience strain

and are less able to engage in learning, innovation, or discretionary improvement work (Bakker & Demerouti, 2007). GenAI adds new demands even when it later promises efficiencies. Without time, the professional development becomes performative and the innovation becomes unpaid overtime.

Capacity is not simply a matter of how many years have passed since GenAI emerged. Several years from now, time will still matter because the technology, policy environment, and assessment use cases will continue changing. New models, embedded platform features, institutional data rules, accreditation expectations, and faculty practices will create new learning cycles. The relevant question is not whether everyone has had enough time in a historical sense. The question is whether assessment professionals have protected time for the specific GenAI-mediated work they are being asked to do.

The amount of time needed will vary. A professional using GenAI to brainstorm workshop titles needs less protected time than one testing GenAI-assisted qualitative coding of student reflections. A professional with prior experience in data analysis, prompt design, or assessment software may need less initial tool learning but still needs time for validation and documentation. A high-risk use involving student artifacts, program decisions, or accreditation claims requires more review than a low-risk use involving non-sensitive drafting. Capacity is therefore relational. It depends on the person, the task, the risk, the data, and the intended use.

Limited time can also create unequal opportunities for experimentation. Some assessment offices may have flexible roles, supportive supervisors, and colleagues who can share that work. Others may be thinly staffed or face reporting demands that leave little room to test a new workflow. Leaders who interpret limited engagement as resistance risk overlooking those differences. The framework directs them to examine the time available for experimentation before drawing conclusions about professional stance.

For these reasons, time should be treated as assessment infrastructure. Institutions routinely invest in assessment management systems, survey tools, dashboards, and reporting platforms. GenAI engagement requires a parallel investment in temporal infrastructure: release time, pilot periods, learning communities, summer institutes, internal grants, protected meetings, and workload recognition. If GenAI is expected to reshape assessment practice, GenAI learning cannot remain extracurricular.

Defining Disposition: From Reaction to Professional Stance

If capacity explains whether assessment professionals have room to engage, disposition explains how they make sense of that engagement. Disposition refers to the professional stance a person brings to GenAI-mediated work. It includes beliefs about evidence, trust, expertise, academic integrity, equity, efficiency, student learning, and institutional risk. It also includes affect. People may encounter GenAI with anxiety, curiosity, fatigue, possibility, pressure, or a sense of professional threat. Workplace affect shapes attention, motivation, and responses to organizational events (Weiss & Cropanzano, 1996).

Disposition is also cultural. Organizational culture shapes what people consider legitimate, safe, risky, or professionally valued (Schein, 2010). In an assessment culture where annual reports are treated as compliance artifacts, GenAI may be interpreted primarily as a way to produce reports faster. In a culture centered on improvement, it may be interpreted as a way to free time for interpretation and action. In a low-trust culture, professionals may hide experimentation or avoid GenAI entirely. In a psychologically safe culture, they may be more willing to test, question, and report failures (Edmondson, 1999).

This article uses preservationist and generative to describe the two ends of the disposition axis. A preservationist stance begins from continuity, evidence protection, and concern about harm; a generative stance begins from experimentation, redesign, and possibility. These terms describe professional orientations rather than age, ability, or fixed identities.

A preservationist disposition can protect validity, fairness, student privacy, and the integrity of evidence. The commitment to using evidence for educational improvement runs through the AAHE principles of good practice and the literature distinguishing assessment for improvement from assessment for compliance (AAHE, 1992; Ewell, 2009). They provide a basis for examining GenAI rather than a reason to avoid that examination. With students and faculty already reporting GenAI use (Freeman, 2025; Watson & Rainie, 2026), assessment professionals need to consider how that use changes the evidence available to them.

A generative disposition also has strengths and risks. Professionals with this orientation may see GenAI as a way to reduce busywork, support faculty, improve qualitative analysis, develop better communication, and imagine richer evidence systems. But enthusiasm is not judgment. A generative stance becomes shallow if it treats speed as improvement, output as evidence, or automation as innovation. Self-determination theory reminds us that motivation is strengthened by autonomy, competence, and relatedness (Deci & Ryan, 2000). For GenAI engagement to mature, professionals need agency in how GenAI is used, competence in evaluating outputs, and connection to colleagues who can help interpret results.

