US JOURNAL OF NEW INSIGHTS IN TECH & EDUCATION

Vol. 4, No. 2 · Article ID: USJNITE-2403 · Journal Homepage: www.edtechinsights.org

Intelligent Digital Assessment in Higher Education: Examining the Role of Automated Feedback in Improving Students' Learning Progress

Md. Zahidul Hasan Zian [1], Md Rasel Ul Alam [2], Md Shahadat [3]

[1] BRAC University, Dhaka Bangladesh

[2] University of the Cumberlands, Kentucky, USA

[3] World University of Bangladesh, Dhaka-1230

**Corresponding author: zian.businessaffairs@gmail.com*

Received: 3 July 2024; Revised: 12 September 2024; Accepted: 20 October 2024; Published: 29 November 2024

KEYWORDS

Intelligent assessment
Automated feedback
Digital assessment
Higher education
Formative assessment
Personalized feedback
Learning progress

Abstract

Intelligent digital assessment is increasingly used to provide rapid, scalable, and data-informed feedback in higher education. Yet the educational value of automated feedback depends not only on delivery speed but also on its specificity, personalization, credibility, and students' willingness to act on it. This study synthesizes evidence on how automated assessment and feedback contribute to learning progress and examines the mechanisms through which feedback supports revision, self-regulation, and subsequent performance. The synthesis focuses on automated assessment, automated feedback, personalized feedback, learning analytics, writing evaluation, programming assessment, and formative learning. The evidence indicates that automated feedback most consistently supports immediate error identification, revision, task performance, engagement with assessment activities, and reduced feedback delays. However, outcomes vary according to feedback specificity, contextualization, student trust, task alignment, and opportunities for revision. Automated systems appear particularly effective when integrated into a continuous formative cycle rather than used solely as scoring tools. Human instructor involvement remains important for higher-order explanation, contextual judgment, motivation, and feedback literacy. Based on these findings, the study proposes a continuous intelligent-assessment framework linking assessment data, automated diagnosis, personalized feedback, student action, reassessment, and learning progress. The findings suggest that effective digital assessment is best achieved by combining automated precision and immediacy with human pedagogical judgment.

1. Introduction

Assessment is not simply a mechanism for assigning grades; it is also a source of information that can guide learners toward improved performance. In digitally mediated higher education, assessment data can be collected continuously through learning management systems, online quizzes, writing platforms, programming environments, knowledge-mapping tools, and other assessment interfaces. This creates an opportunity to move from delayed, episodic feedback toward feedback that is embedded within the learning process. The growing literature on automated assessment indicates that digital systems can score student work rapidly and, in many contexts, attach diagnostic messages that are available immediately after submission (Cavalcanti et al., 2021; Deeva et al., 2021).

The central educational problem, however, is not whether feedback can be automated. The more consequential question is whether automated feedback helps students make meaningful progress. A technically accurate score can still have limited educational value when it is too generic, poorly aligned with the task, difficult to interpret, or disconnected from an opportunity for revision. Conversely, contextualized feedback can support revision and learning, as demonstrated in research on automated formative assessment of scientific argument writing (Zhu et al., 2020). This distinction shifts the research problem from automation as a technical capability to automation as a pedagogical process.

Recent review evidence reinforces this distinction. Cavalcanti et al. (2021) reported that a substantial share of studies showed positive effects of automatic feedback on student performance, while evidence for reducing instructor workload was less consistent. Deeva et al. (2021), after reviewing 109 papers, showed that automated feedback systems vary considerably according to the data used, feedback generation method, feedback target, and the way learners interact with the system. Maier and Klotz (2022) further showed that personalized feedback operates at multiple levels of digital learning environments and that educational effects are not uniform across implementations. Banihashem et al. (2022) similarly concluded that learning analytics can enhance feedback practices, but that the analytic method must be guided by the feedback objective.

Higher education therefore faces a dual challenge. First, institutions need assessment systems capable of producing timely and actionable information at scale. Second, these systems must be designed around how students interpret and use feedback. Studies of automated writing evaluation show that acceptance depends on perceived trust, usefulness, self-efficacy, facilitating conditions, and the cognitive quality of the feedback itself (Zhai & Ma, 2021). Ranalli (2021) likewise found that engagement with automated writing feedback is shaped by contextual and individual factors, including trust. Such findings indicate that automated feedback is best understood as an interaction among technology, assessment design, learner cognition, and instructional context.

This paper addresses that interaction by examining the role of intelligent digital assessment in improving students' learning progress. The focus is deliberately broader than automated marking. The review includes automated writing evaluation, automated programming assessment, personalized feedback systems, knowledge-map assessment, learning-analytics-supported feedback, and broader AI-enabled assessment approaches. Across these areas, the paper asks how automated feedback contributes to learning, which feedback characteristics matter most, why some systems produce stronger outcomes than others, and how human instructors can complement rather than compete with automated systems.

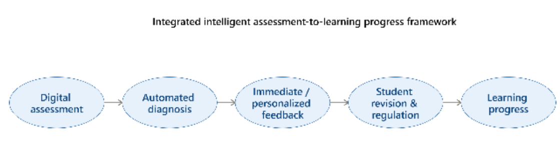

Figure 1. Integrated intelligent assessment-to-learning progress framework.

1.1 Research problem and motivation

The literature reveals a persistent gap between technical automation and pedagogical effectiveness. Many systems are designed to detect errors, assign scores, or compare responses against predefined expectations. These functions are valuable, but they do not automatically ensure that students understand why an answer is weak or what they should do next. The strongest studies in the corpus therefore examine not only score generation but also revision behavior, engagement, acceptance, and learning gains (Zhu et al., 2020; Ranalli, 2021; Thi & Nikolov, 2022).

A second problem concerns scale. Large classes make individual feedback difficult to provide frequently. Automated systems can reduce the time between submission and information delivery, which is especially valuable in programming and engineering where students benefit from rapid cycles of attempt, diagnosis, and correction (Beneroso & Robinson, 2021; Zampirolli, 2021; Insa et al., 2021). Nevertheless, the efficiency argument is incomplete unless faster feedback is educationally useful. The aim should therefore be feedback at scale without sacrificing relevance, clarity, or opportunities for learning.

