ISSN: 2458-8989

Natural and Engineering Sciences

NESciences, 2026, 11 (3): 432-443 doi: 10.28978/nesciences.263036

# Artificial Intelligence Applications in Dialogic Feedback Systems: A Four-Phase Model for Critical Appraisal in Biomedical and Natural Sciences Domains

[1*] Reinaldo Antonio Guerrero-Chirinos, [2]Herlinda Gervacio-Jiménez, [3]Jenny Marleny Bejarano-Huamaní, [4]Fernando Montemayor-Ibarra,[5]Laura Elena Ruiz-Avendaño, [6]Sandra Elizabeth Pozo-Prado

[1]Universidad Técnica Particular de Loja, UTPL, Ecuador. ORCID: https://orcid.org/0000-0003-0499-7453, E-mail: raguerrero12@utpl.edu.ec

[2]Universidad Autónoma de Guerrero, Mexico. ORCID: https://orcid.org/0000-0003-3037-9528, E-mail: lindagervacio@uagro.mx

[3]Universidad Americana de Europa, Spain. ORCID: https://orcid.org/0009-0003-5915-2308, E-mail: Jmarleny@bejaranoh.com

Universidad Autónoma de Nuevo León, UANL, Mexico. ORCID: https://orcid.org/0000-0003-3503-0149, E-mail: fernando.montemayorib@uanl.edu.mx

Universidad Autónoma de Sinaloa, Mexico. ORCID: https://orcid.org/0000-0003-2270-5661, E-mail: laura.ruiza@uas.edu.mx

Unidad Educativa Eugenio Espejo, Ecuador. ORCID: https://orcid.org/0009-0000-9138-0755, E-mail: sandra.pozo@educacion.gob.ec

**Abstract**

Feedback has become plentiful, immediate, and dialogic, and the problem of feedback has shifted to one of eliciting, comparing, evaluating, and applying it: a shift with high stakes for the critical appraisal of evidence in biomedical and natural sciences education. This conceptual article is a theoretical synthesis across three literatures of first-quartile quality: literature on feedback literacy and feedback processes for learning; scholarship on internal feedback, evaluative judgment, and self-regulation; and new research on AI-generated feedback. The article introduces the Critical Appraisal Agency Model (CAAM), which consists of a recursive four-phase model (dialogic elicitation, comparison and internal feedback, critical evaluative judgment and self-regulated enactment), two moderating layers (AI feedback literacy and hybrid regulation of learning), and one enabling condition (pedagogical and curricular design). It is formalized into six testable propositions and a series of conceptual figures, and implications for teaching, institutional policy, and research in biomedical and natural sciences education are discussed.

**Keywords:**
*feedback; generative artificial intelligence; feedback literacy; evaluative judgment; self-regulated learning; critical appraisal; biomedical education; higher education*

**Corresponding Author: Reinaldo Antonio Guerrero-Chirinos, E-mail: raguerrero12@utpl.edu.ec*

Available Online: 01/09/2026

## Introduction

Despite the fact that feedback is among the most salient and effective influences on learning, it is also one of the most variable, with meta-analyses showing both large average effects and high levels of heterogeneity, depending on the type of feedback information provided and how the learner uses it (Wisniewski et al., 2020). This variation has heralded a new paradigm in educational research, shifting the focus from feedback as information sent by the teacher to feedback as a process in which the learner seeks, interprets, and uses information to improve their performance and strategies (Carless, 2022; Henderson et al., 2019).

In this new paradigm, the concept of "feedback literacy" – defined as the understandings, capabilities, and dispositions required to make sense of information and act on it – plays a central role (Carless & Boud, 2018; Molloy et al., 2020). This conceptual shift has been enabled by Generative AI. Currently, drafts can be submitted to large language models (LLMs) via natural-language dialogue and, in seconds, receive comments on drafts, explanations of errors, suggestions for improvement, and alternative examples (Kasneci et al., 2023; Yan, Greiff, et al., 2024). Empirical results show that feedback produced by such systems can meet the standards of human feedback across most metrics, even though it is not yet as good, and that such feedback can boost revision, motivation, and students' positive emotions (Banihashem et al., 2024; Escalante et al., 2023; Meyer et al., 2024; Steiss et al., 2024).

Feedback is no longer a limited resource that the instructor provides but becomes an abundant resource available on demand for students' lab reports, case analyses, and research drafts, for instance, in domains where they regularly need to analyze, compare, and respond to technical feedback on their work. With this abundance, there is a theoretical issue that the literature has so far failed to address satisfactorily. AI and feedback research has largely focused on assessing the quality of the system's output and what the user thinks about it, but has not done much to conceptualize what changes in the learner's role are possible when the source of feedback is a dialogic, fallible, and always-available machine (Dawson et al., 2025; Er et al., 2025).

Experimental research also raises concerns about concrete risks: the use of AI tools can foster metacognitive laziness, which enhances the product but not the learning process (Fan et al., 2025), and learner agency may be diminished when AI tools are used, since learners not only rely on them to complete tasks but may also feel helpless when they are removed (Darvishi et al., 2024). Additionally, dominant AI discourses in higher education often foreground the learner as a recipient of AI-driven services, rather than as an agent (Bearman et al., 2023). There is thus a gap among three literatures that have developed in parallel: research on feedback literacy and new feedback paradigms; research on internal feedback, evaluative judgment, and self-regulation; and research on the hybrid regulation of learning between humans and AI.

This study attempts to synthesize and integrate them by answering the following guiding question: How does generative AI redefine feedback processes, and what capacities and circumstances enable the learner to become the generative agent of their own feedback, specifically in disciplines such as biomedicine and the natural sciences, where critical evaluation of evidence, data, and technical argument is a fundamental skill? The intention is to model this reconfiguration theoretically and redefine the learner's role on the basis of the feedback-literacy tradition initiated by Carless and Boud (2018). The study makes four contributions. First, it categorizes generative feedback as a type of feedback, one that can be connected to generative technology and the learner's internal feedback. Second, it proposes the Critical Appraisal Agency Model (CAAM), which comprises four phases, two layers of moderation, and one condition of enablement. Third, it establishes the model in six propositions that can be tested empirically. Fourth, it presents the model's synthesis as a conceptual set of figures and implications for practice, especially as a way to train critical appraisal skills in the education of biomedical and natural sciences, and outlines a research agenda.