The productive disposition is critical openness. It accepts that GenAI is consequential enough to require serious engagement and risky enough to require disciplined caution. It asks what GenAI can help assessment professionals do better, where human judgment must remain central, how use should be disclosed, and what evidence would justify implementation. This stance is where assessment literacy and AI literacy meet. AI literacy includes understanding capabilities, limitations, ethical risks, and appropriate uses across AI technologies, including GenAI (Chan, 2023; Long & Magerko, 2020; Ng et al., 2021). For assessment literacy, Horst and Prendergast's (2020) Assessment Skills Framework specifies knowledge and skills in alignment, reliability and validity, interpretation and use, and ethics, diversity, and inclusion. Those domains provide the assessment questions that GenAI experimentation must address.

Institutional Support: The Runway for Responsible Engagement

Capacity and disposition develop within an institutional environment. Support determines whether professionals can move from awareness to practice, from informal experimentation to documented pilots, and from individual curiosity to shared stewardship. Organizational readiness for change depends on shared commitment to change and confidence in the collective ability to carry it out, including judgments about task demands and resources (Weiner, 2009). For GenAI in assessment, readiness takes shape through policy, access, professional learning, workload recognition, and risk calibration.

The first layer of support is policy clarity. Assessment professionals need to know which tools are approved, what data may be entered, how student privacy is protected, whether de-identified qualitative comments can be analyzed, how GenAI-assisted work should be disclosed, and who is accountable when GenAI output is used in official reporting. Vague policy invites either avoidance or shadow practice.

The second layer is access. It is difficult to ask professionals to become GenAI literate while leaving them to use free public tools, personal subscriptions, or unclear data agreements. Unequal access also shapes disposition. Professionals with secure institutional tools may experience GenAI as legitimate and supported. Professionals without access may experience it as risky, unofficial, or personally burdensome.

The third layer is professional development, but not the one-and-done kind. Assessment professionals need repeated opportunities to test use cases, compare outputs, critique prompts, review risks, and share failures. Communities of practice are useful because GenAI learning is not just technical learning; it is identity development within a changing profession (Wenger et al., 2002). A single workshop can introduce the vocabulary. It cannot build judgment.

The fourth layer is workload recognition. If an assessment professional is asked to serve on a GenAI policy group, consult with faculty, build prompt resources, test approved tools, design GenAI-related assessment guidance, and answer emerging questions, that labor must be recognized. Otherwise, the institution benefits from innovation while denying the cost of producing it.

The fifth layer is risk calibration. Not all GenAI uses carry the same level of consequence. Drafting a workshop agenda is different from scoring student artifacts. Generating sample rubric language for faculty review is different from using GenAI output as evidence in an accreditation report. Institutions need risk tiers that connect oversight to consequence. Low-risk uses should not require excessive bureaucracy. High-risk uses should not proceed without validation, documentation, and human review.

The Four Conditions of the Capacity-Disposition Framework

The four conditions below are intended to make institutional conversations more concrete. Their names are deliberately memorable. Diffusion of innovations theory reminds us that innovations travel through social systems, and adoption is shaped partly by whether people can see, discuss, and try an innovation in ways that fit their context (Rogers, 2003). Plain language labels can function as boundary

objects: they give people a nonpunitive way to talk about what is happening without reducing the issue to individual virtue or resistance.

The Calendar Trap

The Calendar Trap describes low time capacity paired with a preservationist disposition. GenAI is not simply rejected here because professionals are uninformed or unwilling. It is held at a distance because workload and caution reinforce each other. The calendar is full, the risk feels high, and the safest move appears to be staying with familiar processes.

This condition may arise in offices where assessment work is compliance-heavy, staffing is thin, and GenAI expectations arrive without additional resources. A professional in The Calendar Trap may believe GenAI threatens authentic evidence, student privacy, or academic integrity. Those concerns may be legitimate. But because there is little time to examine GenAI carefully, concern hardens into avoidance. The institution may then misread avoidance as lack of vision when it is partly a rational response to scarcity.