1.2 Research questions

- RQ1. What forms of intelligent digital assessment and automated feedback are most commonly used in higher education?
- RQ2. Which feedback characteristics are associated with stronger evidence of student learning progress?
- RQ3. How does automated feedback influence revision, engagement, self-regulation, and subsequent performance?
- RQ4. What barriers limit the educational effectiveness of automated feedback systems?
- RQ5. How can automated and human feedback be combined to support continuous formative assessment?

2. Literature Review

2.1 From automated scoring to intelligent assessment

Automated assessment initially developed around relatively constrained tasks in which correctness could be evaluated programmatically. Programming environments provide a clear example: student code can be executed against tests and assigned immediate results. Over time, these systems have expanded toward richer diagnostic feedback and more learner-facing features. Zampirolli (2021) reported large-scale experience with automated programming assessment, while Polito and Temperini (2021) demonstrated how automated testing could be integrated with gamification to support engagement.

Delgado-Pérez and Medina-Bulo (2020) emphasized customization and scalability, showing the practical importance of automated assessment infrastructure.

Intelligent digital assessment extends this logic by using data about the learner, the task, and the response to generate feedback that is potentially more informative than a binary correct/incorrect result. Deeva et al. (2021) proposed a classification framework for automated feedback systems that distinguishes dimensions such as the input data, analytic mechanism, feedback output, and feedback target. This shift is conceptually important because intelligence in assessment should not be equated solely with algorithmic complexity; it should be evaluated by whether the system can provide information that supports meaningful learning decisions.

2.2 Automated feedback as formative assessment

Formative assessment depends on information that can be used before learning opportunities have ended. Automated feedback can strengthen this function by shortening the feedback loop. Cavalcanti et al. (2021) found that many studies reported improved performance when automatic feedback was incorporated into online learning environments. Zhu et al. (2020) found that student revisions following automated feedback were positively associated with score increases and that contextualized feedback was more effective than generic feedback. These findings point to a key principle: the educational value of automation lies in supporting action, not merely in producing output.

The same principle appears in programming education. Automated assessment systems can test code, identify failures, and provide results while students still have time to revise. Insa et al. (2021) noted that continuous assessment is pedagogically valuable because immediate feedback can be integrated into learning rather than delayed until after a course activity has ended. Beneroso and Robinson (2021) similarly described personalized formative feedback as a "stage-gate" in understanding, with large-scale feedback delivered quickly enough to support ongoing coursework.

2.3 Personalized and adaptive feedback

Personalization is a second defining feature of intelligent feedback. Rather than sending the same comment to every student who makes the same type of error, personalized systems use characteristics of the learner, task, or response to modify the message. Maier and Klotz (2022) reviewed personalized feedback systems across digital learning environments and classified implementations at micro, meso, and macro levels. Their review suggests that personalization can operate on individual responses, sequences of learning activities, or broader learning trajectories.

The evidence also indicates that personalization must be purposeful. Feedback can become overwhelming when too many issues are presented simultaneously, particularly when students lack the assessment literacy needed to prioritize revisions. Zhai and Ma (2021) demonstrated that students' acceptance of automated writing evaluation depends on trust, perceived usefulness, facilitating conditions, self-efficacy, and cognitive feedback. Personalized feedback is therefore not simply a technical feature; it is a communication design problem that must consider how learners interpret and prioritize information.

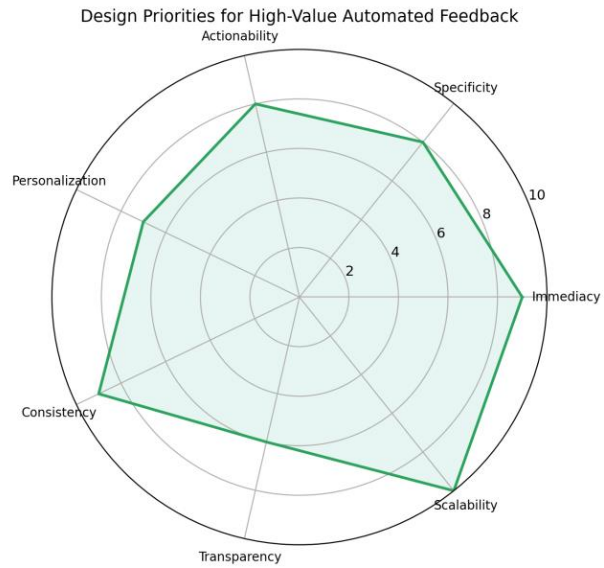

Figure 2. Design priorities for high-value automated feedback.

2.4 Automated writing evaluation and revision

Automated writing evaluation (AWE) represents one of the most studied forms of intelligent feedback because natural-language processing enables systems to assess grammar, vocabulary, mechanics, organization, and aspects of content. Zhang and Hyland (2018) showed that students engage with teacher and automated feedback in different ways, while Lee (2020) examined the longer-term contribution of automated writing evaluation feedback to writing development. Zhai and Ma (2021) shifted attention toward student acceptance, showing that trust and perceived cognitive value matter for continued use. Ranalli (2021) further emphasized that engagement with automated feedback is not automatic; learners may accept, question, or ignore suggestions depending on context and prior beliefs.

The complementarity of human and automated feedback is particularly evident in Thi and Nikolov (2022). Their comparison of teacher and Grammarly feedback suggests that automated tools can complement rather than replace teacher responses, particularly when automation addresses recurring language-level issues while teachers focus on content, organization, and higher-order concerns. Link et al. (2022) likewise examined how AWE affects teacher feedback, student revision, and writing improvement, reinforcing the view that automated feedback is most useful when embedded into an instructional process.

2.5 Automated assessment in programming and STEM

Programming education provides a comparatively mature environment for automated assessment because solutions can be executed and tested. Gordillo (2019), Daradoumis et al. (2019), Marquès Puig et al. (2020), Delgado-Pérez and Medina-Bulo (2020), Zampirolli (2021), Polito and Temperini (2021), and Insa et al. (2021) all illustrate variations of this approach. Across these studies, major advantages include immediate results, repeatable evaluation, scalable checking, and opportunities for iterative practice.