The rest of the article is organized into a Materials and Methods section, a Results and Discussion section that builds the theory and presents the model, a Limitations and scope for future research section, and a conclusion. In particular, this restructuring is relevant to biomedical and natural sciences education, where information is already evaluated in a highly systematic way—reading a laboratory result relative to a hypothesis, a clinical finding relative to prior probability, or a published study relative to methodological criteria. Generative AI can fit into these routines, providing explanations for statistical findings, critiques of experimental design, or comments on differential diagnosis as needed (Chiu et al., 2023; Kasneci et al., 2023). The answer to whether abundance enhances or undermines the appraisal skills these fields have long developed is the same question the model poses to feedback in general: does the student become the one who elicits, compares, judges, and acts, or does the student relinquish that task to the system?

## Materials and Methods

The study is a theory-synthesis conceptual study dedicated to theory building. Theoretical synthesis, following Jaakkola (2020), aims to bring together research streams that have been disconnected to provide a new, more holistic understanding of a phenomenon. It is evaluated based on the clarity of the proposed constructs, the rigor of the theoretical argument, and the ability to spark future research, not solely on the basis of disclosed primary data. The article is integrated into MacInnis's (2011) typology of conceptual contributions as a revision, a reconfiguration (of how the learner's role in feedback is understood), and a delineation (mapping the relationships among constructs in an integrated model). The synthesis architecture differs from multiple integrating lenses and the domain theory (Jaakkola, 2020). Research on learning-centered feedback processes and feedback literacy (Carless & Boud, 2018; Molloy et al., 2020; Nieminen & Carless, 2023) is the domain theory.

Within this domain, three lenses have been integrated: internal feedback as a comparison process (Nicol, 2021), the construct of evaluative judgment (Bearman et al., 2024; Tai et al., 2018), and models of hybrid regulation of learning between humans and AI (Järvelä et al., 2023; Molenaar, 2022a, 2022b). The propositions are supported by empirical evidence on AI-generated feedback, especially in scientific and health professions education when available, and are contrasted with evidence on human-generated feedback. Articles for the literature were carefully and critically selected, with a focus on those published after 2021 in journals ranked in the first quartile of the international journals used for the analysis. In addition, a few seminal papers from earlier years that contain key constructs were also included.

The process of intentional selection is also aligned with the purpose of theory building, which is not to gather as much evidence as possible but to integrate ideas into an argument (Jaakkola, 2020). The criteria proposed for theorizing in the form of propositions and process models, namely explicit definition of constructs, logical consistency between premises and propositions, parsimony, and generative potential for empirical research (Cornelissen, 2017; MacInnis, 2011), were followed to protect the rigor of the conceptual product. The literature on AI and feedback was identified and searched in the Scopus and Web of Science databases using the following combinations of selected descriptors: feedback; feedback literacy; internal feedback; evaluative judgment; self-regulated learning; generative artificial intelligence; large language models, within a timeframe of 2021 to 2025.

The status of the first quartile was validated using the Scimago Journal Rank and Journal Citation Reports available at the time of manuscript preparation, based on each journal's main subject category. For works prior to this period (such as those that established feedback literacy, evaluative judgment, and internal feedback), backward citation tracking from more recent publications was used, a method common in developing theoretical syntheses. Empirical, conceptual, and review articles that directly addressed one of the four constructs above, were published within the primary research section of the source journal, and were available in full text via the authors' institutional subscription were more likely to be included at the article level. Works were excluded if they focused on automatic feedback outside the formal learning context, on the technical architecture of a system without an explicit learning outcome, or if their empirical content could not be identified through a verifiable, indexed source. This filtering process resulted in the corpus synthesized below, which the authors believe is sufficiently deep but not broad for the aims of the study, namely building theories.

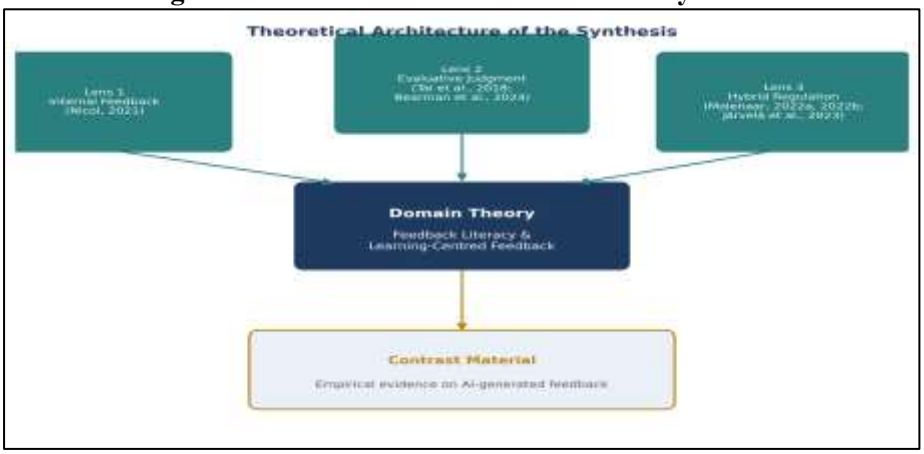
## Results and Discussion

For decades, feedback has been viewed as a transmission process – the teacher provides comments, and the student, in theory, uses them. However, modern studies have shown that this view is incomplete and have established another paradigm in which feedback is understood as a process through which students receive information about their work and act on it to enhance their future work or strategies (Carless, 2022; Henderson et al., 2019). In this paradigm, the impact of feedback is not in the feedback itself but in the learner's reactions to it, making the learner's capacities a key element of the system (Panadero & Lipnevich, 2022). These capacities are summarized in the notion of feedback literacy. Carless and Boud (2018) described it as comprising four interrelated parts: appreciating the value of feedback, forming judgments about quality, managing the affect that comments evoke, and acting on feedback. Molloy et al. (2020) added an empirically grounded, learning-centered approach to the construct, whereas Malecka et al. (2022) identified three channels for introducing the construct in courses: eliciting, processing, and enacting feedback.

The present study is an ideal candidate for these mechanisms because feedback is framed as a process the learner engages in rather than a process the learner simply receives. The construct has been expanded and critiqued in ways that warrant incorporation. Chong (2021) introduced an ecological perspective on feedback literacy, focusing on how individual and contextual factors interact, while Gravett (2022) expanded the concept into a socio-material practice that accounts for networked people, technologies, and artifacts, including AI as an artifact. However, in a

critical review, Nieminen and Carless (2023) cautioned against reducing the construct to a set of separate skills without considering design and power, and Carless and Winstone (2023) demonstrated its dependence on teacher feedback literacy. Furthermore, in digital environments, implicit conceptualizations of feedback influence practice, as Jensen et al. (2021) found in their metaphor analysis.

