



OPEN ACCESS

EDITED BY

Gavin T. L. Brown,
The University of Auckland, New Zealand

REVIEWED BY

Syamsul Nor Azlan Mohamad,
MARA University of Technology, Malaysia
Jason M. Lodge,
The University of Queensland, Australia
Kim Schildkamp,
University of Twente, Netherlands

*CORRESPONDENCE

Therese N. Hopfenbeck
✉ Therese.hopfenbeck@unimelb.edu.au

RECEIVED 01 August 2023

ACCEPTED 02 November 2023

PUBLISHED 23 November 2023

CITATION

Hopfenbeck TN, Zhang Z, Sun SZ,
Robertson P and McGrane JA (2023)
Challenges and opportunities for classroom-
based formative assessment and AI: a
perspective article.
Front. Educ. 8:1270700.
doi: 10.3389/educ.2023.1270700

COPYRIGHT

© 2023 Hopfenbeck, Zhang, Sun, Robertson
and McGrane. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Challenges and opportunities for classroom-based formative assessment and AI: a perspective article

Therese N. Hopfenbeck^{1,2,3*}, Zhonghua Zhang¹, Sundance Zhihong Sun¹, Pam Robertson¹ and Joshua A. McGrane^{1,2}

¹Assessment and Evaluation Research Centre, Graduate School of Education, The University of Melbourne, Parkville, VIC, Australia, ²Kellogg College, University of Oxford, Oxford, England, ³The University of Science and Technology, Trondheim, Norway

The integration of artificial intelligence (AI) into educational contexts may give rise to both positive and negative ramifications for teachers' uses of formative assessment within their classrooms. Drawing on our diverse experiences as academics, researchers, psychometricians, teachers, and teacher educators specializing in formative assessment, we examine the pedagogical practices in which teachers provide feedback, facilitate peer- and self-assessments, and support students' learning, and discuss how existing challenges to each of these may be affected by applications of AI. Firstly, we overview the challenges in the practice of formative assessment independently of the influence of AI. Moreover, based on the authors' varied experience in formative assessment, we discuss the opportunities that AI brings to address the challenges in formative assessment as well as the new challenges introduced by the application of AI in formative assessment. Finally, we argue for the ongoing importance of self-regulated learning and a renewed emphasis on critical thinking for more effective implementation of formative assessment in this new AI-driven digital age.

KEYWORDS

artificial intelligence, formative assessment, self-regulation, critical thinking, classroom based assessment

Introduction

In an era marked by rapid technological advancements, artificial intelligence (AI) is now increasingly used in diverse sectors of our society, fundamentally transforming the way we live, work, and learn. Within the field of educational assessment, the introduction of AI has raised both concerns and optimisms, particularly with respect to the dynamics around AI and formative assessment classroom practices. In the current paper, we explore the opportunities and challenges AI offers and underscore the continued significance of self-regulated learning and critical thinking as essential skills in this AI-driven digital age.

A brief background of classroom-based formative assessment

Classroom-based assessment has been internationally researched for decades, both with respect to summative assessments that typically occur at the end of a learning process (e.g., [McMillan, 2013](#); [Brookhart, 2016](#)), as well as formative assessments that

involve feedback processes promoting students' learning as it happens (e.g., [Brown, 2018](#); [Lipnevich and Smith, 2018](#)). [Black and Wiliam \(1998\)](#) emphasized the pivotal role of formative assessment in providing valuable information not only to teachers but also to students, guiding improvements in teaching and learning to optimize student outcomes. Since the publication of their classic work, they have continued to refine their model through subsequent theoretical papers (e.g., [Black and Wiliam, 1998](#)). Additionally, they have supported their theoretical insights with empirical studies, documenting the tangible impact of formative assessment practices on students' learning within classroom settings (e.g., [Wiliam et al., 2004](#)).

While there is a consensus among researchers regarding the positive effects of formative assessment on students' learning ([Hattie 2009](#); [Lipnevich and Smith, 2018](#)), the term "formative assessment" itself has faced critique for lacking a cohesive definition. Instead, it has been argued to be a collection of varied definitions and practices, making it challenging to conduct rigorous evaluations of its effects ([Bennett 2011](#); [Stobart and Hopfenbeck, 2014](#)). We aim to navigate these complexities by adopting [Black and Wiliam's \(2009\)](#) definition of formative assessment: "Practice in a classroom is formative to the extent that the evidence about student achievement is elicited, interpreted, and used by teachers, learners, and their peers to make decisions about their next steps in instruction that are likely to be better, or better founded than the decisions they would have taken in the absence of the elicited evidence" (p.9).