However, automated programming feedback also demonstrates the limits of pure correctness signals. A student may receive a failing test result without understanding the underlying concept. Consequently, the systems that provide meaningful feedback, diagnostics, logs, progress indicators, or gamified representations have greater pedagogical potential than systems that merely return a score. The DSLab work of Marquès Puig et al. (2020) illustrates how automated assessment can be situated within a richer interactive environment, while Polito and Temperini (2021) show how real-time automated assessment can be linked to learner engagement and motivation.

2.6 Knowledge mapping and nontraditional automated assessment

Automated assessment is not limited to writing and programming. Ho et al. (2018) developed an online knowledge-mapping assessment tool that automatically assessed student-submitted concept maps in medical education. Their results indicated a significant relationship between map scores and performance on modified essay questions, while students viewed the tool positively. This example expands the definition of intelligent assessment because the system evaluates representations of conceptual knowledge rather than only textual correctness or executable code.

2.7 Learning analytics and feedback practices

Learning analytics adds another layer of intelligence by using patterns across student activity rather than solely analyzing a single response. Banihashem et al. (2022) reviewed the role of learning analytics in enhancing feedback practices in higher education and emphasized that the type of data, analytic method, feedback objective, and stakeholder need should be aligned. Information visualization was one of the most common approaches, suggesting that intelligent assessment may involve not only automated comments but also dashboards and analytics that help learners and instructors interpret progress over time.

Hooda et al. (2022) further connected artificial intelligence with assessment and feedback in higher education and argued for data-driven approaches that can support more timely and constructive assessment processes. The broader implication is that intelligent assessment can be viewed as a continuum: from automated scoring, to automated feedback, to analytics-supported diagnosis, to adaptive feedback that is informed by the learner's history and learning goals.

2.8 Student acceptance, trust, and feedback literacy

Technological accuracy does not guarantee student engagement. Automated feedback is mediated by learners' perceptions of credibility and relevance. Akçapınar (2015), for example, showed that automated feedback based on text mining could alter plagiarism behavior, demonstrating that automated feedback can shape student conduct as well as learning outcomes. Zhai and Ma (2021) found that perceived trust and cognitive feedback contributed to acceptance of AWE, while Ranalli (2021) highlighted the importance of trust in actual engagement with automated feedback. These studies suggest that feedback literacy—the ability to interpret, evaluate, and use feedback—is a central condition for intelligent assessment to produce educational value.

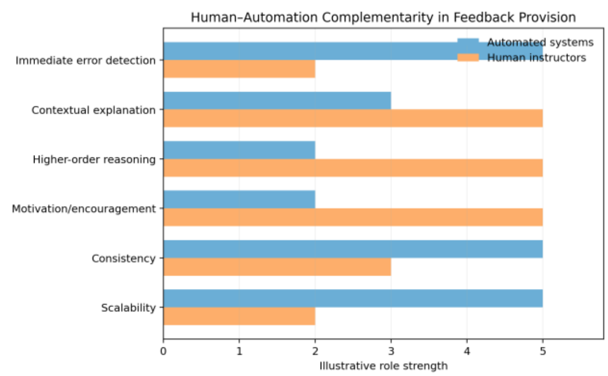

Figure 3. Human–automation complementarity in feedback provision.

2.9 Research gap

Taken together, the literature shows that the field has moved beyond a simple question of whether automated feedback works. Reviews now document a heterogeneous ecosystem of feedback technologies, while empirical studies demonstrate effects on revision, performance, engagement, and acceptance. Yet three gaps remain. First, many studies evaluate immediate task outcomes rather than longitudinal learning progress. Second, evidence is concentrated in constrained assessment contexts such as writing and programming, limiting generalizability. Third, fewer studies conceptualize automated feedback as one stage in an ongoing cycle that connects assessment, diagnosis, student action, reassessment, and subsequent learning. The present synthesis addresses these gaps by framing automated feedback as a continuous learning-progress mechanism rather than an isolated technological feature.

3. Conceptual Framework

The conceptual model proposed in this paper integrates insights from automated feedback, formative assessment, personalized learning, learning analytics, and human–AI complementarity. The model assumes that intelligent assessment begins with evidence produced by a learner's response or activity. The system then applies an analytic procedure to identify errors, misconceptions, patterns, or performance levels. The resulting diagnostic information is transformed into feedback that varies in immediacy, specificity, contextualization, and personalization. Students interpret the feedback, decide whether it is credible and useful, and take action through revision, practice, reflection, or strategy change. Finally, the learner encounters a subsequent assessment opportunity, allowing the cycle to continue.

The framework therefore distinguishes between five proximal mechanisms: diagnostic accuracy, feedback quality, student interpretation, student action, and reassessment. Contextual conditions—including instructor mediation, feedback literacy, task design, data quality, and transparency—shape the strength of each link. This framing addresses a recurring limitation in the literature: a system can be accurate at detecting an error but educationally weak if its feedback cannot be understood or acted upon.

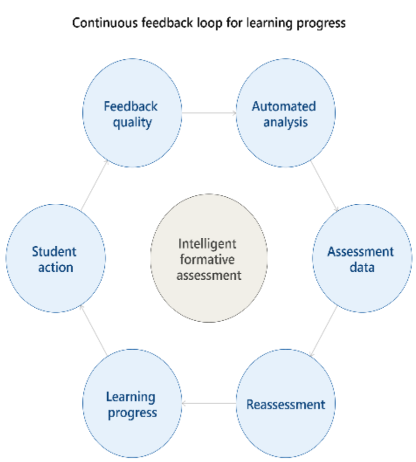

Figure 4. Continuous feedback loop for learning progress.

4. Methodology

4.1 Research design

A structured literature review and evidence synthesis design was adopted. The purpose was not to calculate a pooled statistical effect because the reviewed studies differ substantially in assessment type, outcome definition, learner population, technology, and research design. Instead, the review identifies recurrent mechanisms, design characteristics, reported outcomes, and implementation barriers across 30 studies published between 2015 and 2022.