**Figure 1. Theoretical Architecture of the Synthesis**




Note. The domain theory (feedback literacy and learning-centred feedback) is integrated with three theoretical lenses and contrasted against empirical evidence on AI-generated feedback. Own elaboration.

The second line of research takes the learner's perspective, moving from comments to the learner's internal processes. Ultimately, all feedback is internal, meaning learners compare what they know or have produced with reference materials, whether that feedback comes from an instructor, a model text, peer feedback, or a rubric (Nicol, 2021). The important thing is not merely to deliver comments but to deliberately engage in comparisons, and the various combinations of reference points will generate different levels of internal feedback. Nicol and McCallum (2022) showed that making the multiple comparisons that occur during peer review explicit helps students provide their own feedback, which can be content-matched with the instructor's feedback. This view aligns with the construct of evaluative judgment, the ability to make judgments about the quality of one's own and others' work (Tai et al., 2018). Appropriate notions of quality, therefore, are a condition of possibility for internal feedback, if comparisons are to be anything more than blind. It also aligns with self-regulated learning, as integrative feedback models have reached consensus that information can only enhance learning when it is part of information cycles of goal-setting, monitoring, and strategic adjustment (Panadero & Lipnevich, 2022). This second pillar – the competent learner does not wait for feedback – they produce it – is complete with the concept of agentic engagement with feedback (Winstone et al., 2017) and the enabling role of feedback literacy for self-assessment (Yan & Carless, 2022). Automated feedback is not new—automated feedback systems are well discussed and categorized (Deeva et al., 2021).

The qualitatively different thing about generative AI is its dialogic, open-ended, and pervasive nature. While large language models offer personalization and availability, they also have drawbacks regarding reliability, bias, privacy, and transparency (Chiu et al., 2023; Holmes & Tuomi, 2022; Kasneci et al., 2023). These systems are not only new and innovative but also pose a number of ethical concerns, including their opacity and their ability to produce plausible but false information, which is especially significant when they are intended to support learning and is particularly salient in biomedical and natural sciences settings where factual and technical accuracy is safety-relevant (Yan, Sha, et al., 2024). Comparative data is available but complex. On most quality criteria for writing feedback, well-trained human raters perform better than ChatGPT, with relatively small gains, but with much greater time savings, as found by Steiss et al. (2024).

In a study involving students who received both ChatGPT and peer feedback, Banihashem et al. (2024) found that the two types of feedback have complementary strengths: ChatGPT feedback is more descriptive and product-oriented, while peer feedback is more problem-oriented. Previous studies (Escalante et al., 2023) found no significant difference in learning gains between the two types of feedback, with a split in student preferences. An experiment showed that language model feedback led to more text revisions, greater motivation, and more positive emotions among secondary students (Meyer et al., 2024). The relational dimension is added by learners' perceptions and attributions. Students can vary in their belief in the reliability of AI (Ruwe & Mayweg-Paus, 2023), in their differential assessment of the usefulness and trust in AI feedback compared with instructor feedback (Dawson et al., 2025), and in their perception and use of AI in authentic scenarios (Er et al., 2025). Taken together, this makes

it clear that generative AI is already a valid source of performance information, quite adequate and, to be sure, very accessible. The educational challenge now becomes the acquisition of knowledge rather than feedback, which is the territory of feedback literacy and, in technical domains, of critical appraisal—how to ask for feedback, how to compare, how to evaluate, how to convert it into knowledge. The fourth part of the synthesis is the danger that learning is undermined by excessive help.

However, when Darvishi et al. (2024) used AI prompts to help students, they found that performance increased during assistance but declined when prompts were withdrawn, suggesting a possible lack of transfer of learning and a risk of dependency. Fan et al. (2025) use the term "metacognitive laziness" to describe how students may not develop knowledge and motivation but instead seek immediate support for the task by delegating monitoring and evaluation to ChatGPT's support system. Yan, Greiff, et al. (2024) place these findings in the broader context of a dichotomy between the promise and the dangers of personalization, cognition, metacognition, and creativity. Given these dangers, it is important to consider the distribution of control between learner and AI in light of the literature on hybrid regulation of learning. Molenaar (2022a, 2022b) theorizes about levels of automation in which regulatory control is dynamically shared between technology and the learner, with educational value focused at intermediate levels, where AI supplements the learner. Järvelä et al. (2023) present a shared human–AI regulation model that enables the system to detect events that trigger the learner to execute a regulatory process. What used to be the question of whether to use AI is now who will take on metacognitive control at every step of the process. This framework is completed by the capacities learners need for a world with AI. Markauskaite et al. (2022) and Gašević et al. (2023) share the view that such capacities are not necessarily technical; they extend beyond cognitive, social, and ethical aspects. Bearman and Ajjawi (2023) offer a pedagogy for working with the black box by setting out the limits of trust in opaque systems.

Convergently, Bearman et al. (2024) suggest that as the world enters the age of generative AI, it is crucial to cultivate students' critical and reflective understanding of AI-generated content and algorithms, a uniquely human skill in an age of rapid technological change. These contributions constitute the raw material for the proposed model, which creates its moderating layers. This conflict is not just a conflict between the biomedical and natural sciences. Critical appraisal of the literature is a core competency in evidence-based practice and in the training of researchers in general, and is usually developed through structured checklists, journal clubs, and supervised critiques of published articles. Generative AI also extends the object of appraisal beyond the published record to the system's recorded output as text, which is already in the form of commentary, explanation, or suggestion, rather than a declaration to be verified against a source. The skills these disciplines already provide—the ability to analyze methods, evaluate evidence, and distinguish between well-supported and plausible claims—are exactly the skills the model identifies in its third stage, and the synthesis outlined here suggests that they must now be applied reflexively to the AI dialogue itself, not just to the literature it refers to.

**The Critical Appraisal Agency Model (CAAM): The Learner as Agent of Their Own AI-Mediated Feedback.**

The Critical Appraisal Agency Model (CAAM) is based on three assumptions derived from the synthesis above. First, the processes the learner uses when learning from feedback are more important than the information itself (Carless & Boud, 2018; Nicol, 2021). Second, generative AI transforms feedback into an abundant, dialogic, and fallible resource, shifting the burden from the availability of feedback to the judgment of whether it is correct (Kasneci et al., 2023; Steiss et al., 2024); this has particular consequences in the biomedical and natural sciences, where the cost of accepting a plausible but incorrect explanation can be high. Third, the role of the AI as an assistant rather than a capacity-remover depends on the distribution of regulatory control between the learner and the AI (Darvishi et al., 2024; Molenaar, 2022a). On these premises, generative feedback is considered the process by which the learner is able to produce understandings and decisions for improvement through conversation with generative AI systems; the process in which the learner can request information about their performance; the comparison of their work with other reference points; the critical judgment of the quality of the performance; and the act of implementing the results in the revision and strategic adjustments of the performance.

