Research from the Office of RIDIL

Higher Education Assessments in the Age of GenAI: A Systematic Review

This white paper provides a summary of the work from:

Crompton, H., & Burke, D. (in press). Higher education assessments in the age of GenAI: A systematic review. *International Journal of Educational Technology in Higher Education*.

Research Author Bios:

Dr. Helen Crompton is a Professor of Instructional Technology in the Department of STEM and Professional Studies, and the Executive Director of the Research Institute for Digital Innovation in Learning at ODU. Dr. Diane Burke is a Senior Research Associate at the Research Institute for Digital Innovation in Learning at ODU.

Overview

Assessment in higher education has long centered on core principles such as validity, authenticity, and alignment with learning outcomes, alongside ongoing debates about how best to support academic integrity and meaningful feedback. The emergence of generative artificial intelligence (GenAI) is now placing renewed pressure on these established principles, prompting institutions to reconsider how assessment is designed, delivered, and evaluated. This systematic review examines empirical research published between 2024 and 2026 on how GenAI influences assessment practices and supports educators in assessing student learning.

Purpose of Research

This study examines assessment practices in HE in the GenAI milieu. Four questions guided this study. In studies examining assessment and GenAI in HE, which academic disciplines are represented? Which research purposes are represented in the studies? What assessment design approaches are HE educators adopting in response to GenAI? What affordances does GenAI offer educators within the assessment process in HE?

Method

A PRISMA systematic review methodology was used to determine the articles that would be included in this systematic review. A grounded coding approach then revealed trends in responses to this research question.

Findings and Discussion

The study addresses several research questions. In this white paper, two questions provide contextual insight into the disciplinary and research characteristics of the included studies, while two overarching questions examine the assessment designs emerging from the literature and how GenAI is being used to support assessment.

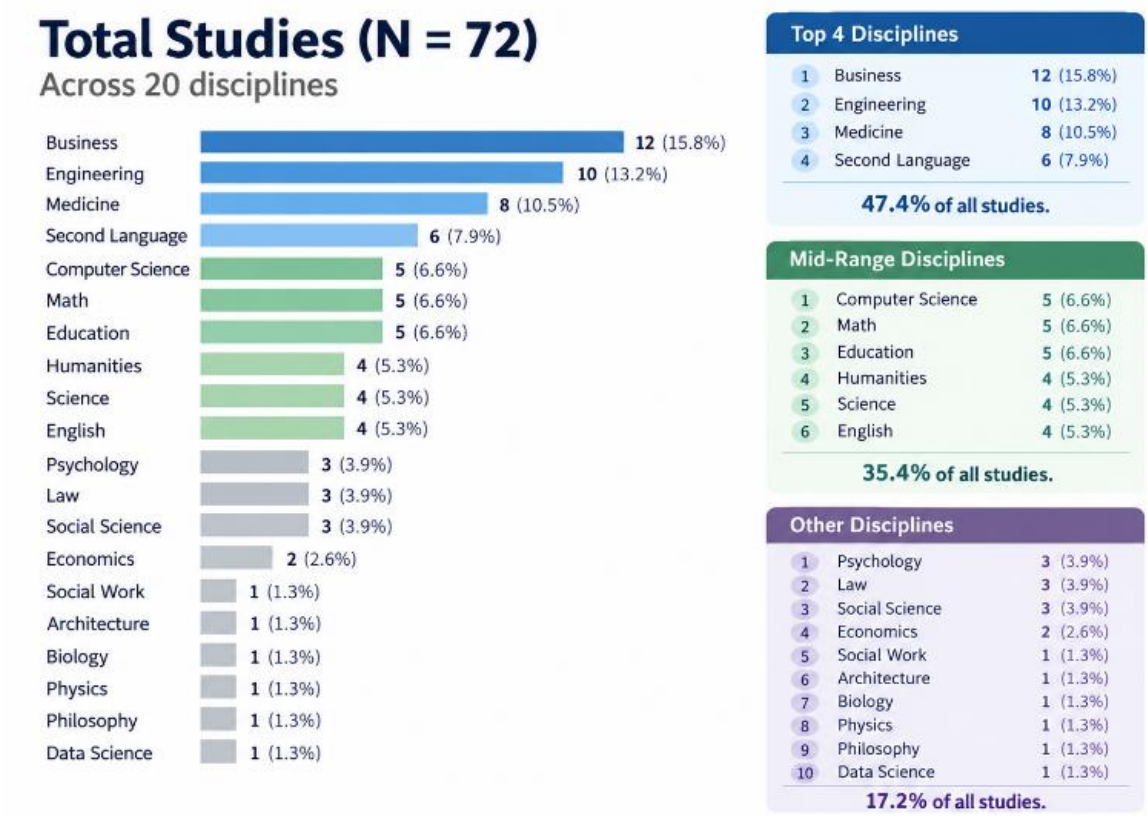
While the broader study gathered a wide range of background data, this white paper centers primarily on its two overarching questions: what assessment designs are emerging and how GenAI is being used to support

them. To ground these findings without detailing the full study's broader contextual survey, we draw on two selected questions that highlight the disciplinary domains and research purposes of the literature examined.