In other words, for a teacher to conduct formative assessment, they will need to know each student, their learning progress, and how to support them to achieve their learning goals. In traditional classrooms, formative assessment challenges teachers, as it requires them to find ways of following up a whole class or classes of students and provide individualized feedback to everyone, either through teacher-assessment, peer assessment, self-assessment, group-assessment, or by other means ([Double et al., 2020](#)). As we will discuss in the next section, research has shown that these practices are difficult to implement at scale and in ways that are sustainable over time ([Hopfenbeck and Stobart, 2015](#)).

Challenges to implementing formative assessment

Several challenges to the implementation of formative assessment have been documented by researchers as presented in the January 2015 Special Issue of the *Assessment in Education* journal. [Wylie and Lyon \(2015\)](#) found substantial variation in the quality of implementing formative assessments among 202 Mathematics and Science teachers in the US context. They suggest that more targeted professional development is needed to secure high-quality implementations of formative assessment practices. Further, since formative assessment requires teachers to have high competency across a range of knowledge and skills (e.g., domain content knowledge, pedagogical content knowledge, assessment and data literacy, and knowledge of measurement fundamentals), such professional development needs to be wide in scope.

Challenges have also been found when stakeholders involved in the assessment do not share a mutual understanding of its purpose(s). For example, in the same Special Issue, [Hopfenbeck et al. \(2015\)](#) conducted an evaluation of a large-scale implementation of an assessment for learning program in Norway and found that implementation was weaker in schools where the assessment was perceived as part of an accountability system, while in schools with a high degree of trust between teachers, head teachers and the school owners at the municipality level, the quality of the implementation was better. Similarly, a study of school-based assessment in Singapore found that their high-stakes examination-focused education system created tensions when trying to implement formative assessment processes, thus demonstrating how context matters in terms of the challenges that arise between different stakeholders in the interaction between formative assessment and accountability systems ([Ratnam-Lim and Tan, 2015](#)). These findings indicate how teachers' formative assessment implementations are influenced by accountability structures, educational leadership, resources, workload and social pressures within their context.

Thirdly, formative assessment practices have primarily been researched and developed in contexts where students and teachers have access to a wealth of resources, and, thus, do not necessarily generalize to more challenging contexts. [Halai et al. \(2022\)](#) evaluated an implementation of assessment for learning practices in six schools in Dar es Salaam, Tanzania, and documented the challenges created by very large class sizes, with one teacher responsible for up to 180 students at a time, as well as under-resourced classrooms. Furthermore, it was found that the cultural assumptions of the student role assumed in most of the Western, English-speaking literature did not fit what is seen as a good student in Tanzania. Formative assessment practices expect students to be self-regulated and proactive so they can participate in peer-discussions and assessment, and as part of this, they are supposed to engage in dialog in groups and with the teacher and be able to ask critical questions. In contrast, a good student in Tanzania is expected to listen to the teacher, not ask questions or be too critical, and overall follow instructions and do what the teacher tells them to do. This is enforced by the parents' expectations of how schools and teachers need to help raise the child. Thus, the interactive dialog between teachers, students and peers that are at the heart of formative assessment can be culturally and contextually sensitive, which poses challenges for implementing a 'one size fits all' formative assessment practice across different contexts.

Finally, studies have reported that teachers find it challenging to provide enough feedback to students, particularly at crucial times in the learning process, as well as with the quality of feedback required to further each student's learning, due to time and other resource constraints ([Brooks et al., 2019](#); [Gamlem and Vattøy, 2023](#)). As a result, formative assessment theory suggests that teachers need to design classrooms where students can provide feedback to each other to help reduce this workload ([Wiliam, 2011](#)). However, teachers still report that they struggle to manage classrooms where these peer assessment practices are established ([Dignath et al., 2008](#); [Halai et al., 2022](#)). So, even in well-resourced contexts where teachers endeavor to engage best-practices in the implementation of formative assessment in their classrooms, the high workload such practices engender continues to be a barrier.

Evolving and revolutionizing formative assessment: what can AI bring?