4.2 Search and selection logic

The evidence corpus was assembled from scholarly sources addressing automated assessment, automatic feedback, personalized feedback, automated writing evaluation, programming assessment, digital assessment, learning analytics, and AI-supported feedback. Studies were prioritized when an openly accessible full text or repository version was available. The publication window was deliberately closed at 2022 so that the synthesis represents evidence available before 2023.

4.3 Inclusion criteria

- Published in 2022 or earlier.
- Focused on assessment, automated feedback, personalized feedback, learning analytics, or closely related intelligent assessment functions.
- Included higher education directly or addressed mechanisms transferable to higher-education digital assessment.
- Contained empirical, review, design, or evaluation evidence sufficient for thematic coding.
- Provided an openly accessible full-text version or openly accessible scholarly record where possible.

4.4 Coding framework

Each study was coded across the following dimensions: publication year, assessment context, feedback type, degree of automation, personalization, feedback timing, student response, reported learning outcome, human instructor role, and implementation barriers. The coding was used to build descriptive figures and cross-study thematic comparisons. Because the corpus includes both empirical and review studies, numerical values in figures represent counts or author-coded emphasis across the literature rather than pooled effect sizes.

Table 1. Cross-study coding framework used for the evidence synthesis.

Dimension	Coding categories	Purpose
Assessment context	Writing/language; programming; STEM/engineering; medicine; general online	Identify where intelligent assessment is concentrated
Feedback mode	Immediate; personalized; analytics-mediated;	Compare technological approaches

	automated writing; automated coding	
Learner action	Revision; retry; reflection; practice; acceptance; monitoring	Connect feedback to learning behavior
Outcome	Performance; revision quality; engagement; self-regulation; efficiency; acceptance	Identify reported educational effects
Human role	None/minimal; complementary; central mediation	Examine human–automation division of labor
Barrier	Trust; genericity; overload; technical limitations; alignment	Identify implementation risks

4.5 Reliability and interpretation

The review treats numerical coding as descriptive rather than inferential. Counts indicate how often a theme was present or emphasized in the selected corpus; they do not imply statistical relationships between variables. Where the evidence is mixed or context-dependent, the synthesis reports the distinction rather than forcing a single positive or negative conclusion.

5. Results

5.1 Profile of the evidence base

The 30-study corpus spans 2015–2022 and shows a marked growth in research during the later part of the period. The number of studies increases sharply from 2020 onward, corresponding with the expansion of AI-supported assessment, automated writing evaluation, scalable programming assessment, and learning-analytics-supported feedback. The evidence base is therefore both interdisciplinary and increasingly oriented toward intelligent, continuous, and personalized assessment.

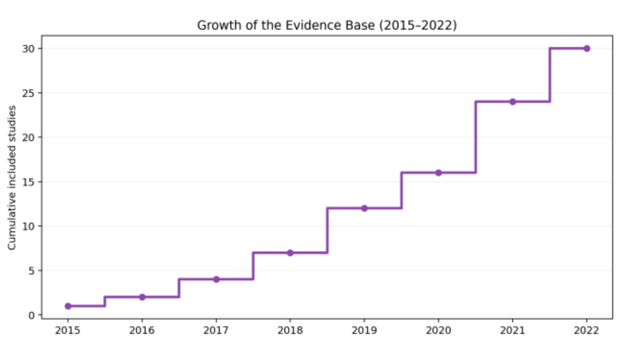

Figure 5. Growth of the included evidence base, 2015–2022.

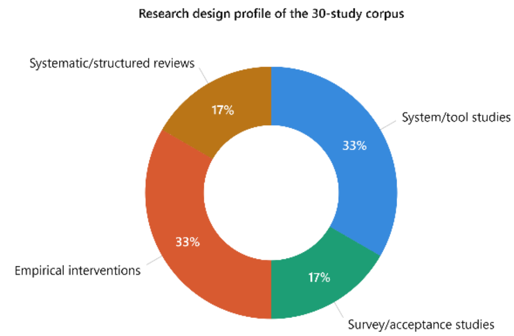

Figure 6. Research design profile of the 30-study corpus.

5.2 Assessment contexts

Programming and computer science form the largest contextual cluster, reflecting the relative ease with which software-based tasks can be automatically tested and graded. Writing and language assessment form the second major cluster, where natural-language processing enables grammar, mechanics, and selected discourse features to be evaluated. Engineering and STEM studies demonstrate another important use case: automated tools can process structured numerical work and return personalized formative information at a scale difficult to achieve through manual feedback alone. General digital assessment, medicine, and other contexts broaden the evidence base beyond these two dominant areas.

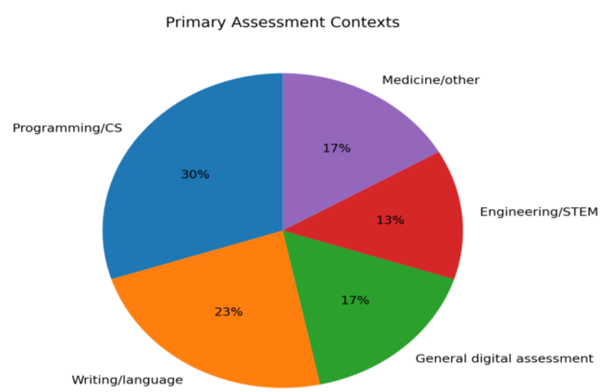

Figure 7. Primary assessment contexts represented in the reviewed studies.

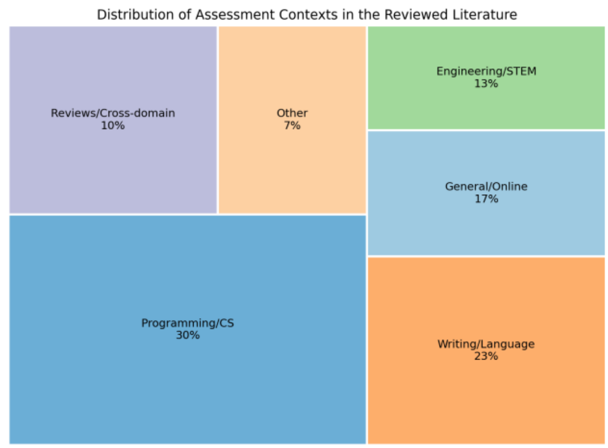

Figure 8. Context distribution of the reviewed literature.