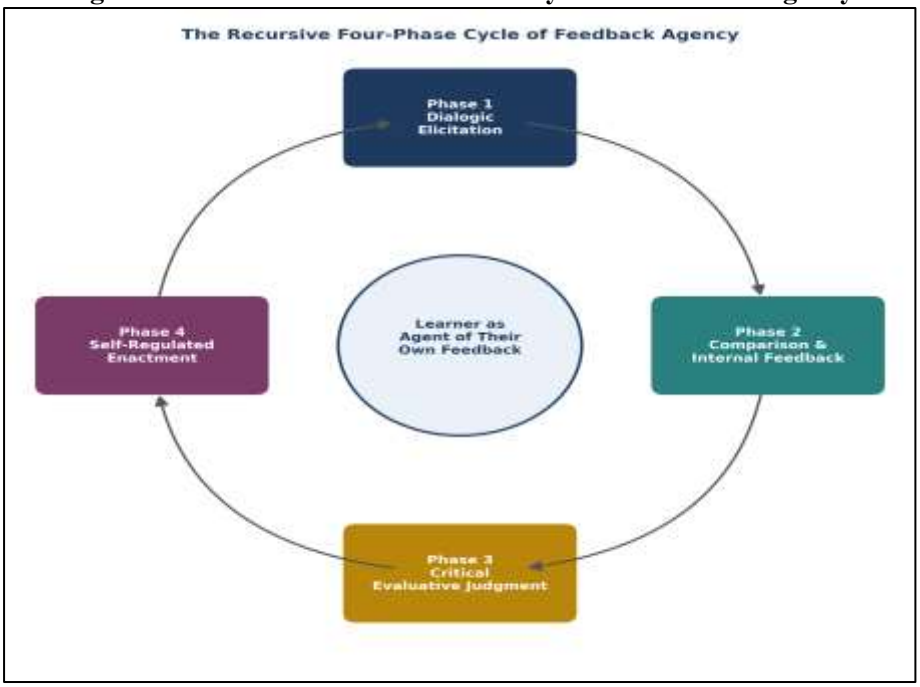
It is not only the technology that produces information, but also the feedback the learner generates in his or her own mind through comparisons (Nicol, 2021). AI is not a service that provides readers with complete feedback; it provides the content for the reader to produce, if they have the necessary capacities and conditions. Based on this, the model's main category can be specified. The term feedback agency is used here to refer to the learner's attitude and skill in intentionally initiating, directing, and assessing feedback processes: when and how to seek information, how to use it to compare with other information, and what information to accept and what to reject. This definition radicalizes Winstone et al.'s (2017) concept of proactive recipience and extends the agentic approach to feedback, as the active role ascribed to the learner in the feedback construct (Carless & Boud, 2018): In the generative environment, agency is no longer an attitude toward received comments, but rather the mechanism that makes up feedback itself.

**The Four-Phase Cycle**

The first step is dialogic elicitation: feedback from AI differs from an instructor's comment in that it is not provided automatically but must be requested, set up with conditions, tailored to the purpose, and formulated by the learner. Here, the first approach proposed by Malecka et al. (2022) for curricular feedback literacy is emphasized: making a request is equivalent to stating explicit quality criteria and learning goals, and it allows multiple iterations in the dialogue, treating each response as input rather than output. A poor, generic, or delegatory elicitation – if this is what is generally expected – suggests a poor cycle; a deliberate elicitation, grounded in criteria, will begin a rich one. The second step is comparison and incoming feedback. As in Nicol (2021), the rubric, exemplar, and peer comments are not the most important part of an AI's output—that lies in the comparison the learner makes between the AI's output, their work, and other references. As anticipated, because it can generate alternate options, counterexamples, and explanations on the fly, AI creates even more opportunities for comparison, similar to the many comparisons Nikol and McCallum (2022) wrote about in the area of peer reviews. Feedback, as defined in this phase, is the internal feedback generated by the system (i.e., other than the system's text). Third phase– Critical Evaluative Judgment. Since generative systems produce information that may be correct, incorrect, or biased, and either generic or not, the learner needs to assess its quality, relevance, and validity before adding it.

Given that generative systems can produce information that is plausible yet potentially incorrect, biased, or generic, the learner must assess the quality, relevance, and reliability of the information before incorporating it. This step implements Bearman et al.'s (2024) proposal for operationalizing methods for evaluating AI outputs and processes and reflects a classic construct in the literature (Tai et al., 2018). It also addresses how to handle attributions of authority – while more and more people document how often and to what extent they assume AI is knowledgeable (Ruwe & Mayweg-Paus, 2023), equally important is how to use errors to calibrate the amount of trust placed in the system by the human (Bearman & Ajjawi, 2023). In the context of the biomedical and natural sciences, this step comes closest to the critical evaluation of evidence as practiced by the disciplines, but now it concerns the "evidence" provided by AI commentary rather than published evidence. The step of self-regulated enactment is the fourth phase. The cycle ends only when internal feedback deemed useful is translated into action: feedback is incorporated into the work, strategies are revised, new goals are set, or another round of feedback on the work is initiated, which takes the work to a new level. This phase includes the action component of feedback literacy (Carless & Boud, 2018), the proactive recipience described by Winstone et al. (2017), and the self-regulation cycles of integrative feedback models (Panadero & Lipnevich, 2022). The agent separates itself from the consumer: of the suggestions AI gives them, the consumer accepts some, while the agent chooses whether to implement them or discard them, and the lessons learned relate to its own choices.