Given the challenges to implementing formative assessment in the classroom that are outlined in the previous section, the advancement of AI and relevant products (e.g., ChatGPT) may provide opportunities to overcome some of these challenges, such as high number of students, and only one teacher to provide feedback. Indeed, researchers have identified assessment as one of the most significant areas of opportunity that AI and related technologies offer in education (Cope et al., 2021; Swiecki et al., 2022; Zhai and Nehm, 2023). However, new challenges may also be introduced through the widespread use of AI, including practical and ethical challenges (Milano et al., 2023). This section focuses on the changing landscape of the practice of formative assessment, especially under the influence of AI, and discusses how AI can help support teachers to provide formative assessment for students on a large and sustainable scale.

What do we mean by AI?

In our discussion, AI refers to the application of sophisticated algorithms that allow computers and machines to simulate human intelligence for successfully completing tasks (Murphy, 2019). Although different technical approaches and methods, e.g., supervised learning versus deep learning, have been used to develop AI systems, the essence of AI is to use data to teach machines to make classifications, diagnoses, decisions, predictions, and/or recommendations (Gardner et al., 2021). More specifically, the application of AI typically involves collecting large multivariate data relevant to the task of interest, applying statistical methods and sophisticated algorithms to process sets of input data to build a model(s) that identifies and weights features and/or patterns of the input variables relevant to the task, then using a different pre-collected dataset to validate the model(s) for the task where the correct output is known in advance, and then applying the model(s) to generate a task output(s) (e.g., classification, prediction, decision, or recommendation) in a context where the correct output is unknown (Murphy, 2019; Gardner et al., 2021).

Opportunities for implementing formative assessment with AI

As mentioned above, one of the challenges in formative assessment is to provide individualized, high-quality feedback to students. It is highly resource intensive for teachers to personally give or find other ways (e.g., peer assessment, self-assessment, group assessment) to provide individualized feedback to each student on a large scale. However, AI can make some of the assessment procedures fully or partly automated, making the assessment practices more feasible to maintain, which can then reduce the time burden on teachers (Swiecki et al., 2022). A typical example that has been widely discussed is automated essay scoring systems (Ke and Ng, 2019; Gardner et al., 2021). The application of AI in automated essay scoring frees teachers from the labor-intensive grading process and allows them to assign more extended writing tasks to students, automate the grading process, and, more importantly, with the integration of

natural language processing-based AI, provide timely formative feedback to help students revise and improve their writing (Murphy, 2019).

Liu et al. (2016) showed that a machine learning enabled automated scoring tool, c-rater-ML, could produce scores that were comparable to human raters in scoring students' responses to constructed response questions about science inquiry, offering a promising solution to improving the efficiency of not only obtaining the summative scores but also generating instant formative feedback (Linn et al., 2014). Another example of how AI can help is by using computers to support the management and delivery of formative assessments (e.g., Webb et al., 2013; Tomasik et al., 2018). These systems have the capacity to discern distinct learning pathways in students' progress, enabling the identification of the most suitable tasks or questions for each student at different points in time. In addition, computerized formative assessment systems can optimize the administration of formative assessments by determining their frequencies and schedules customized for every individual student (Shin et al., 2022). These findings demonstrate how AI can improve the efficiency and flexibility of formative assessment practices at the individual student level.

Another significant opportunity that AI offers for formative assessment is the improvement of feedback both in quantity and quality (Gardner et al., 2021). The main goals of formative assessment are to provide constructive feedback based on students' responses and to help teachers design differentiated instructional strategies and sustain students in self-regulating their learning. AI can delve into the data to identify the patterns on which dynamic, customized, individualized, and visualized feedback can be automatically generated (Verma, 2018; Tashu and Horvath, 2019; Lee, 2021, 2023). For example, the adaptive nature of some computerized formative assessment systems and intelligent tutoring systems enables every student's attainment to be individually and more precisely assessed, which facilitates more appropriate and targeted feedback based on their individual learning stage and trajectory (Ma et al., 2014; Tomasik et al., 2018; Mousavinasab et al., 2021). Adaptive multi-strategy feedback models, based on AI methods, have been applied in the context of such systems to automatically adapt the feedback generating strategy to individual students, which have, in turn, been found to generate more effective feedback than the traditional feedback generation methods (Gutierrez and Atkinson, 2011).