The support strategy should begin with time and respect, not persuasion. These professionals need low-stakes opportunities to examine GenAI through familiar assessment values: validity, fairness, transparency, and use. A useful starting point might be reviewing GenAI-generated rubric language for construct alignment, not asking them to redesign assessment reporting overnight.

The Margin Lab

The Margin Lab describes a generative disposition paired with low time capacity. The name is intentional: experimentation is happening, but it is happening in the margins rather than inside an institutional structure. Professionals in this condition are curious and often already using GenAI to draft communications, summarize ideas, brainstorm faculty development activities, or explore qualitative themes. But the work is informal, under-documented, and often uncompensated.

The Margin Lab is valuable and vulnerable. Because these professionals are willing, they often become informal translators for colleagues. They are asked to join committees, give demonstrations, answer faculty questions, and explain policy implications. They may be moving the institution forward, but without protected time, approved tools, or workload recognition, the work depends on personal energy. That is not sustainable.

The support strategy is institutionalization. Leaders should identify promising informal experiments and convert them into supported pilots. That means approved tools, written procedures, documentation standards, modest release time, and opportunities to evaluate whether the workflow actually improves assessment practice.

The Evidence Gate

The Evidence Gate describes high time capacity paired with a preservationist disposition. Professionals in this condition may have access to tools, training, and institutional support, yet remain cautious. They are not simply blocking change. They may be asking the questions the institution most needs: What

would count as adequate evidence that GenAI-assisted coding is trustworthy? How will we know whether GenAI-generated rubric language is biased? What student data can be used? What human review is required? What happens when GenAI output is wrong?

The Evidence Gate becomes unproductive only when every uncertainty becomes a reason to postpone any test. Responsible assessment practice does not require certainty before inquiry; it requires inquiry designed well enough to produce usable evidence. The value of The Evidence Gate is that skepticism can become a method.

The support strategy is disciplined experimentation. Rather than asking cautious professionals to become cheerleaders, institutions should ask them to lead validation checks, risk reviews, item-review protocols, inter-rater comparisons, disclosure standards, and human review procedures. Their caution becomes an asset when it is given a constructive task.

The Practice Commons

The Practice Commons describes high time capacity paired with a generative disposition. Professionals in this condition have room to experiment and the willingness to do so, but they are not simply enthusiasts. They are stewards. They test GenAI workflows, document decisions, share resources, invite critique, protect student data, and connect GenAI use to improvement rather than mere production.

This condition aligns with Miller's (2024) call for assessment professionals to help shape ethical and inclusive GenAI use and with Hathcoat et al.'s (2026) description of assessment professionals as curators and auditors of GenAI-mediated evidence. The Practice Commons is not one person with advanced prompting skill. It is a shared environment where practices become visible, reviewable, and improvable.

The risk is that The Practice Commons collapses back into one person. Once a campus finds someone who can explain GenAI clearly, that person can quickly become responsible for everything: policy, training, faculty consultation, tool testing, ethical review, and implementation. A commons requires shared infrastructure. Otherwise, it becomes another version of The Margin Lab with better branding. Table 1 summarizes the institutional responses associated with each condition. These conditions support institutional planning and self-study. They should not be used to rank individual professionals or assign permanent identities.

Applying the Framework to Assessment Work

The framework becomes useful when applied to ordinary assessment tasks. The examples that follow illustrate possible responses to each condition rather than established effects. Consider learning outcomes and curriculum mapping. GenAI can draft outcomes, suggest possible alignment gaps, and propose clearer language. That can help faculty move past the blank page. But outcomes represent program values and disciplinary expectations. In The Calendar Trap, a scheduled review of one draft outcome offers a manageable way to examine concerns about disciplinary fit. A professional in The Margin Lab may generate promising drafts without time to facilitate faculty interpretation. A

professional in The Evidence Gate can lead a structured review of whether the drafts represent the intended construct. In The Practice Commons, faculty and assessment professionals can share that review process, document revisions, and connect the outcomes to curriculum decisions.