5.3 Reported learning outcomes

Across the evidence base, the most consistently reported benefits concern speed of feedback, task performance, revision support, and student engagement with assessment activities. Cavalcanti et al. (2021) reported that 65.07% of reviewed studies found evidence that automatic feedback increased performance in activities. Zhu et al. (2020) found that revision after automated feedback was positively related to score increases, and contextualized feedback was more effective than generic feedback. Beneroso and Robinson (2021) documented substantial improvements after personalized formative feedback in engineering coursework. The evidence therefore supports a meaningful learning benefit, but it does not imply that every automated feedback system improves outcomes equally.

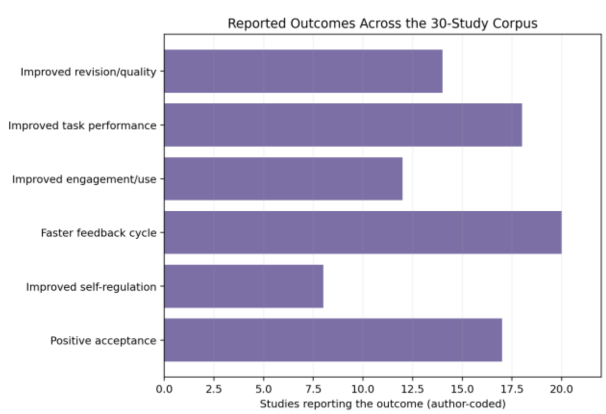

Figure 9. Reported outcomes across the 30-study corpus.

5.4 Feedback quality characteristics

The coding suggests seven recurring design priorities: immediacy, specificity, actionability, personalization, consistency, transparency, and scalability. Immediacy and scalability are the strongest technical advantages because automation can substantially shorten feedback turnaround and handle large numbers of submissions. Specificity and actionability are more closely tied to learning because they determine whether students can identify what to change. Personalization becomes increasingly important when learners differ in prior knowledge or make distinct error patterns. Transparency and trust remain necessary because students must understand why a system has produced a recommendation before they decide to act on it.

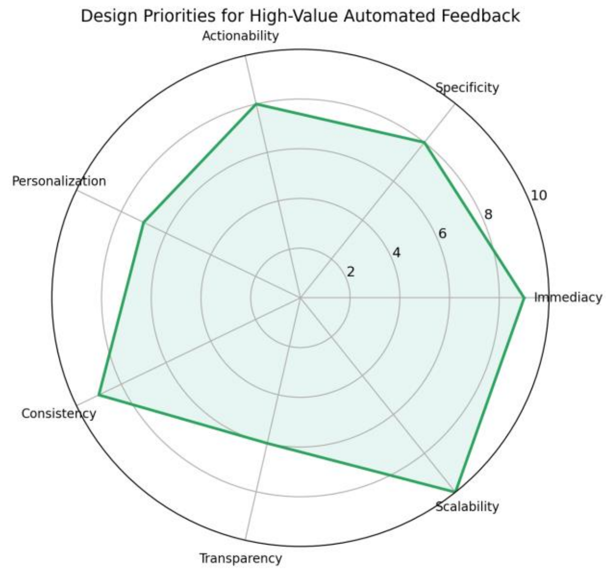

Figure 10. Design priorities for high-value automated feedback.

5.5 Automated feedback as a learning cycle

A second major finding is that feedback produces value through a cycle of action. The strongest studies do not end the intervention at score generation. Instead, they provide feedback, allow revision or repeated attempts, and evaluate the subsequent response. Zhu et al. (2020), for example, examined how students revised scientific arguments after automated feedback. Programming systems similarly create rapid loops in which students submit, receive automated diagnostics, revise code, and resubmit. This repeated cycle is more compatible with formative assessment than one-off automated grading.

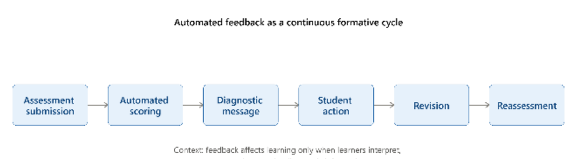

Figure 11. Automated feedback as a continuous formative cycle.

5.6 Reported barriers

Despite the benefits, the literature identifies recurring limitations. Generic feedback remains a major concern because it may tell students that an error exists without explaining the reasoning needed to correct it. Trust and credibility are also important, particularly in automated writing evaluation. Some students may question suggestions or avoid using them if they do not understand how the system generated the feedback (Ranalli, 2021;

Zhai & Ma, 2021). Other barriers include feedback overload, mismatch between automated criteria and task goals, technical limitations, and uncertainty about how automation affects instructor workload. Cavalcanti et al. (2021) explicitly found that evidence for workload reduction was less consistent than evidence for performance benefits.

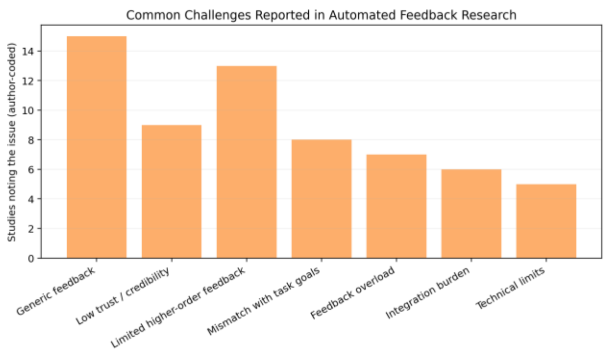

Figure 12. Common challenges reported in automated feedback research.

5.7 Learning-progress pathway

The synthesis supports a learning-progress pathway in which initial performance generates diagnostic evidence, automated feedback identifies a target for improvement, student revision or practice changes the next response, and reassessment provides evidence of progress. This pathway is consistent with formative feedback theory and with empirical evidence from writing, programming, and engineering studies. It also highlights the difference between immediate task improvement and durable learning: a better second attempt is an important indicator, but learning progress is stronger when improvements transfer to subsequent tasks.