Disciplines Represented in the Study

The research studies for this systematic review spanned across 20 disciplines, indicating that GenAI assessment research is interdisciplinary and covers a wide range of academic fields. However, the distribution was uneven, with a small number of disciplines accounting for a large share of the literature. Business had the most studies, with 12. Engineering accounted for 10 studies, medicine for 8, and second language for 6. Together, these four disciplines accounted for 36 of the 72 studies. See Figure 1 for total disciplines represented.

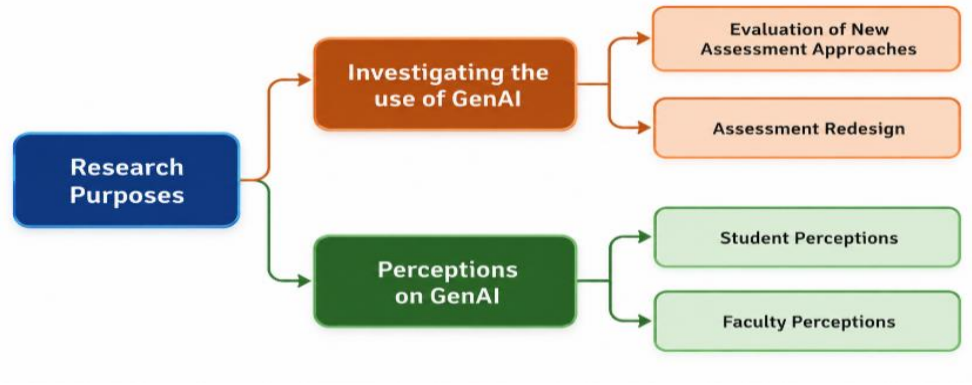
Figure 1: Disciplines Represented in the Study



Research Purposes

Two main categories of research purposes were identified: studies investigating the use of GenAI and studies on perspectives on GenAI. Of those that investigated the use of GenAI, two axial codes were uncovered: 1) evaluation of new assessment approaches and 2) assessment redesign. For the investigation of perceptions of GenAI, two axial codes emerged: 1) student perceptions and 2) faculty perceptions. See Figure 2.