In addition, AI can improve the quality and effectiveness of peer assessment in classrooms with large class sizes. Peer assessment can be supported with prompts from language models (Er et al., 2021). This approach to peer assessment supports students in not only providing feedback to peers but also reflecting on and justifying their judgments, providing further opportunities for them to develop their self-regulated learning skills (Liu and Carless, 2006), and has been found to provide useful peer feedback to students (Luaces et al., 2018). In addition, the peer assessment reviews can both help the teachers to better understand the performance of the students in their classroom and also provide additional data (e.g., the review text) that can be analyzed using AI-based techniques (e.g., semantic, lexical, and psycho-linguistic analyses; Vincent-Lamarre and Larivière, 2021) to further enhance teachers' understanding of the performance of their students.

AI can also aid teachers in collecting and analysing longitudinal formative assessment data, and in generating learner profiles to trace

students' learning progression over time (Swiecki et al., 2022). This application of AI enables the scalable implementation of formative assessment in both cross-sectional and longitudinal contexts, which makes it more sustainable, allowing teachers to efficiently monitor the growth of student learning and identify the knowledge and skill gaps in their learning over time (Barthakur et al., 2023). Another contribution of AI to facilitating longitudinal formative assessment lies in its ability to analyse the large-scale longitudinal formative assessment data to trace the learning trajectories of the students and predict their future learning states. For example, some of the widely applied statistical methods in AI (e.g., hidden-Markov models, artificial neural networks) have been combined with traditional cognitive diagnostic models (CDM) to analyse longitudinal formative assessment data to track the changes of students' learning over time (e.g., Chen et al., 2018; Wen et al., 2020).

The application of AI also allows teachers to get an in-depth understanding of students' learning processes based on the analysis of large volumes of 'process data' rather than just the assessment artefacts (e.g., responses to questions, items or tasks) produced by the students. With modern technologies, the processes leading to assessment responses can be captured in time-stamped log stream data (Cope et al., 2021). For example, students' actions (e.g., keystrokes, editing, chat history, video watching) during an assessment can be captured. Which potentially contain additional information for understanding how students produce their responses. With the support of AI, the process data can be analyzed to investigate their strategies (e.g., identifying misconceptions), which can then provide invaluable information for individualized feedback. In addition, taking advantage of AI's increasing capacity to deal with complex, multi-media data, more authentic assessment tasks (e.g., multimedia, game-based problem solving, essay writing, performance-based tasks) can be effectively used in undertaking formative assessments (Swiecki et al., 2022).

AI also offers opportunities for assessing some hard-to-measure constructs. Students' non-scholastic attributes, including social-emotional traits (e.g., classroom engagement, self-efficacy, motivation, resilience) and social-cognitive skills (e.g., metacognition, collaborative problem-solving, critical thinking, digital literacy, self-regulated learning), have been attracting more attention and are increasingly recognized as equally important as their scholastic achievement (Durlak and Weissberg, 2011). Advances in AI and related technologies allow for these constructs to be more validly assessed, instead of purely relying on students' self-reported beliefs and behavior through questionnaires. Now, the data collected through different channels (e.g., time-stamped process data, eye contact, feedback, facial expression, eye movements, body posture and gesture) can be mined to develop indicators for assessing these different aspects of student learning. For example, MOOC data has been used to design indicators through a thorough analysis of students learning behaviors in online courses to measure students' self-regulated learning (e.g., Milligan and Griffin, 2016) and leadership development in workplace learning in an online environment (e.g., Barthakur et al., 2022). Another example is the measurement of collaborative problem-solving skills through process data that captured the actions and chats of pairs of team members collaboratively solving tasks (Griffin and Care, 2014). AI-based Large Language Models provide further promise for mining chat history data to support assessing how team members explore, define, plan, execute and solve tasks in a collaborative way.

Challenges arising from using AI in formative assessment

AI introduces not only opportunities but also challenges to formative assessment practices (Swiecki et al., 2022). A primary challenge that needs to be addressed before teachers can apply AI in their formative assessment practices is their lack of knowledge and skills relevant to AI techniques as well as their limited access to big data. Thus, although AI can potentially ease the workload of teachers by automating some aspects of formative assessment (e.g., automatising scoring and tracing students' learning progress), it adds further burden through the need for professional development in its use (Engeness, 2021). Moreover, despite the promising future for formative assessment brought by big data, with the possibility of collecting the process data through students' learning, a new challenge arises in identifying which part of the collected data is most helpful and relevant to improve student learning. In addition, the unique features of current big data (e.g., time-stamped process data, sparse data) are significantly different from that of the traditional assessment data and pose a variety of challenges to the psychometric methods for analyzing the data. To deal with this, scholars have been endeavoring to introduce new methods to integrate data science and machine learning into psychometrics (e.g., von Davier et al., 2022).