Table 1

Institutional Moves Associated with Each Framework Condition

Condition	Primary institutional risk	Developmental need	Useful institutional move
The Calendar Trap	GenAI is avoided, treated only as a threat, or disconnected from actual assessment practice.	Protected time, low-stakes exposure, and respect for legitimate validity concerns.	Begin with assessment values: data protection, evidence integrity, fairness, and small tests of non-sensitive tasks.
The Margin Lab	Innovation stays informal, undocumented, or dependent on personal subscriptions and uncompensated labor.	Governance, tool access, workload recognition, and pilot structure.	Convert promising experiments into supported pilots with approved tools, written procedures, and time credit.
The Evidence Gate	Skepticism becomes delay, policy paralysis, or an endlessly deferred readiness conversation.	A constructive role for skepticism.	Ask cautious practitioners to lead validation checks, risk reviews, item-review protocols, and human review standards.
The Practice Commons	Expertise is concentrated in one person or office and becomes unsustainable.	Distributed stewardship and shared infrastructure.	Build communities of practice, reusable templates, disclosure norms, risk tiers, and cross-functional governance.

Rubric development offers a similar example. GenAI can generate dimensions and performance descriptors quickly, and Brookhart (2013) reminds us that rubrics are useful only when criteria communicate meaningful expectations and support learning. GenAI-generated rubrics may be polished but generic. They may miss disciplinary nuance, equity concerns, or the features of performance that faculty most value. For someone in The Calendar Trap, reviewing one generated criterion during a scheduled norming session may be a manageable starting point. The same session gives a professional in The Margin Lab time to test descriptors that otherwise remain informal drafts. Professionals in The Evidence Gate can lead construct and fairness reviews using sample student work. In The Practice Commons, the results become shared rubrics and a documented norming process, with periodic checks on scoring consistency.

Qualitative analysis is a promising use case for offices facing large sets of open-ended survey comments, reflections, focus group notes, or program narratives. GenAI can summarize themes

quickly, but qualitative meaning is contextual. Slotnick and Boeing (2025) examine the promise of GenAI in qualitative assessment while emphasizing human expertise and contextual interpretation. A small comparison of human and GenAI coding can give professionals in The Calendar Trap a bounded way to investigate their concerns, using data authorized for the selected tool. In The Margin Lab, the priority is time to check preliminary themes against the comments, including discrepant cases. Colleagues in The Evidence Gate can lead that comparison and specify review criteria. In The Practice Commons, shared coding protocols, recorded prompts, and peer review make the analysis available for scrutiny and reuse.

Annual reporting and accreditation create another test. GenAI can draft executive summaries, findings narratives, and action-plan language. Used carefully, it may reduce writing burden and free time for sense-making. Used poorly, it can produce smoother reports without better use of evidence. In The Calendar Trap, a reporting deadline may leave concerns about GenAI-assisted claims unexamined; in The Margin Lab, a promising draft may circulate before its sources are checked. Protected time changes the task. Professionals in The Evidence Gate can trace claims to their supporting evidence. In The Practice Commons, that review becomes part of a shared reporting process, with time reserved for interpretation and improvement planning. The practical question is whether GenAI helps the institution use evidence to improve learning.

Faculty development is also affected. Faculty are asking what GenAI means for assignment design, academic integrity, feedback, and evidence of learning. Assessment professionals are well positioned to support those conversations because they understand learning outcomes, evidence, and use. Yet faculty development requires relationship-building. A GenAI-generated resource alone offers little opportunity for the discussion and feedback involved in sustained professional learning (Desimone, 2009). A brief discussion of one familiar assignment may give a professional in The Calendar Trap a feasible entry point, while professionals in The Margin Lab need protected time to turn informal advice into reviewed guidance. Colleagues in The Evidence Gate can help faculty decide what evidence would justify a proposed use. The Practice Commons sustains this work through recurring workshops where faculty test, critique, and revise assessment approaches together.