**Figure 2. The Recursive Four-Phase Cycle of Feedback Agency**



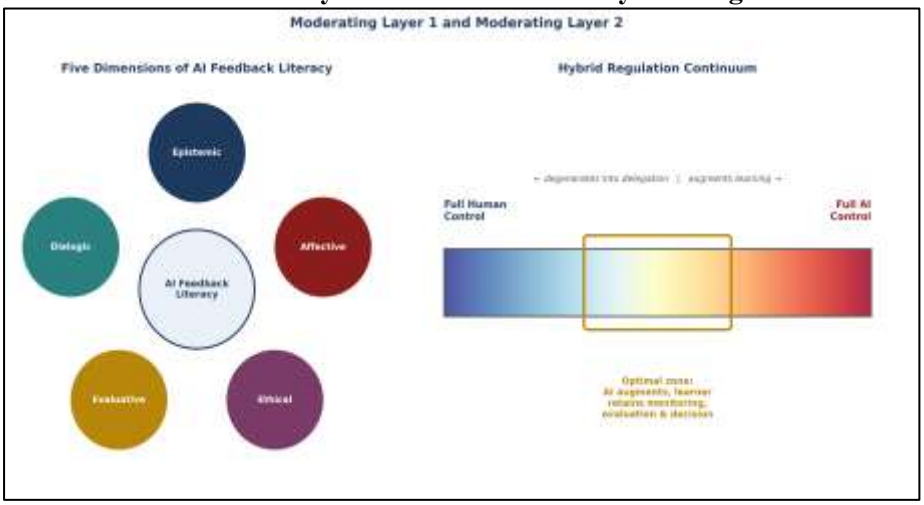

Note. The four phases unfold clockwise and recursively around the learner's core feedback agency. Own elaboration.

**Moderating Layers and Enabling Condition.**

The cycle is not a closed unit. The first of the two moderating dimensions is the AI feedback literacy extended component of feedback literacy, according to the model, which is broken down into: epistemic – understanding what the AI output is and how it is generated, including its limitations; dialogic – knowing how to have high-quality dialogues and then responding by asking for further iterations of the process; evaluative – having the judgment necessary to calibrate what the outputs are and how they are produced, including what is considered appropriate and inappropriate; ethical – understanding the limitations of academic integrity, privacy, and bias while using these systems (Holmes & Tuomi, 2022; Yan, Sha, et al., 2024); and affective – managing the frustrations and complacency that dialogue with the machine might induce – part of the original construct (Carless & Boud, 2018) and related to the broader capacities needed for a world with AI (Gašević et al., 2023; Markauskaite et al., 2022).

The second element that moderates learning is hybrid regulation of learning. As in Molenaar (2022a, 2022b) and Järvelä et al. (2023), this layer represents the dynamic interplay between the learner's and the AI's control over the metacognitive regulator throughout the cycle. The theory suggests that the best learning outcomes from using AI occur when it is used at a moderate level of automation – with the AI offering information and detection – and the learner remains in charge of monitoring, evaluation, and decision-making. However, when control is too high (very much in the system), control becomes a matter of delegation. The product improves, but learning doesn't (see helper-helper effects by Darvishi et al., 2024 and metacognitive laziness by Fan et al., 2025).

**Figure 3. AI Feedback Literacy Dimensions and the Hybrid Regulation Continuum**




Note. Left: the five dimensions of AI feedback literacy. Right: the continuum of metacognitive control between learner and AI, with the optimal zone at intermediate levels of automation. Own elaboration.

To enable this, pedagogical design and curriculum are required. There is no such thing as a self-reliant feedback agency; it needs tasks that require explicit comparison and iterative improvement, scaffolded by researchers/examiners, readily accessible quality criteria, and a well-entrenched culture of evaluation that values giving feedback on the process rather than its product (Henderson et al., 2019; Malecka et al., 2022). It takes into account warnings by Nieminen and Carless (2023) about the need for more than just individual reading of the construct, its interdependence with teacher feedback literacy (Carless & Winstone, 2023), and the ecological and sociomaterial sensitivity that places the cycle within a network of people, technologies, and artifacts (Chong, 2021; Gravett, 2022). Design also fulfills an equity agenda, of course, as access to high-quality systems and the ability to use them well are socially distributed (Yan, Sha, et al., 2024); feedback agency is often reserved for those who already have access to systems and experience using them well, a danger that is especially acute in critical-appraisal training in resource-unequal biomedical and science programs.

**Theoretical Propositions.**

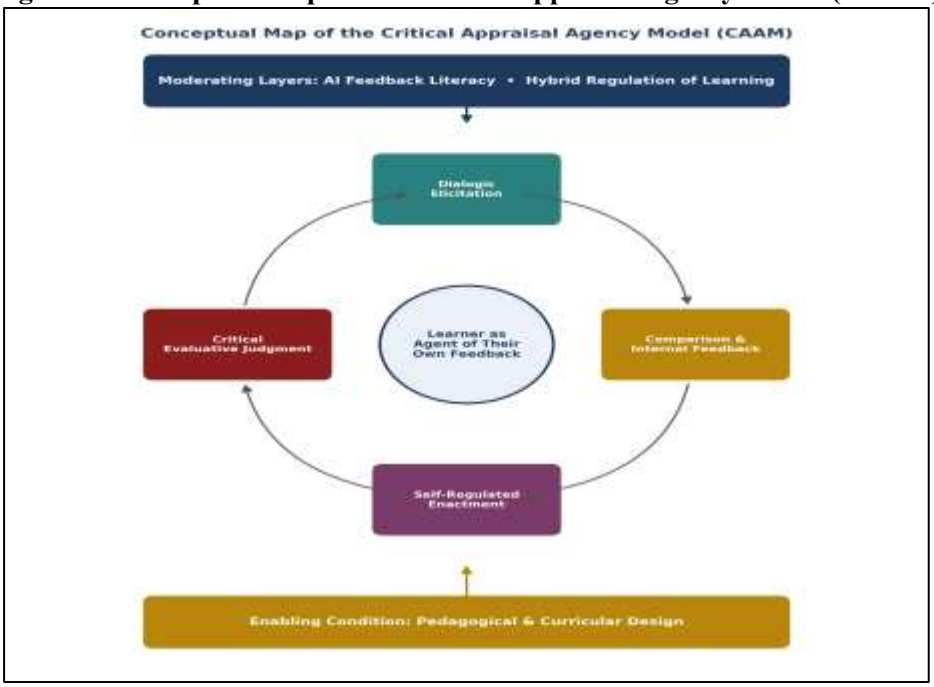
The integration just described is documented in six propositions, which were formed in accordance with the suggestions offered for conceptual theorizing ( Cornelissen, 2017):

P1. The quality of the learner's elicitation, comparison, judgment, and enactment processes is more important for learning from AI-mediated feedback than the informational quality of system outputs.

P2. Using dialogue with generative AI increases the number of points of comparison and, therefore, opportunities for feedback, as the learner intends to compare or not compare generative AI outputs with their own resources.