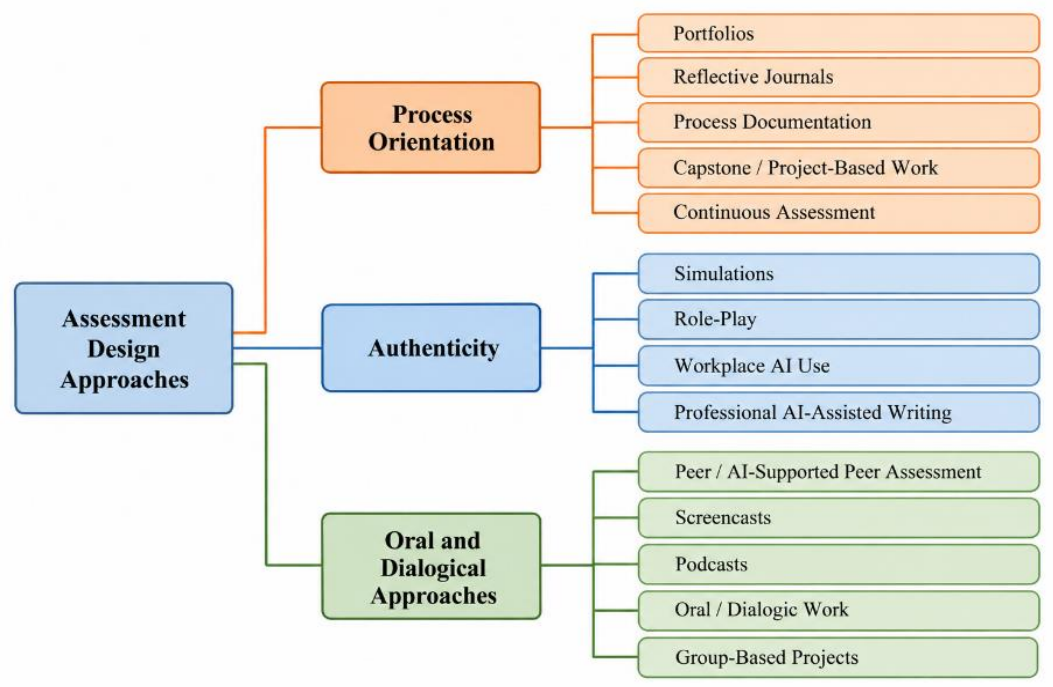
Figure 2: Research Purposes



Assessment Design Approaches

Three assessment design approaches emerged from the data: 1) Process Orientation, 2) Authenticity, and 3) Oral and Dialogical Approaches. These approaches reflect the main ways HE assessments are being redesigned in response to the opportunities and challenges GenAI presents. Process orientation captures the shift from judging a final artifact alone to assessing how learning develops, is documented, and is reflected upon. This approach prioritizes the "how" of learning rather than just the final, potentially AI-assisted, "what." Authenticity focuses on assessments that resemble real-world practice, making assessments more authentic, applied, and professionally aligned. Oral and dialogic approaches reflect a shift away from solitary, static learning toward socially mediated, interactive forms. Figure 3 shows specific examples of what these approaches could look like in practice.

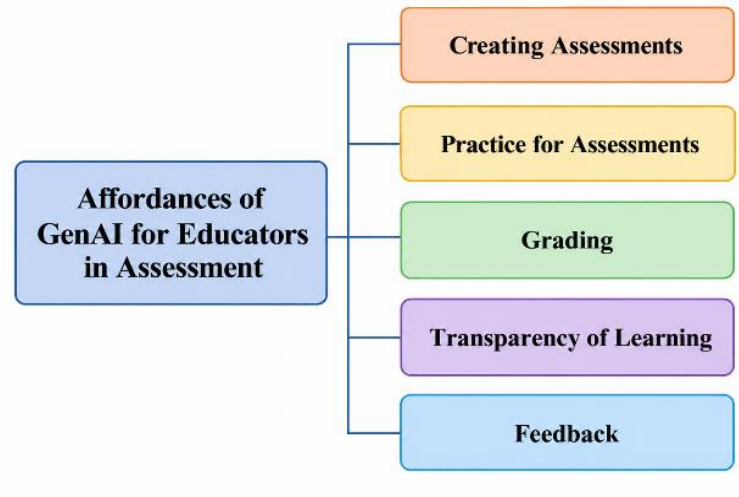
Figure 3: Assessment Design Approaches



Affordances

The final research question serves as the overarching focus of this study in examining what GenAI offers educators across the assessment process. Five codes emerged from the data: 1) creating assessments, 2) practice for assessments, 3) grading, 4) transparency of learning, and 5) feedback, with axial codes delineating the affordances within each main code. Taken together, the codes span the entire assessment cycle, from task design through student preparation to evaluating work and communicating results. Creating assessments and feedback were the most frequently identified affordances. See Figure 4.