Student use of GenAI also changes the questions assessment professionals need to help faculty ask about learning evidence. When GenAI contributes to a final product, faculty need to determine which features of that product support the intended claim about student learning. A professional in The Evidence Gate can help identify weaknesses in that argument before an assignment or rubric is adopted, then specify what additional evidence is needed. Protected review time makes this work feasible for professionals in The Calendar Trap and gives those in The Margin Lab a way to examine informal redesigns. In The Practice Commons, faculty and assessment professionals can compare and document approaches such as iterative drafts, oral explanations, or process reflections. The framework directs attention to who can undertake this review, with what time and support.

Evidence Quality as a Test of the Framework

A report's fluency alone does not establish the quality of its underlying evidence. The framework directs leaders to ask whether professionals have time to verify claims and whether their concerns

inform that review. In The Calendar Trap, concerns about evidence may remain untested because no review time is available. In The Margin Lab, experimentation may outpace scrutiny of the resulting outputs. The Evidence Gate gives skeptical review a defined role, while The Practice Commons supports shared and documented review. Whether these conditions produce the expected differences in evidence quality remains an empirical question.

If GenAI is used to draft assessment items, prompts, or survey questions, item review remains essential. Reviewers should examine construct alignment, reading level, accessibility, bias, clarity, and unintended cueing. If GenAI is used to revise rubrics, faculty should compare descriptors against actual student work and intended performance levels. If GenAI is used to score or code artifacts, institutions should compare GenAI-assisted results with human raters, calculate agreement where appropriate, examine disagreement patterns, and decide whether GenAI is supporting formative insight or consequential judgment. Agreement checks contribute evidence about the dependability of scoring or coding (Stemler, 2004). They give professionals in The Evidence Gate a constructive role in validation and provide procedures that can be shared through The Practice Commons.

Fairness review asks whether GenAI-assisted interpretations privilege dominant language patterns, miss culturally specific meanings, or penalize multilingual expression. Culturally responsive assessment scholarship has long warned against treating dominant norms as neutral (Montenegro & Jankowski, 2017; Randall, 2021). Those questions explain the value of a preservationist disposition when paired with time to examine diverse student artifacts and local contexts. The more consequential the intended use, the stronger the supporting evidence must be.

Disclosure also makes movement out of The Margin Lab visible: an informal workflow becomes one colleagues can inspect and improve. Shared documentation should identify the data used, GenAI's contribution, human review, and remaining limitations. In The Practice Commons, maintaining these records becomes part of the workflow. That requires protected time alongside the analysis and reporting themselves.

What Is General and What Is Assessment Specific?

The Capacity-Disposition Framework is not claiming that capacity and disposition matter only for assessment professionals. They likely matter across knowledge work. Lawyers, clinicians, analysts, faculty, advisors, and administrators all face versions of the same problem: GenAI can help, but responsible use depends on time, trust, expertise, and context. The assessment-specific contribution is the location of these dimensions inside the evidentiary responsibilities of assessment work.

Assessment professionals occupy a boundary role. They translate between faculty, students, accreditors, institutional leaders, data systems, and improvement processes. They are expected to help others make claims about learning and then use those claims responsibly. That makes GenAI engagement especially consequential. A poorly reviewed GenAI-generated interpretation of student learning evidence can shape curriculum decisions, accreditation narratives, resource allocation, or public claims about institutional quality.

Several factors may moderate how professionals experience the framework: prior AI literacy, role authority, staffing level, institutional trust, professional identity, disciplinary background, data sensitivity, access to secure tools, and exposure to earlier technology failures. These factors are not outside the model. They help explain movement within it. For example, strong policy clarity may increase capacity by reducing uncertainty. Prior technical experience may reduce the time needed for basic tool learning. A low-trust culture may push even a generative professional into cautious practice. The framework is best used as a diagnostic lens, not a predictive formula.

Implications for Assessment Leaders

The first implication is simple: assess the conditions before judging the people. If assessment professionals are not engaging GenAI, leaders should ask why. Do they have protected time? Approved tools? Policy clarity? Training? A reason to believe GenAI can support better assessment rather than simply faster reporting? Labeling people as resistant is easier than examining the conditions that produce resistance.