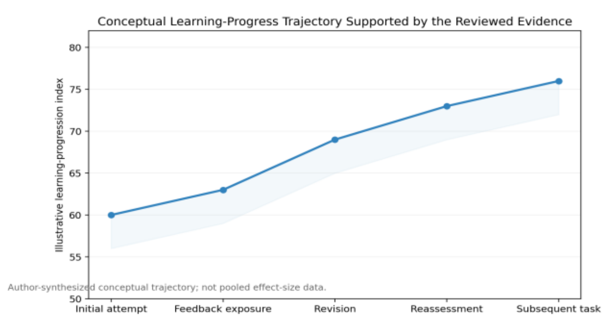

Figure 13. Conceptual learning-progress trajectory supported by the reviewed evidence.

5.8 Thematic co-occurrence

Thematic coding reveals that automated scoring is frequently discussed together with immediate feedback, while personalized feedback is more closely associated with adaptivity, actionability, and learner interpretation. Revision and learning gains are more strongly associated with specific and actionable feedback than with automation alone. Instructor mediation shows a particularly strong relationship with trust-sensitive elements. These co-occurrences suggest that technical sophistication and pedagogical effectiveness are related but distinct constructs.

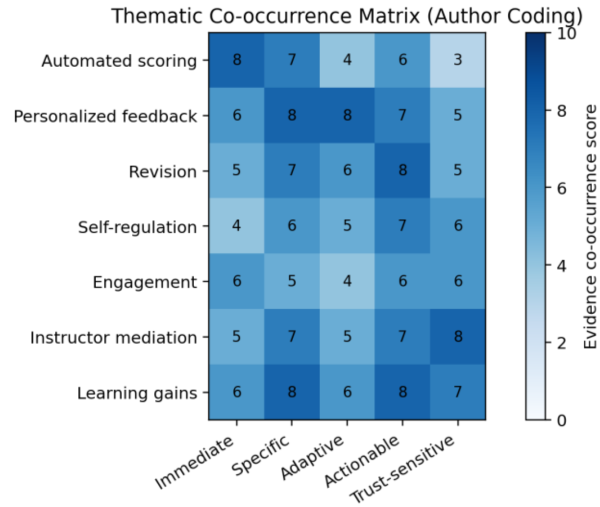

Figure 14. Thematic co-occurrence matrix across the literature.

5.9 Value propositions of intelligent assessment

The literature identifies six major value propositions: immediate diagnosis, personalized guidance, revision support, mastery monitoring, partial support for instructor workload, and increased learner autonomy. Immediate diagnosis and revision support are especially consistent across programming, writing, and STEM contexts. Mastery monitoring becomes more visible in systems that incorporate learning analytics or gamification. Autonomy is supported when students can use feedback without waiting for an instructor response, although autonomy depends on sufficient feedback literacy.

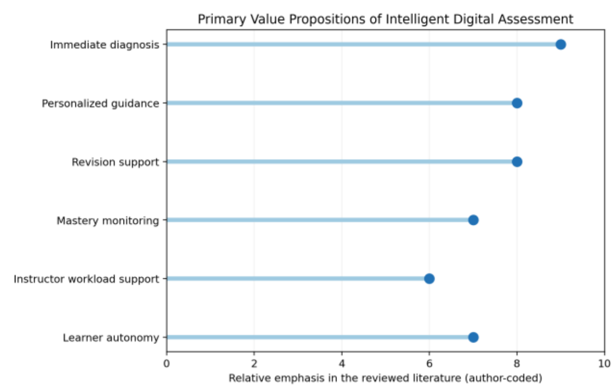

Figure 15. Primary value propositions of intelligent digital assessment.

5.10 Mechanisms linking assessment to learning progress

The evidence suggests several interlocking mechanisms. Automated assessment produces evidence; analytics convert evidence into diagnosis; feedback translates diagnosis into guidance; students interpret and trust that guidance; revision or practice changes behavior;

reassessment determines whether the change is retained or extended. Human instructors can strengthen the process by clarifying ambiguous feedback, addressing higher-order issues, and supporting motivation. Personalization and transparency improve the fit between the message and the learner's needs, while feedback literacy determines whether students can convert information into action.

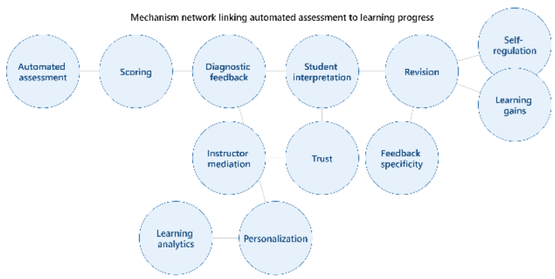

Figure 16. Mechanism network linking automated assessment to learning progress.

5.11 Human–AI complementarity

The reviewed evidence does not support a simple replacement model in which automated feedback displaces instructors. Instead, the strongest conceptual pattern is complementarity. Automated systems are well suited to repeated, high-frequency, rule-based, and scalable operations such as error detection and initial diagnostics. Instructors remain comparatively important for contextual explanations, higher-order reasoning, motivation, and the interpretation of complex disciplinary work. Thi and Nikolov (2022) illustrate this complementarity in writing feedback, while Banihashem et al. (2022) emphasize that analytics should be aligned with feedback objectives and stakeholder needs.

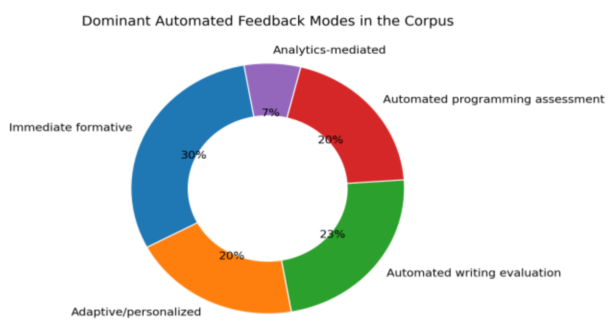

Figure 17. Dominant automated feedback modes in the reviewed corpus.

6. Discussion

6.1 Automated feedback works best as part of a process

The most important conclusion from the synthesis is that automated feedback should be treated as a process rather than a feature. Systems that only grade are useful for efficiency, but systems that grade, explain, support revision, and reassess create a more complete formative cycle. This interpretation is consistent with Cavalcanti et al. (2021), Deeva et al. (2021), and Banihashem et al. (2022), all of whom emphasize that educational value depends on how automated systems are configured and used.