P3. Evaluative judgment serves as the filter in the cycle: The more detailed the evaluative judgment, the greater the likelihood of selectively transferring valuable outputs and the lower the likelihood of selectively accepting defective outputs.
P4. The relationship between the intensity of AI feedback use and the development of self-regulation is positive if the learner maintains metacognitive control, but becomes negative when such control is continually handed over to the AI.
P5. AI feedback literacy moderates the relationships in this cycle in a differentiating manner: The epistemic and dialogic dimensions increase the quality of elicitation and comparison (P1, P2); the evaluative dimension enhances the filtering function of judgment (P3); and the ethical and affective dimensions maintain a distribution of regulatory control favorable to the learner (P4).
P6. Pedagogical and curricular design are enabling conditions for the cycle: the lack of tasks, scaffolding, and an evaluative culture that requires them will result in the use of AI that implies delegation and dependence rather than agency.
The aggregate, or sum, of the CAAM is shown in Figure 4. The embodiment of this learning model category places the individual at the center, with four components of feedback cycles that rotate clockwise on the model, each of which is followed recursively: dialogic elicitation is followed by comparison and internal feedback, then by critical evaluative feedback, and finally by self-regulated enactment. The two moderating layers (AI feedback literacy and hybrid regulation of learning) act on the cycle upstream in the diagram, with pedagogical and curricular design sustaining the cycle as an enabling condition at the bottom. This sustained and moderated cycle yields the model's outcomes: sustainable evaluative judgment, self-regulation, and learning transfer to an unassisted context.

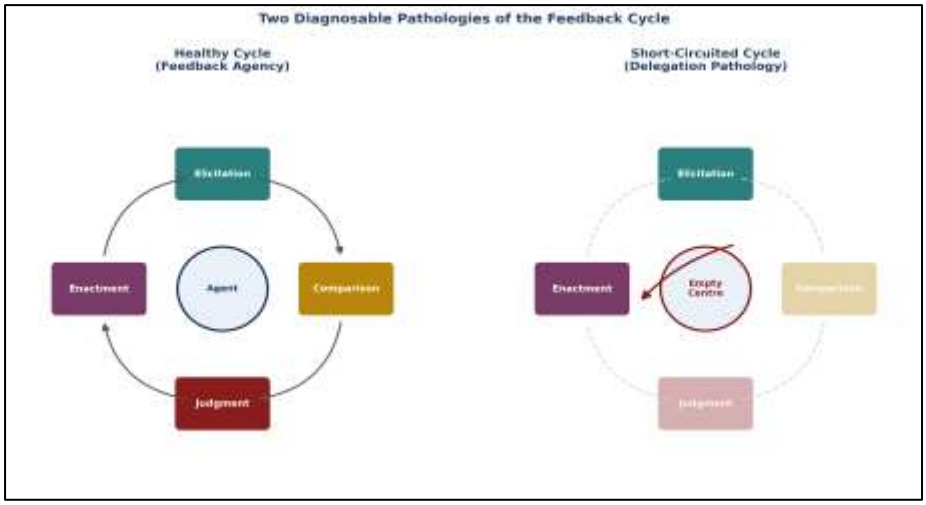
**Figure 4. Conceptual Map of the Critical Appraisal Agency Model (CAAM)**




Note. The learner's core feedback agency sits at the centre of the four-phase cycle, moderated from above by AI feedback literacy and hybrid regulation, and sustained from below by pedagogical and curricular design. Own elaboration.

By reading the map, two pathologies become visible that are diagnosed by the model. The first is the "short-circuiting" of the cycle—the shortcut that involves no comparison and no judgment. This results in an empty center of the diagram, where agency is lost, and configures a delegation associated with dependence and metacognitive laziness, as evidence suggests (Darvishi et al., 2024; Fan et al., 2025). The second is the lack of a base, of being placed somewhere where the moderating layers can develop. If pedagogical design demands and scaffolds the cycle, then the capacity of the moderating layers could also be achieved, and this could be an achievement of design (Gravett, 2022; Nieminen & Carless, 2023).

**Figure 5. Two Diagnosable Pathologies of the Feedback Cycle**




Note. Left: a healthy cycle in which all four phases are traversed and the centre is occupied by an active agent. Right: the short-circuited, delegation pathology, in which comparison and judgment are bypassed and the centre is left empty. Own elaboration.

In this study, the learner in the era of generative AI shifts from a mere recipient, merely recording comments on the document, to an agent that generates its own comments with the help of AI, using a comparison document. This repositioning expands the theory of internal feedback (Nicol, 2021) by integrating evaluative judgment in the age of AI (Bearman et al., 2024) and hybrid regulation (Järvelä et al., 2023; Molenaar, 2022a) within a single architecture of phases, layers, and propositions. By contrast, discourses that treat learners as consumers of automated services (Bearman et al., 2023) provide a theoretical tool for designing and researching the opposite, directly applicable to disciplines involving critical analysis of evidence.

The idea of generative feedback offers another lens for interpreting the evidence that has accumulated over the years. Apparent contradictions, such as AI feedback increasing product quality and engagement but not always improving learning outcomes or improving on its own as evidenced in a few studies (Escalante et al., 2023; Meyer et al., 2024), no longer appear paradoxical once one uses the distinction between receiving feedback from the machine and generating feedback within oneself. The model makes a directly testable prediction that the effects of these uses are durable and become focused in those uses that span the entire cycle. To make it easier to teach, the model recommends shifting the design from controlling access to AI to designing the cycle. In biomedical and natural sciences courses, this means activities that require drafts and iterations, such as multiple versions of a laboratory report or a critical appraisal of a research article, with prompts for learners to document requests, the rationale for accepting or rejecting, and the explanation, similar to multiple comparisons that have been demonstrated to have a significant positive impact on learning in the context of peer review (Nicol & McCallum, 2022).

The teacher's role and the teacher's resource as a designer of criteria, a modeler of the process of judging, and a guarantor of the culture of evaluation are preserved in tandem with the interdependence between teacher and student feedback literacies, as described by Carless & Winstone (2023). In the institutional context, the model shifts the dichotomy of prohibition vs. permission to AI policy. The five dimensions of ‘AI feedback literacy' can be used to design cross-curricular AI feedback training programs, which are much sought-after within accreditation requirements for health-professions and science education, where critical appraisal and evidence-based reasoning are already core to the learning experience. Training programs for which students and educators are already primed to design, build, and assess tools within these dimensions (Molenaar, 2022b) are readily adaptable to the criteria of the hybrid regulation layer: systems that keep metacognitive control with the learner, metacognize the fallibility of the system, and allow the level of assistance to be adjustable (Molenaar, 2022b).