Figure 4: Affordances

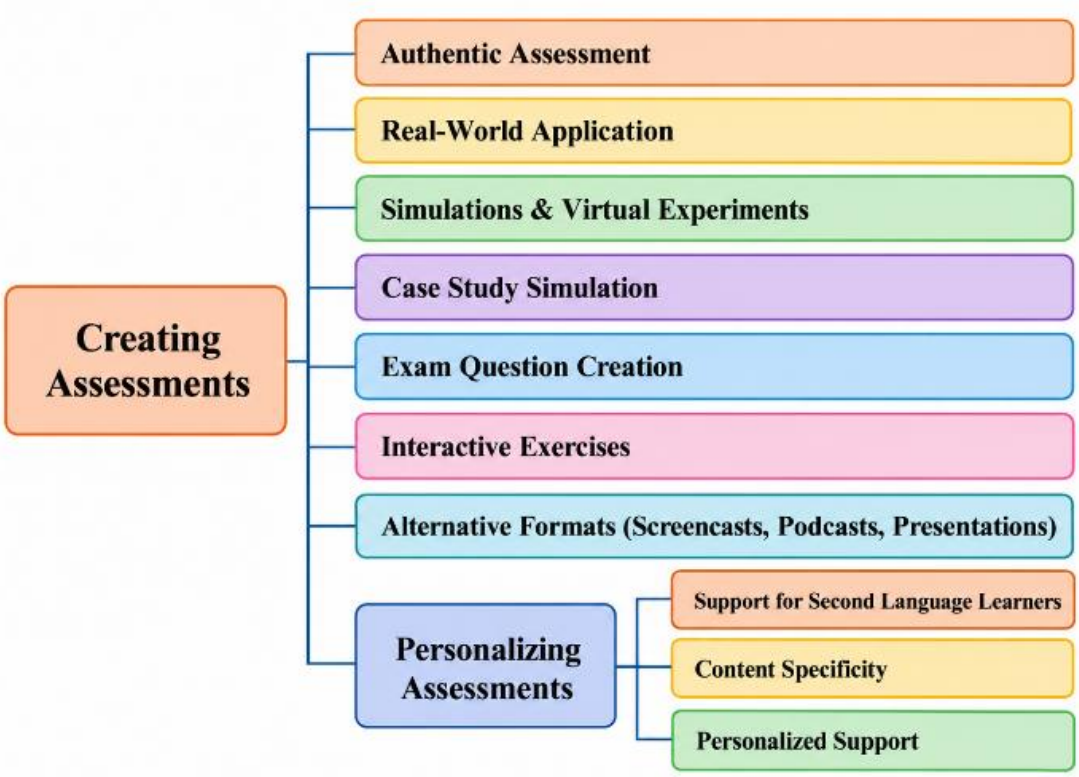


Each affordance from the findings of the study are each described in further detail below.

Creating Assessments

The most common affordance identified in the studies was educators' use of GenAI to create assessments. Eight axial codes emerged: authentic assessment, real-world application, simulations and virtual experiments, case study simulation, exam question creation, interactive exercises, alternative formats (e.g., screencasts, podcasts, and presentations), and personalizing assessments. The personalizing assessments axial code was further delineated into three sub-codes: support for second language learners, content specificity, and personalized support. See Figure 5.

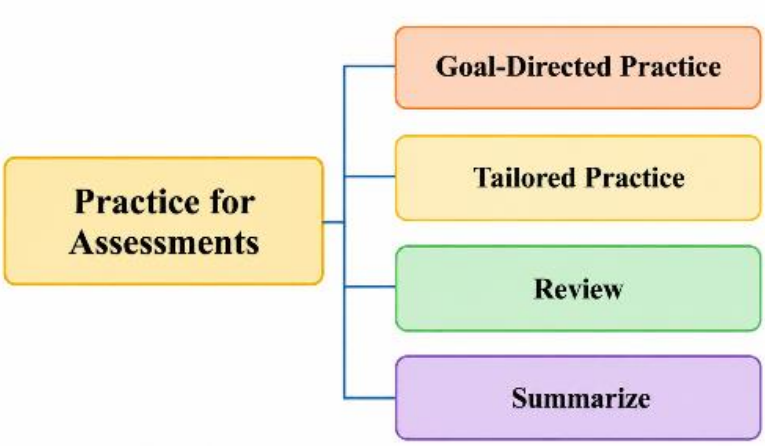
Figure 5: Creating Assessments



Practice for Assessments

The second code captured educators using GenAI to provide students with opportunities to practice for assessments. Four axial codes emerged: goal-directed practice, tailored practice, review, and summarize. The findings suggest that GenAI extends practice opportunities by offering unlimited, on-demand, low-stakes rehearsal without requiring a dedicated tutoring platform, repositioning assessment preparation from a resource constrained by instructor time to a formative activity students can access at the point of need. See Figure 6.

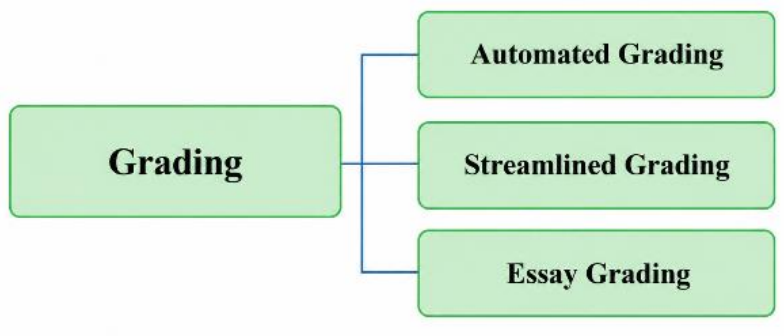
Figure 6: Practicing for Assessments



Grading

The third code involved using GenAI for grading. Three axial codes emerged: automated grading, streamlined grading, and essay grading. Grading is among the most established AI affordances in education. What has changed is that general-purpose GenAI tools can now grade sophisticated open-ended work. The studies in this review temper this affordance with caution, as questions of reliability, accuracy, and fairness remain, indicating that GenAI grading is best positioned as support to educator judgment rather than a replacement for it. See Figure 7.