Second, leaders should audit workload. GenAI should not be layered onto every existing assessment obligation. Institutions should ask which reports are read, which processes are duplicative, which data requests support improvement, and which practices survive only because they have always been done. If GenAI becomes an excuse to do more low-value assessment work faster, it will worsen the problem.

Third, leaders should create protected experimentation structures. These might include GenAI assessment labs, prompt review sessions, summer pilot cohorts, internal mini-grants, assessment innovation fellowships, or cross-functional governance teams. Each structure should produce something reusable: a tested workflow, a disclosure template, a risk matrix, a faculty guide, an item-review protocol, or a case example.

Fourth, leaders should pair dispositions intentionally. A productive GenAI pilot should include someone who sees possibilities and someone who sees risks. The generative professional can help identify what might be improved. The preservationist professional can help determine what evidence would make the improvement defensible. That pairing moves the work away from hype versus resistance and toward professional judgment.

Fifth, leaders should normalize disclosure and risk tiering. Low-risk uses might include brainstorming workshop titles, drafting non-sensitive explanations, or generating sample rubric language for human review. Moderate-risk uses might include summarizing de-identified comments or drafting report sections from verified data. High-risk uses might include scoring student work, making program-level claims, or producing evidence for consequential decisions. Institutions should match oversight to risk.

Finally, leaders should connect GenAI engagement to assessment culture. Wilton and Methot (2020) argue that assessment culture can either move institutions forward or hold them back. GenAI can strengthen assessment culture if it frees time for faculty conversation, student partnership, equity review, and improvement planning. It can weaken assessment culture if it produces more polished documentation without deeper use of evidence.

A Research Agenda

The Capacity-Disposition Framework is meant to be tested, revised, and improved. A first research need is survey work that measures protected time, AI literacy, disposition, institutional support, risk perception, and GenAI use among assessment professionals. A useful design would test whether capacity and disposition interact to predict documented GenAI use, informal use, avoidance, or leadership in governance. If the framework is supported, professionals in The Practice Commons should report more documented and shared GenAI practices than those in The Margin Lab, who should report more informal use and workload strain. The Calendar Trap and The Evidence Gate offer a parallel comparison: whether protected time is associated with a shift from avoidance toward structured validation and governance.

Second, longitudinal studies could examine movement across conditions. Researchers might follow assessment offices across an academic year, collecting time-use diaries, interviews, policy artifacts, and examples of GenAI-assisted work. The framework would predict that protected time, secure tools, and communities of practice help professionals move from The Margin Lab to The Practice Commons, while rising workload or unclear policy pushes them back toward the margins.

Third, design-based studies could evaluate institutional interventions. For example, a campus might implement a GenAI assessment fellowship that provides release time, approved tools, and structured peer review. Researchers could examine whether participants develop stronger AI literacy, produce more transparent workflows, and report greater confidence in evaluating GenAI outputs. Another campus might assign preservationist professionals to lead risk audits and examine whether that role increases engagement without requiring them to abandon caution.

Fourth, validation studies are needed for specific workflows. Researchers could compare GenAI-assisted and human-only qualitative coding of student reflections, examine agreement and theme quality, and analyze whether certain student groups are misrepresented. Rubric studies could compare faculty-generated, GenAI-generated, and collaboratively revised rubrics for construct alignment, clarity, fairness, and scoring consistency. Reporting studies could examine whether GenAI-assisted annual reports are more readable but not necessarily more useful for improvement.

Fifth, scholarship should examine the labor politics of GenAI adoption. Who is asked to learn GenAI? Whose time is protected? Whose experimentation is recognized? Who becomes responsible for translating policy to faculty and faculty anxiety to institutional leaders? The ethics of GenAI adoption include the ethics of professional workload.