Another challenge arising from using AI in formative assessment is to tackle relevant issues about investigating the best way to use AI in formative assessment practices. One of the hotly debated issues is whether AI will replace teachers. We argue that AI should not replace but facilitate teachers' formative assessment practices and promote the role of formative assessments in supporting instruction and learning. As stated by Murphy (2019), "the best use of AI in education is to augment teacher capacity by helping teachers deliver more effective classroom instruction" (p. 14). Teachers need to understand the limitations of the AI techniques when they review the assessment results. For example, automated scoring systems have long been criticized for their inability to measure higher-order aspects of writing (e.g., creativity, argumentation, reasoning) (Gardner et al., 2021). One of the primary aims of formative assessment is to diagnose gaps in students' learning based on the well-established interpretability of the measurement scales. However, many approaches based on machine learning are designed for prediction involving complicated models for improving accuracy but sacrificing the ease of interpretation. Therefore, any inferences from the results of formative assessments involving the integration of AI techniques should only be made after having examined the assessments' validity and interpretability (Bejar et al., 2016; Scalise et al., 2021). Teachers need to critically review how the assessment results are reached and identify any sources of bias introduced by the application of AI techniques in assessment, which in turn adds to their workloads (Murphy, 2019). Finally, but not at least, the introduction of AI in the classroom cannot happen without ethical considerations for the use and application of it. Scholars have emphasized the importance of having conversations with students on the productive, ethical and critical relationship around the use of AI and future technologies (Bearman et al., 2023) and improving knowledge on data privacy for children (Johnston, 2023). With these considerations in mind, we will now turn to one example of formative assessment practices and AI.

The formative uses of rubrics and the opportunities and challenges of using AI

The formative use of rubrics, i.e., the scoring guides that are used to make judgments about the quality of students' constructed responses, such as writing, performances or products (Popham, 1997), has been shown to have a positive influence on learning. Specifically, Panadero and Jönsson (2013) argued that the use of quality rubrics plays a key role in enhancing the effectiveness of formative assessment practices. There are two main ways that rubrics are thought to improve formative assessment. They make assessment expectations explicit, thereby assisting with understanding and feedback, and they support self-regulated learning by supporting learners to monitor their own learning and make decisions about areas for improvement (Jönsson and Panadero, 2016). While the use of rubrics for formative assessment purposes has a positive effect on learning, this effect is amplified when it is accompanied by teacher-given rubric feedback that addresses the three feedback questions (Wollenschläger et al., 2016). Consequently, we now examine the formative use of rubrics as a specific example of how AI can address existing challenges, provide additional opportunities, and present new challenges to this formative assessment practice.

Firstly, AI can support the formative uses of rubrics by helping teachers to overcome some of the time needed to construct rubrics and to teach students how to use them, as these have been found to be a constraint to rubric use (English et al., 2022). Generative AI can speed up rubric design, as teachers can use it to provide samples of rubrics to assess specified constructs, and a teacher can choose to directly use these rubrics or to use them as a source of ideas for designing their own. AI also has the potential to assist students as they learn to use rubrics by providing work samples matching different levels on a rubric, by assessing student-generated work samples against a rubric so a student can check the accuracy of self-assessments, and by providing written feedback to accompany a rubric assessment. These possible AI-augmented rubric uses by students help build agency, as the students can have more control over the timing and style of feedback they receive. Therefore, AI has the potential to help teachers overcome some of the common challenges to using rubrics in the classroom.

Nonetheless, the challenges presented by potential biases in training data are also applicable to rubrics generated with AI (Li et al., 2023). Rubrics for constructs with a greater cultural influence, such as communication, are likely to be more affected than those for constructs where the subject matter is more consistent irrespective of culture, like chemistry. In addition, while there is an acknowledged need for more research on rubric design (e.g., English et al., 2022), the findings of such research often fail to be commonly adopted by teachers. One example of this is that most rubrics have structurally aligned categories, e.g., all parts of the rubric have five levels of quality. Unfortunately, there is empirical evidence that this common structure is ineffective because it increases construct-irrelevant variance by facilitating scoring based on a halo effect where the assessor makes a global judgment of quality and simply aligns the ratings on different criteria of the rubric to match rather than making independent decisions for each (Humphry and Heldsinger, 2014). Rubrics, thus, support more accurate judgments when the number of levels of quality are tailored to the specific criterion being evaluated where some criterion (e.g., quality of argumentation) have more levels than

others (e.g., use of paragraphs). Consequently, without careful curation of training data sets to ensure they meet best practice in rubric design, AI generated rubrics will likely propagate common design flaws. Moreover, exploration by researchers of the different ways AI is already augmenting rubric use in classrooms is warranted, especially in cultures and contexts that are not well represented in training data sets.