For developers, the model proposes designing interfaces that request criteria before rendering comments, present alternatives, and support integration into the user's decision-making without forcing a choice. Instead, they should make users aware of the available options—thereby making technological design an ally of the cycle rather than a shortcut around it (Gašević et al., 2023). The model also suggests that in biomedical and natural science programs, rather than taking a single shot at a draft, learners should develop a critical appraisal log that records each question posed to an AI system, the reference point(s) used to assess the AI's answer, and the outcome of the process, whether

it was accepted, modified, or rejected. Such arrangements are already widely used in these fields and could provide a pre-existing context for modeling this practice as a group—where a walk-through of AI commentary might be useful.

**Limitations and Scope for Future Research**

This study has a few limitations. The model is not meant to be exhaustive and was developed by excluding asymmetries across the entire field of dialogue research to ensure focus on the integrated streams, though it may be enriched and challenged by other lenses (e.g., sociocultural approaches to dialogue or the political economy of platforms). The literature that is accessible is, for the most part, from higher education, Anglophone, and related to higher education; this means care should be taken when applying the model to other practices at other levels and in different academic cultures, many of which, such as biomedical and scientific programs in many Spanish-speaking and Latin American contexts, are not well represented in the literature surveyed. Lastly, with the rapid pace of technological change, the concrete capacities of the systems cited will change; the model was deliberately formulated at the level of learner processes, as capacities that are more enduring and stable than the tools, but its propositions will have to be reconsidered given future evidence.

Four streams of research agenda emerge from the model. First, operationalization: instrument development and validation to assess students' ability to provide feedback and operate generative systems across the five aspects of the feedback cycle, along with protocols for trace analysis to establish which phases of the cycle are encountered in logs of dialogue between the student and a generative AI system, leveraging opportunities afforded by multimodal data as highlighted in shared regulation research (Järvelä et al., 2023). Second, to test the propositions using experimental and longitudinal designs, with a focus on the assistance withdrawal/withdrawing aspect as a significant test of transfer when implementing the experimental methods.

Third, conditions and equity studies: investigating variation by discipline, level of education, language, and unequal contexts, where access to high-quality systems is unequal and unfamiliarity with them is socially distributed, with a deep focus on contexts of Spanish-speaking students and Biomedical and Natural Sciences programs, both underrepresented in the reviewed literature (Yan, Sha, et al., 2024). Fourth, the relational and affective dimension: understanding how perceptions of AI authority influence evaluative judgment and how perceptions of usefulness and trust change over time as systems are ‘embedded' in institutional ecosystems for assessment, such as the critical appraisal training that takes place in health and science curricula.

## Conclusion

This study presents a theoretical synthesis that integrates research on feedback literacy, internal feedback, evaluative judgment, self-regulation, and hybrid regulation to understand how the feedback process shifts with the assistance of generative AI. The outcome is the Critical Appraisal Agency Model (CAAM), which resituates the learner as the agent of their own feedback, presents a recursive system of dialogic elicitation, comparison, critical evaluative judgment, and enacting SRL (self-regulated enactment), mediated by the distribution and enactment of regulatory control and supported by AI feedback literacy, and enabled by pedagogical design. The model is set out in six propositions, and a set of conceptual figures provides the formal framework and subjects the model to empirical testing. Feedback is not a problem that generative AI can solve; it's a problem that it can change. The challenge now is to produce learners able to generate, judge, and utilize information, not just deliver information—as previously it was to get comments; today it is to get judgment, very recently, in a way that would undoubtedly and distinctly include. For instance, in biomedical and natural science, where the critical appraisal of evidence is a skill and a guard against error. Where the art of education designs for its full cycle, the copiousness of artificial feedback may become the best source in history for teaching evaluative judgment and self-regulation; where it does not, it will become the best source for teaching delegation and dependence. The model developed is intended to serve as a guide for researchers and instructors who want to take the first path.

**Conflicts of Interest**: The authors declare no conflicts of interest.

## References

1. Banihashem, S. K., Kerman, N. T., Noroozi, O., Moon, J., & Drachsler, H. (2024). Feedback sources in essay writing: Peer-generated or AI-generated feedback? International Journal of Educational Technology in Higher Education, 21, Article 23. https://doi.org/10.1186/s41239-024-00455-4
2. Bearman, M., & Ajjawi, R. (2023). Learning to work with the black box: Pedagogy for a world with artificial intelligence. British Journal of Educational Technology, 54(5), 1160–1173. https://doi.org/10.1111/bjet.13337

3. Bearman, M., Ryan, J., & Ajjawi, R. (2023). Discourses of artificial intelligence in higher education: A critical literature review. Higher Education, 86(2), 369–385. https://doi.org/10.1007/s10734-022-00937-2
4. Bearman, M., Tai, J., Dawson, P., Boud, D., & Ajjawi, R. (2024). Developing evaluative judgement for a time of generative artificial intelligence. Assessment & Evaluation in Higher Education, 49(6), 893–905. https://doi.org/10.1080/02602938.2024.2335321
5. Carless, D. (2022). From teacher transmission of information to student feedback literacy: Activating the learner role in feedback processes. Active Learning in Higher Education, 23(2), 143–153. https://doi.org/10.1177/1469787420945845
6. Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. Assessment & Evaluation in Higher Education, 43(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354
7. Carless, D., & Winstone, N. (2023). Teacher feedback literacy and its interplay with student feedback literacy. Teaching in Higher Education, 28(1), 150–163. https://doi.org/10.1080/13562517.2020.1782372
8. Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. Computers and Education: Artificial Intelligence, 4, Article 100118. https://doi.org/10.1016/j.caeai.2022.100118
9. Chong, S. W. (2021). Reconsidering student feedback literacy from an ecological perspective. Assessment & Evaluation in Higher Education, 46(1), 92–104. https://doi.org/10.1080/02602938.2020.1730765
10.Cornelissen, J. (2017). Editor's comments: Developing propositions, a process model, or a typology? Addressing the challenges of writing theory without a boilerplate. Academy of Management Review, 42(1), 1–9. https://doi.org/10.5465/amr.2016.0196
11.Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D., & Siemens, G. (2024). Impact of AI assistance on student agency. Computers & Education, 210, Article 104967. https://doi.org/10.1016/j.compedu.2023.104967
12.Dawson, P., Bearman, M., Dollinger, M., & Boud, D. (2025). Comparing generative AI and teacher feedback: Student perceptions of usefulness and trustworthiness. Assessment & Evaluation in Higher Education. Advance online publication. https://doi.org/10.1080/02602938.2025.2502582
13.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
14.Er, E., Akçapınar, G., Bayazıt, A., Noroozi, O., & Banihashem, S. K. (2025). Assessing student perceptions and use of instructor versus AI-generated feedback. British Journal of Educational Technology, 56(3), 1074–1091. https://doi.org/10.1111/bjet.13558
15.Escalante, J., Pack, A., & Barrett, A. (2023). AI-generated feedback on writing: Insights into efficacy and ENL student preference. International Journal of Educational Technology in Higher Education, 20, Article 57. https://doi.org/10.1186/s41239-023-00425-2
16.Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gašević, D. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. British Journal of Educational Technology, 56(2), 489–530. https://doi.org/10.1111/bjet.13544
17.Gašević, D., Siemens, G., & Sadiq, S. (2023). Empowering learners for the age of artificial intelligence. Computers and Education: Artificial Intelligence, 4, Article 100130. https://doi.org/10.1016/j.caeai.2023.100130
18.Gravett, K. (2022). Feedback literacies as sociomaterial practice. Critical Studies in Education, 63(2), 261–274. https://doi.org/10.1080/17508487.2020.1747099
19.Henderson, M., Phillips, M., Ryan, T., Boud, D., Dawson, P., Molloy, E., & Mahoney, P. (2019). Conditions that enable effective feedback. Higher Education Research & Development, 38(7), 1401–1416. https://doi.org/10.1080/07294360.2019.1657807
20.Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. European Journal of Education, 57(4), 542–570. https://doi.org/10.1111/ejed.12533
21.Jaakkola, E. (2020). Designing conceptual articles: Four approaches. AMS Review, 10(1–2), 18–26. https://doi.org/10.1007/s13162-020-00161-0
22.Järvelä, S., Nguyen, A., & Hadwin, A. (2023). Human and artificial intelligence collaboration for socially shared regulation in learning. British Journal of Educational Technology, 54(5), 1057–1076. https://doi.org/10.1111/bjet.13325
23.Jensen, L. X., Bearman, M., & Boud, D. (2021). Understanding feedback in online learning: A critical review and metaphor analysis. Computers & Education, 173, Article 104271. https://doi.org/10.1016/j.compedu.2021.104271
24.Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of