Conclusion

GenAI is changing the environment in which assessment professionals make claims about learning, support faculty, prepare accreditation evidence, and help institutions improve. Existing work addresses AI literacy through competency frameworks (Long & Magerko, 2020), policy through institutional guidance (Chan, 2023), and evidence quality through assessment-specific discussions of human oversight and interpretation (Miller, 2024; Slotnick & Boeing, 2025). The Capacity-Disposition

Framework asks what time and professional stance assessment professionals need to put those commitments into practice.

Assessment professionals cannot responsibly explore and implement GenAI without temporal capacity. Enthusiasm without time becomes The Margin Lab, where innovation happens after hours and depends on unrecognized labor. Caution without time becomes The Calendar Trap, where legitimate concerns harden into distance. Caution with time can become The Evidence Gate, where skepticism becomes review, testing, and governance. Generative engagement with time and support can become The Practice Commons, where GenAI practice is shared, ethical, documented, and connected to improvement.

Research on technology adoption, professional learning, and organizational readiness emphasizes facilitating conditions, sustained learning, and resources for change (Desimone, 2009; Venkatesh et al., 2003; Weiner, 2009). The Capacity-Disposition Framework brings those insights into assessment practice by pairing protected time with professional stance. Institutions can act on that argument by providing approved tools, clarifying policy, recognizing labor, and creating communities where preservationist and generative professionals work together.

Assessment's purpose remains the use of evidence to improve student learning (AAHE, 1992; Ewell, 2009; Kuh et al., 2015). GenAI introduces new questions about how that evidence is produced and interpreted. The framework offers a way to examine whether assessment professionals have the time, professional stance, and institutional support to answer those questions. Its value should be judged by whether it helps institutions create the conditions for careful decisions about where GenAI belongs.

References

- American Association for Higher Education. (1992). *Principles of good practice for assessing student learning*. AAHE Assessment Forum.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Bakker, A. B., & Demerouti, E. (2007). The job demands-resources model: State of the art. *Journal of Managerial Psychology*, 22(3), 309–328. <https://doi.org/10.1108/02683940710733115>
- Banta, T. W., & Palomba, C. A. (2015). *Assessment essentials: Planning, implementing, and improving assessment in higher education* (2nd ed.). Jossey-Bass.
- Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. ASCD.
- Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 20, Article 38. <https://doi.org/10.1186/s41239-023-00408-3>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

- Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181–199.
<https://doi.org/10.3102/0013189X08331140>
- Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2), 350–383. <https://doi.org/10.2307/2666999>
- Ewell, P. T. (2009). Assessment, accountability, and improvement: Revisiting the tension (NILOA Occasional Paper No. 1). National Institute for Learning Outcomes Assessment.
- Freeman, J. (2025). *Student Generative AI Survey 2025* (HEPI Policy Note 61). Higher Education Policy Institute. <https://www.hepi.ac.uk/reports/student-generative-ai-survey-2025/>
- Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Teaching*, 8(3), 381–391. <https://doi.org/10.1080/135406002100000512>
- Hathcoat, J. D., Slotnick, R., & Miller, W. (2026). The end of assessment? Disruption and transformation in the age of AI. *Research & Practice in Assessment*, 21(3), 5–21.
- Horst, S. J., & Prendergast, C. O. (2020). The Assessment Skills Framework: A taxonomy of assessment knowledge, skills and attitudes. *Research & Practice in Assessment*, 15(1), 1–25.
<https://www.rpajournal.com/dev/wp-content/uploads/2020/05/The-Assessment-Skills-Framework-RPA.pdf>
- Hutchings, P. (2010). Opening doors to faculty involvement in assessment (NILOA Occasional Paper No. 4). National Institute for Learning Outcomes Assessment.
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1–73. <https://doi.org/10.1111/jedm.12000>
- Kuh, G. D., Ikenberry, S. O., Jankowski, N. A., Cain, T. R., Ewell, P. T., Hutchings, P., & Kinzie, J. (2015). Using evidence of student learning to improve higher education. Jossey-Bass.
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). Association for Computing Machinery. <https://doi.org/10.1145/3313831.3376727>
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). American Council on Education/Macmillan.
- Miller, W. (2024). Adapting to AI: Reimagining the role of assessment professionals. *Intersection: A Journal at the Intersection of Assessment and Learning*, 5(4), 99–113.
<https://doi.org/10.61669/001c.121439>
- Mollick, E. R., & Mollick, L. (2024). Instructors as innovators: A future-focused approach to new AI learning opportunities, with prompts. *SSRN Electronic Journal*.
<https://doi.org/10.2139/ssrn.4802463>
- Montenegro, E., & Jankowski, N. A. (2017). Equity and assessment: Moving towards culturally responsive assessment (NILOA Occasional Paper No. 29). National Institute for Learning Outcomes Assessment.
- National Institute of Standards and Technology. (2024). *Artificial intelligence risk management framework: Generative artificial intelligence profile* (NIST AI 600-1).
<https://doi.org/10.6028/NIST.AI.600-1>
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, Article 100041.
<https://doi.org/10.1016/j.caeai.2021.100041>

- Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory Into Practice*, 48(1), 4–11. <https://doi.org/10.1080/00405840802577536>
- Randall, J. (2021). 'Color-neutral' is not a thing: Redefining construct definition and representation through a justice-oriented critical antiracist lens. *Educational Measurement: Issues and Practice*, 40(4), 82–90. <https://doi.org/10.1111/emip.12429>
- Randazzo, S., Lifshitz-Assaf, H., Kellogg, K. C., Dell'Acqua, F., Mollick, E. R., Candelon, F., & Lakhani, K. R. (2025). Cyborgs, centaurs and self-automators: The three modes of human-GenAI knowledge work and their implications for skilling and the future of expertise (Harvard Business School Working Paper No. 26-036). <https://doi.org/10.2139/ssrn.4921696>
- Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.
- Schein, E. H. (2010). *Organizational culture and leadership* (4th ed.). Jossey-Bass.
- Slotnick, R. C., & Boeing, J. Z. (2025). Enhancing qualitative research in higher education assessment through generative AI integration: A path toward meaningful insights and a cautionary tale. *New Directions for Teaching and Learning*, 2025(182), 97–112. <https://doi.org/10.1002/tl.20631>
- Slotnick, R. C., Boeing, J., & Grillo Pinnelli, B. (2026). *GenAI Pulse Survey for higher education assessment professionals: Spring 2026 results*. Bridgewater State University. <https://www.genaiassessment.com/https://genaiassessment.neocities.org/Spring2026/Spring2026>
- Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research, and Evaluation*, 9(4). <https://doi.org/10.7275/96jp-xz07>
- Suskie, L. (2018). *Assessing student learning: A common sense guide* (3rd ed.). Jossey-Bass.
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478. <https://doi.org/10.2307/30036540>
- Watson, C. E., & Rainie, L. (2026). *The AI challenge: How college faculty assess the present and future of higher education in the age of AI*. American Association of Colleges and Universities; Elon University's Imagining the Digital Future Center. <https://imaginingthedigitalfuture.org/wp-content/uploads/2026/01/Elon-AACU-faculty-AI-survey-full-report-1-21-26.pdf>
- Weiner, B. J. (2009). A theory of organizational readiness for change. *Implementation Science*, 4, Article 67. <https://doi.org/10.1186/1748-5908-4-67>
- Weiss, H. M., & Cropanzano, R. (1996). Affective events theory: A theoretical discussion of the structure, causes and consequences of affective experiences at work. *Research in Organizational Behavior*, 18, 1–74.
- Wenger, E., McDermott, R., & Snyder, W. M. (2002). *Cultivating communities of practice: A guide to managing knowledge*. Harvard Business School Press.
- Wilton, S., & Methot, M. (2020). Moving forward or holding back? Creating a culture of academic assessment. *The Canadian Journal for the Scholarship of Teaching and Learning*, 11(2). <https://doi.org/10.5206/cjsotl-rcacea.2020.2.8009>

About the Author

Will Miller, Associate Vice President, Continuous Improvement and Institutional Performance, Embry-Riddle Aeronautical University will.miller@erau.edu