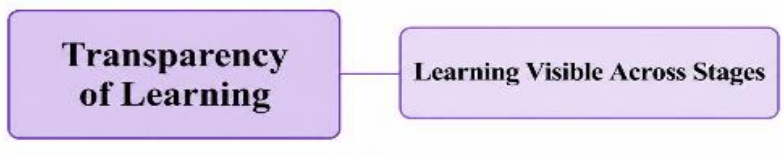
Figure 7: Grading



Transparency of Learning

The fourth code, transparency of learning, captured the use of GenAI to make student achievement visible to educators in connection with learning objectives, allowing educators to see how students are progressing toward the intended goals of instruction. See Figure 8.

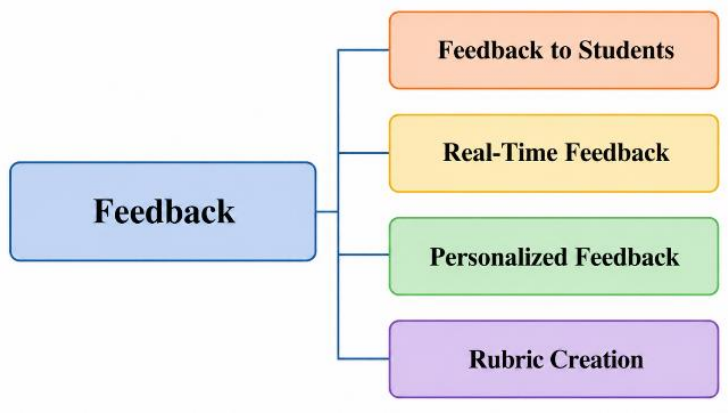
Figure 8: Transparency of Learning



Feedback

The final code, and one of the most frequent, was using GenAI to provide feedback. Four axial codes emerged: feedback to students, real-time feedback, personalized feedback, and rubric creation. See Figure 9.

Figure 9: Feedback



Implications of Findings

This review's findings have practical implications for educators and educational leaders. The identified assessment design approaches indicate that responding to GenAI through detection and restriction alone is unlikely to be sufficient; instead, the studies point to redesigning assessment around process, authenticity, and oral and dialogical interaction, and to leveraging the affordances GenAI can provide across the entire assessment cycle. Educators can approach GenAI not only as a challenge to be managed but as a design partner in assessment development. However, realizing these affordances responsibly requires educators to verify the accuracy and alignment of GenAI-created materials and to retain evaluative judgment over AI-supported grading and feedback, positioning GenAI as a support to professional judgment rather than a replacement for it.

For institutional leaders and policymakers, these findings indicate that faculty development is central to an effective response to GenAI. The redesigned approaches identified in this review demand more of educators, including more time, clearer criteria, and the capacity to evaluate learning processes rather than final products. Institutions should therefore pair clear policies on acceptable GenAI use with sustained professional learning that builds these capacities, rather than relying on integrity rules alone. Equity also needs attention.

Finally, because students will use GenAI in the assessment process, they need training and guidance to use it ethically and responsibly. This cannot be left to chance, and institutional leaders and policymakers need to put mechanisms in place to make this happen.

Conclusion

The findings of this review suggest that GenAI is not merely introducing a new set of tools into HE assessment but is contributing to a broader reconsideration of what assessment should value and how learning should be evidenced. The analysis of assessment redesign in response to GenAI indicates that GenAI appears to be prompting a move away from assessment centered on static written products and toward models that focus on process, authenticity, and oral/dialogic experiences. This indicates that the challenge posed by GenAI is not limited to concerns about misuse or task completion, but extends to fundamental questions about assessment purpose, design, and validity. The findings reveal that GenAI offers affordances across the entire assessment cycle. Capabilities that once required purpose-built systems are now available to individual faculty through general-purpose tools, democratizing AI affordances in education. Together, the design approaches and affordances identified suggest that responding to GenAI

responsibly means reimagining both what is assessed and how that assessment is supported, rather than treating these as separate problems to solve in isolation.

References

Crompton, H., & Burke, D. (in press). Higher education assessments in the age of GenAI: A systematic review. *International Journal of Educational Technology in Higher Education*.