large language models for education. Learning and Individual Differences, 103, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274
25.MacInnis, D. J. (2011). A framework for conceptual contributions in marketing. Journal of Marketing, 75(4), 136–154. https://doi.org/10.1509/jmkg.75.4.136
26.Malecka, B., Boud, D., & Carless, D. (2022). Eliciting, processing and enacting feedback: Mechanisms for embedding student feedback literacy within the curriculum. Teaching in Higher Education, 27(7), 908–922. https://doi.org/10.1080/13562517.2020.1754784
27.Markauskaite, L., Marrone, R., Poquet, O., Knight, S., Martinez-Maldonado, R., Howard, S., Tondeur, J., De Laat, M., Buckingham Shum, S., … Siemens, G. (2022). Rethinking the entwinement between artificial intelligence and human learning: What capabilities do learners need for a world with AI? Computers and Education: Artificial Intelligence, 3, Article 100056. https://doi.org/10.1016/j.caeai.2022.100056
28.Meyer, J., Jansen, T., Schiller, R., Liebenow, L. W., Steinbach, M., Horbach, A., & Fleckenstein, J. (2024). Using LLMs to bring evidence-based feedback into the classroom: AI-generated feedback increases secondary students' text revision, motivation, and positive emotions. Computers and Education: Artificial Intelligence, 6, Article 100199. https://doi.org/10.1016/j.caeai.2023.100199
29.Molenaar, I. (2022a). The concept of hybrid human-AI regulation: Exemplifying how to support young learners' self-regulated learning. Computers and Education: Artificial Intelligence, 3, Article 100070. https://doi.org/10.1016/j.caeai.2022.100070
30.Molenaar, I. (2022b). Towards hybrid human-AI learning technologies. European Journal of Education, 57(4), 632–645. https://doi.org/10.1111/ejed.12527
31.Molloy, E., Boud, D., & Henderson, M. (2020). Developing a learning-centred framework for feedback literacy. Assessment & Evaluation in Higher Education, 45(4), 527–540. https://doi.org/10.1080/02602938.2019.1667955
32.Nicol, D. (2021). The power of internal feedback: Exploiting natural comparison processes. Assessment & Evaluation in Higher Education, 46(5), 756–778. https://doi.org/10.1080/02602938.2020.1823314
33.Nicol, D., & McCallum, S. (2022). Making internal feedback explicit: Exploiting the multiple comparisons that occur during peer review. Assessment & Evaluation in Higher Education, 47(3), 424–443. https://doi.org/10.1080/02602938.2021.1924620
34.Nieminen, J. H., & Carless, D. (2023). Feedback literacy: A critical review of an emerging concept. Higher Education, 85(6), 1381–1400. https://doi.org/10.1007/s10734-022-00895-9
35.Panadero, E., & Lipnevich, A. A. (2022). A review of feedback models and typologies: Towards an integrative model of feedback elements. Educational Research Review, 35, Article 100416. https://doi.org/10.1016/j.edurev.2021.100416
36.Ruwe, T., & Mayweg-Paus, E. (2023). "Your argumentation is good", says the AI vs humans: The role of feedback providers and personalised language for feedback effectiveness. Computers and Education: Artificial Intelligence, 5, Article 100189. https://doi.org/10.1016/j.caeai.2023.100189
37.Steiss, J., Tate, T., Graham, S., Cruz, J., Hebert, M., Wang, J., Moon, Y., Tseng, W., Warschauer, M., & Olson, C. B. (2024). Comparing the quality of human and ChatGPT feedback of students' writing. Learning and Instruction, 91, Article 101894. https://doi.org/10.1016/j.learninstruc.2024.101894
38.Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. Higher Education, 76(3), 467–481. https://doi.org/10.1007/s10734-017-0220-3
39.Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting learners' agentic engagement with feedback: A systematic review and a taxonomy of recipience processes. Educational Psychologist, 52(1), 17–37. https://doi.org/10.1080/00461520.2016.1207538
40.Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. Frontiers in Psychology, 10, Article 3087. https://doi.org/10.3389/fpsyg.2019.03087
41.Yan, L., Greiff, S., Teuber, Z., & Gašević, D. (2024). Promises and challenges of generative artificial intelligence for human learning. Nature Human Behaviour, 8(10), 1839–1850. https://doi.org/10.1038/s41562-024-02004-5
42.Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2024). Practical and ethical challenges of large language models in education: A systematic scoping review. British Journal of Educational Technology, 55(1), 90–112. https://doi.org/10.1111/bjet.13370
43.Yan, Z., & Carless, D. (2022). Self-assessment is about more than self: The enabling role of feedback literacy. Assessment & Evaluation in Higher Education, 47(7), 1116–1128. https://doi.org/10.1080/02602938.2021.2001431