6.2 Specificity and contextualization matter more than automation alone

The evidence repeatedly distinguishes generic messages from context-sensitive guidance. Zhu et al. (2020) found contextualized automated feedback more effective than generic feedback, while Beneroso and Robinson (2021) showed how personalized feedback can be embedded directly into engineering problem solving. The implication for system design is clear: an intelligent assessment system should be judged not only by whether it detects an error but also by whether it explains the error in a way that connects to the learner's current task and knowledge state.

6.3 Student engagement is a mediating mechanism

Automated feedback is only useful when students attend to it and act on it. Ranalli (2021) showed that trust influences engagement with automated feedback, and Zhai and Ma (2021) identified perceived usefulness, facilitating conditions, self-efficacy, and cognitive feedback as important determinants of acceptance. This suggests a mediated pathway in which the quality of system output affects student perceptions, which in turn affect the likelihood of revision and practice.

6.4 Personalization should be purposeful

Personalization is promising but not automatically beneficial. Maier and Klotz (2022) showed that personalization occurs at different levels of digital environments, while the broader literature suggests that additional information can become counterproductive when students receive too many corrections or when the logic of personalization is opaque. Personalization should therefore prioritize the next most useful learning action rather than attempting to explain every possible weakness simultaneously.

6.5 Human mediation remains important

The literature suggests that intelligent assessment is most defensible when automation and instructor expertise are integrated. Automation can absorb high-volume, repetitive diagnostic work, allowing instructors to focus more on explanation, synthesis, mentoring, and complex evaluation. However, this does not imply that automation necessarily reduces workload in every setting. Cavalcanti et al. (2021) found inconsistent evidence regarding workload reduction, which means that implementation costs, system maintenance, interpretation of analytics, and student support must be considered alongside grading efficiency.

6.6 Implications for higher-education design

Higher-education institutions should design assessment systems around feedback use, not only scoring accuracy. Tasks should permit revision where possible, feedback should be aligned with learning outcomes, and students should be taught how to interpret automated suggestions. Dashboards and analytics should emphasize actionable patterns rather than excessive data. Instructors should receive information that helps them identify where human intervention adds value. These design principles shift the purpose of intelligent assessment from replacing manual marking to creating a more responsive learning environment.

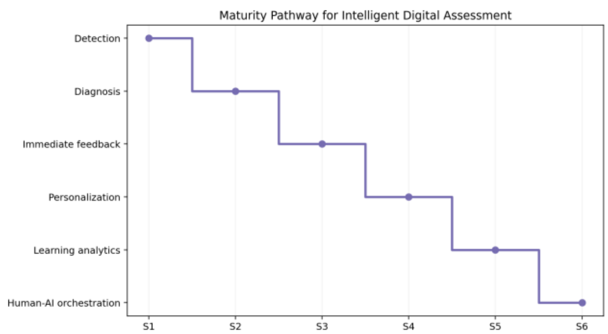

Figure 18. Maturity pathway for intelligent digital assessment.

7. Practical Implications

- For instructors: use automated systems for rapid, recurring diagnostics and reserve human feedback for explanation, disciplinary judgment, higher-order reasoning, and motivation.
- For institutions: invest in assessment infrastructure that supports iterative attempts, revision, and progress monitoring rather than single-attempt scoring alone.
- For system developers: prioritize contextualized, actionable, transparent, and appropriately personalized feedback.
- For students: develop feedback literacy so that automated recommendations are interpreted critically rather than accepted or rejected automatically.
- For quality assurance: evaluate systems using both technical accuracy and learning outcomes, including revision behavior and subsequent performance.

8. Limitations

This synthesis has several limitations. First, the 30-study corpus is purposive and focused on studies that were accessible through open sources or accessible scholarly records; it is therefore not a complete census of all publications. Second, the studies use heterogeneous methods and outcome measures, which prevents calculation of a single pooled effect. Third, the evidence is concentrated in programming, writing, STEM, and selected professional contexts, so generalization to all higher-education disciplines should be cautious. Fourth, author-coded counts describe recurrence of themes and should not be interpreted as effect sizes. Finally, the publication cutoff ends at 2022, which intentionally excludes later developments in generative AI assessment and newer forms of automated feedback.

9. Conclusion

Intelligent digital assessment has the potential to change higher education by shortening the feedback loop, scaling diagnostic support, and enabling more frequent opportunities for revision. The evidence reviewed in this paper indicates that automated feedback can improve task performance, support revision, increase engagement with assessment activities, and reduce delays between student action and instructional information. Yet the results also demonstrate that automation itself is not the mechanism of learning. Learning progress emerges when students receive useful information, understand it, trust it, act on it, and encounter a subsequent opportunity to test the change.

The most promising model is therefore a continuous human–AI assessment cycle. Automated systems provide speed, consistency, and scalable diagnosis. Personalized feedback improves the match between the message and the learner's needs. Learning analytics can identify patterns that are invisible within a single response. Human instructors provide context, higher-order judgment, motivation, and interpretive support. When these functions are connected, intelligent assessment can move from automated marking toward continuous formative learning. The central implication is that future assessment design should be evaluated not by how intelligently a system scores work, but by how effectively it helps students make and sustain progress.

References

Ai, H. (2017). Providing graduated corrective feedback in an intelligent computer-assisted language learning environment. ReCALL, 29(3), 313–334. https://doi.org/10.1017/S095834401700012X

Akçapınar, G. (2015). How automated feedback through text mining changes plagiaristic behavior in online assignments. Computers & Education, 87, 123–130. https://doi.org/10.1016/j.compedu.2015.04.007

Banihashem, S. K., Noroozi, O., van Ginkel, S., Macfadyen, L. P., & Biemans, H. J. A. (2022). A systematic review of the role of learning analytics in enhancing feedback practices in higher education. Educational Research Review, 37, Article 100489. https://doi.org/10.1016/j.edurev.2022.100489

Beneroso, D., & Robinson, J. P. (2021). A tool for assessing and providing personalised formative feedback at scale within a second in engineering courses. Education for Chemical Engineers, 36, 38–45. https://doi.org/10.1016/j.ece.2021.02.002

Bey, A., Jermann, P., & Dillenbourg, P. (2018). A comparison between two automatic assessment approaches for programming: An empirical study on MOOCs. Educational Technology & Society, 21(2), 259–272.