The role of self-regulated learning and critical thinking in formative uses of AI

Based on the formative assessment cycle in Ruiz-Primo and Brookhart (2018), there is a natural bridge between self-regulated learning and formative assessment, as formative assessment can be considered as a self- and co-regulated process of improving learning, which starts with defining and sharing learning goals and then through a process of gathering or eliciting information, analysing and interpreting the collected information, and finally using the collected information to make a reflective judgment on whether the pre-defined learning goal has been achieved or not. There has been a call for linking the research into self-regulation with formative assessment, as it is recognized that self-regulation will enhance students' ability to act as peer-assessors, do self-assessment and take on the proactive role needed for formative assessment practices (Brandmo et al., 2020). Despite decades of educational research into what improves students' learning, few researchers have tried to combine the two fields of formative assessment and self-regulation, although exceptions include Allal (2010), Andrade and Brookhart (2016), Brown (2018), Butler and Winne (1995), Nicol and MacFarlane-Dick (2006), and Panadero et al. (2018). Moreover, more recent research has demonstrated how students can benefit more from formative assessment practices if they are self-regulated learners (Allal, 2020; Andrade and Brookhart, 2020; Perry et al., 2020). With the rise of AI, students' ability to self-regulate will be even more important, as it opens opportunities but also challenges in how we plan, use strategies, and evaluate our learning processes.

Furthermore, we argue for the importance of critical thinking for both teachers and learners to navigate the principled use of AI and leverage the effectiveness of formative assessment as part of their process of self-regulated learning, particularly when confronted with the novel challenges of AI. Although highlighting the importance of critical thinking may seem like an already labored point in educational settings, as it has been acknowledged as a fundamental generic skill necessary for individuals to live and thrive in the 21st century (e.g., Davies and Barnett, 2015), none of the existing research has yet built a connection between critical thinking, formative assessment and self-regulated learning under the impact of AI. Before getting into the specific argument on the role of critical thinking in formative assessment and self-regulated learning, it is worth clarifying that self-regulated learning is used in a broader way in this section, extending beyond learners to encompass teachers who also need to apply their self-regulated learning skills to effectively acquire new knowledge and skills to harness the potential of AI in their teaching practice effectively. In the following part, we will briefly explain our understanding of critical thinking, depicting the role of critical thinking when facing new challenges brought by AI, and then

describing the role of critical thinking in formative assessment and self-regulated learning.

When facing the uncertain, complex issues brought by the advancement of AI, critical thinking, defined as “reasonable and reflective thinking focused on deciding what to believe or do” (Ennis et al., 2005: 1), becomes increasingly pronounced. As a complex competency, good critical thinking practitioners are expected to have a sophisticated level of epistemic beliefs (i.e., attitudes to knowledge and knowing), which are essential to recognizing the uncertainty and complexity of such controversial issues as the ethical use of AI in education. These beliefs lay a foundation for the engagement of their thinking skills (e.g., understanding, applying, analysing, synthesizing, evaluating) to be well-informed of the issue and navigate through a vast amount of potentially conflicting information (King and Kitchener, 1994; Kuhn and Weinstock, 2004; Wertz, 2019; Sun, 2021, 2023). As theorized by Dewey (1910), suspending judgments may be the most effective course of action prior to acquiring a comprehensive understanding of a relevant issue.

For different responses on the issue of whether AI should be used in educational settings, it is not surprising to witness resistance toward emerging technologies because there is a natural fear and unease that often accompanies the introduction of new technologies (Ball and Holland, 2009). However, if critical thinking is engaged before deciding what to believe or do, this natural tendency can be challenged. AI is far from a novel concept and has been an academic discipline since the 1950s (Haenlein and Kaplan, 2019; Gillani et al., 2023), and various AI technologies (e.g., image recognition, smart speakers, self-driving cars) and models (e.g., AlphaGo, Deep Blue and ELIZA) have already significantly impacted our ways of living and working (Haenlein and Kaplan, 2019). Yet, limited transformation has happened in education, with 20th-century traditions and practices still dominating our schools (Luckin and Holmes, 2016). Despite the recognition of the enormous benefits and potential of AI in transforming education, scholars’ impatience is mounting because many of these promising ideas remain confined to the lab or lecture halls with few practical breakthroughs (Luckin and Holmes, 2016).