Cavalcanti, A. P., Barbosa, A., Carvalho, R., Freitas, F., Tsai, Y.-S., Gašević, D., & Mello, R. F. (2021). Automatic feedback in online learning environments: A systematic literature review. Computers and Education: Artificial Intelligence, 2, Article 100027. https://doi.org/10.1016/j.caeai.2021.100027

Cheng, K. S. (2017). The impact of online automated feedback on students' reflective journal writing in an EFL course. The Internet and Higher Education, 34, 18–27. https://doi.org/10.1016/j.iheduc.2017.04.002

Cheniti-Belcadhi, L. (2016). Personalized feedback for self assessment in lifelong learning environments based on semantic web. Computers in Human Behavior, 55, 562–570. https://doi.org/10.1016/j.chb.2015.07.042

Daradoumis, T., Marquès Puig, J. M., Arguedas, M., & Calvet Liñan, L. (2019). Analyzing students' perceptions to improve the design of an automated assessment tool in online distributed programming. Computers & Education, 128, 159–170. https://doi.org/10.1016/j.compedu.2018.09.021

Deeva, G., Bogdanova, D., Serral, E., Snoeck, M., & De Weerdt, J. (2021). A review of automated feedback systems for learners: Classification framework, challenges and opportunities. Computers & Education, 162, Article 104094. https://doi.org/10.1016/j.compedu.2020.104094

Delgado-Pérez, P., & Medina-Bulo, I. (2020). Customizable and scalable automated assessment of C/C++ programming assignments. Computer Applications in Engineering Education, 28(6), 1449–1466. https://doi.org/10.1002/cae.22317

Galan, D., Heradio, R., Vargas, H., Abad, I., & Cerrada, J. A. (2019). Automated assessment of computer programming practices: The 8-years UNED experience. IEEE Access, 7, 130113–130119. https://doi.org/10.1109/ACCESS.2019.2938391

Gordillo, A. (2019). Effect of an instructor-centered tool for automatic assessment of programming assignments on students' perceptions and performance. Sustainability, 11(20), Article 5568. https://doi.org/10.3390/su11205568

Hassan, M. A., Habiba, U., Khalid, H., Shoaib, M., & Arshad, S. (2019). An adaptive feedback system to improve student performance based on collaborative behavior. IEEE Access, 7, 107171–107178. https://doi.org/10.1109/ACCESS.2019.2931565

Ho, V. W., Harris, P. G., Kumar, R. K., & Velan, G. M. (2018). Knowledge maps: A tool for online assessment with automated feedback. Medical Education Online, 23(1), Article 1457394. https://doi.org/10.1080/10872981.2018.1457394

Hooda, M., Rana, C., Dahiya, O., Rizwan, A., & Hossain, M. S. (2022). Artificial intelligence for assessment and feedback to enhance student success in higher education. Mathematical Problems in Engineering, 2022, Article 5215722. https://doi.org/10.1155/2022/5215722

Insa Cabrera, D., Pérez-Rubio, S., Silva, J., & Tamarit Muñoz, S. (2021). Semiautomatic generation and assessment of Java exercises in engineering education. Computer Applications in Engineering Education, 29(5), 1034–1050. https://doi.org/10.1002/cae.22356

Khan, S., & Khan, R. A. (2019). Online assessments: Exploring perspectives of university students. Education and Information Technologies, 24, 661–677. https://doi.org/10.1007/s10639-018-9797-0

Lee, Y.-J. (2020). The long-term effect of automated writing evaluation feedback on writing development. English Teaching, 75(1), 67–92.

Link, S., Mehrzad, M., & Rahimi, M. (2022). Impact of automated writing evaluation on teacher feedback, student revision, and writing improvement. Computer Assisted Language Learning, 35(4), 605–634. https://doi.org/10.1080/09588221.2020.1743323

Maier, U., & Klotz, C. (2022). Personalized feedback in digital learning environments: Classification framework and literature review. Computers and Education: Artificial Intelligence, 3, Article 100080. https://doi.org/10.1016/j.caeai.2022.100080

Marquès Puig, J. M., Daradoumis, T., Calvet Liñan, L., & Arguedas, M. (2020). Fruitful student interactions and perceived learning improvement in DSLab: A dynamic assessment tool for distributed programming. British Journal of Educational Technology, 51(1), 53–70. https://doi.org/10.1111/bjet.12756

McCarthy, K. S., Roscoe, R. D., Allen, L. K., Likens, A. D., & McNamara, D. S. (2022). Automated writing evaluation: Does spelling and grammar feedback support high-quality writing and revision? Assessing Writing, 52, Article 100608.

Polito, G., & Temperini, M. (2021). A gamified web based system for computer programming learning. Computers and Education: Artificial Intelligence, 2, Article 100029. https://doi.org/10.1016/j.caeai.2021.100029

Ranalli, J. (2021). L2 student engagement with automated feedback on writing: Potential for learning and issues of trust. Journal of Second Language Writing, 52, Article 100816. https://doi.org/10.1016/j.jslw.2021.100816

Thi, N. K. K., & Nikolov, M. (2022). How teacher and Grammarly feedback complement one another in

Myanmar EFL students' writing. The Asia-Pacific Education Researcher, 31, 767–779. https://doi.org/10.1007/s40299-021-00625-2

Zampirolli, F. A. (2021). An experience of automated assessment in a large-scale introduction programming course. Computer Applications in Engineering Education, 29(5), 1284–1299. https://doi.org/10.1002/cae.22385

Zhai, N., & Ma, X. (2021). Automated writing evaluation (AWE) feedback: A systematic investigation of college students' acceptance. Computer Assisted Language Learning, 35(9), 2817–2842. https://doi.org/10.1080/09588221.2021.1897019

Zhang, Z., & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. Assessing Writing, 36, 90–103. https://doi.org/10.1016/j.asw.2018.02.004