Specific to the context of formative assessment and self-regulated learning, the role of critical thinking is also pivotal, which equips both teachers and learners to effectively address emerging challenges. For teachers, as the landscape evolves with a growing array of AI products and assessment data, being flooded by abundant online learning resources can be overwhelming (Schwartz, 2020). It would be increasingly important for teachers to critically evaluate what, when, and how to utilize these resources to enhance their teaching methodologies and bolster student learning. When facing an increasing amount of data that has been collected or needs to be collected, teachers need to critically discern how assessment data can best inform their pedagogical strategies rather than have data to dictate their teaching. Additionally, teachers should exercise discernment in determining the level of trust they can place in specific AI models when making judgments about student learning outcomes. This becomes especially crucial for teachers who should critically assess the potential biases that AI models might carry due to the use of training data (Li et al., 2022, 2023). Furthermore, as AI advancements have the potential to liberate teachers from routine and time-consuming tasks like assignment grading and rubric

development, they must engage in critical reflection. They need to consider which skills they should prioritize for their professional development, such as data literacy, and what skills should remain at the core of their teaching, notably critical and creative thinking. This critical assessment of their evolving role is essential in navigating the transformative impact of AI in education.

Regarding individual learners, engaging critical thinking can have positive contribution to the effectiveness of formative assessment and self-regulated learning when facing the opportunities and challenges introduced by the advancement of AI. For instance, the advancement of AI, indeed, can certainly be used to generate text to pass the assessment of a subject, but if learners are practicing their critical thinking and self-regulated learning skills, they may ask themselves some reflective questions, such as what is the purpose of learning? Will a certain way of using AI contribute to achieving their learning goals? When specific solutions have not been produced to address the new challenges brought by AI, individual learner’s practice of their critical thinking and self-regulated learning may contribute to the ethical use of new technologies. Despite some instances of learners exploiting AI to evade plagiarism detection systems, it is encouraging to learn from recent empirical research that many students genuinely benefit from the timely feedback and companionship provided by AI (Skeat and Ziebell, 2023). Moreover, these students display ethical awareness, being cautious and mindful of their AI usage even in the absence of well-developed regulations governing AI in education.

While some scholars have suggested that assessment is holding us back from transforming our education systems (Luckin and Holmes, 2016, p.35), the advancement of AI may catalyze a “Renaissance in Assessment” (Hill and Barber, 2014). Although the acceptance of AI may encounter some resistance, the power of new technologies, if unleashed with principled and research-driven use, may significantly change and improve ways of teaching and learning. In this vein, Australian educational policymakers made a significant shift by granting permission for the use of ChatGPT and generative AI in all government schools. This change followed the release of the Australian Framework for Generative Artificial Intelligence in Schools. It is encouraging to observe the transition from a policy that limited the use of ChatGPT across every Australian state and territory, except South Australia, to a more welcoming and adaptable stance.

Conclusion

As we have outlined in this article, despite decades of research on formative assessment practices, teachers still face several challenges in implementing these practices on a large scale. The use of AI in classrooms has the potential for supporting formative assessment practices, although we will argue, it will require some careful considerations. Based upon what we have outlined in this article, we will conclude with the following suggestions on how to integrate AI into formative assessment:

1. Utilize AI for feedback assistance, particularly in large classes where teachers struggle to give timely feedback to all students.
2. Promote self-regulating skills as students will need to take even more responsibility for their own learning, when using AI. This

includes goal-setting, monitoring progress, and adjusting their study strategies based upon AI feedback.

3. Emphasize the ethical use of AI in formative assessment and discuss the importance of integrity and responsible use of AI tools, to avoid inappropriate uses.
4. Emphasize the role of teachers in guiding the students' use of AI. Teacher can help students interpret AI feedback, set learning goals, and make informed decisions based upon AI recommendations.
5. Encourage collaborative research between educators and researchers to explore the effectiveness of AI in formative assessment. Co-design studies with teachers and students to assess how AI impacts learning outcomes and student engagement.
6. Recognize the evolving role of teachers and facilitators of AI-enhanced learning.

In the changing times of AI, students more than ever need teachers to guide them using AI, and as researchers, we would encourage colleagues to take part in co-designing studies with teachers and students, where we together examine how to improve students learning through formative assessment practices, critical thinking, self-regulated learning and AI.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

References

- Allal, L. (2010). "Assessment and the regulation of learning" in *International encyclopedia of education*. eds. P. Peterson, E. Baker and B. McGaw (Elsevier), 348–352.
- Allal, L. (2020). Assessment and the co-regulation of learning in the classroom. *Assess. Educ.* 27, 332–349. doi: 10.1080/0969594X.2019.1609411
- Andrade, H. L., and Brookhart, S. M. (2020). Classroom assessment as the co-regulation of learning. *Assess. Educ.* 27, 350–372. doi: 10.1080/0969594X.2019.1571992
- Ball, W., and Holland, S. (2009). The fear of new technology: A naturally occurring phenomenon. *The American Journal of Bioethics*, 9, 14–16. doi: 10.1080/15265160802617977
- Barthakur, A., Kovanovic, V., Joksimovic, S., Zhang, Z., Richey, M., and Pardo, A. (2022). Measuring leadership development in workplace learning using automated assessments: learning analytics and measurement theory approach. *Br. J. Educ. Technol.* 53, 1842–1863. doi: 10.1111/bjet.13218
- Barthakur, A., Dawson, S., and Kovanovic, V. (2023). Advancing leaner profiles with learning analytics: a scoping review of current trends and challenges. In LAK23: 13th international learning analytics and knowledge conference, 606–612. doi: 10.1145/3576050.3576083
- Bearman, M., Ajjawi, R., Boud, D., Tai, J., and Dawson, P. (2023) *CRADLE suggest assessment and gen AI*. Centre for Digital Learning, Deakin University, Melbourne, Australia.
- Bejar, I. I., Mislevy, R. J., and Zhang, M. (2016). "Automated scoring with validity in mind" in *The Wiley handbook of cognition and assessment: frameworks, methodologies, and applications*. eds. A. A. Rupp and J. P. Leighton (Wiley), 226–246.
- Bennett, R. E. (2011). Formative assessment: a critical review. *Assess. Educ.* 18, 5–25. doi: 10.1080/0969594X.2010.513678
- Black, P., and Wiliam, D. (1998). Assessment and classroom learning. *Assess. Educ.* 5, 7–74. doi: 10.1080/0969594X.2019.1609411
- Black, P., and Wiliam, D. (2009). Developing the theory of formative assessment. *Educ. Assess.* 21, 5–31. doi: 10.1007/s11092-008-9068-5
- Black, P., and Wiliam, D. (2018). Classroom assessment and pedagogy. *Assess. Educ.* 25, 551–575. doi: 10.1080/0969594X.2018.1441807

Author contributions

TH: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – original draft. ZZ: Conceptualization, Writing – original draft. SS: Conceptualization, Writing – original draft, Writing – review & editing. PR: Writing – original draft, Writing – review & editing. JM: Conceptualization, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Boud, D. (2005). *Enhancing learning through self-assessment* Routledge.

Brandmo, C., Panadero, E., and Hopfenbeck, T. N. (2020). Bridging classroom assessment and self-regulated learning. *Assess. Educ.* 27, 319–331. doi: 10.1080/0969594X.2020.1803589

Brookhart, S. M. (2016). "The use of teacher judgement for summative assessment in the USA" in *International teacher judgement practices*. ed. V. Klenowski (Oxon, New York: Routledge), 69–90.

Brooks, C., Carroll, A., Gillies, R. M., and Hattie, J. (2019). A matrix of feedback for learning. *Australian J. Teach. Educ.* 44, 14–32. doi: 10.14221/ajte.2018v44n4.2

Brown, G. T. L. (2018). *Assessment of student achievement*. Oxon, New York: Routledge.

Butler, D. L., and Winne, P. H. (1995). Feedback and self-regulated learning: a theoretical synthesis. *Rev. Educ. Res.* 65, 245–281. doi: 10.3102/00346543065003245

Chen, Y., Culpepper, S. A., Wang, S., and Douglas, J. (2018). A hidden Markov model for learning trajectories in cognitive diagnosis with application to spatial rotation skills. *Appl. Psychol. Meas.* 42, 5–23. doi: 10.1177/0146621617721250

Cope, B., Kalantzis, M., and Searsmith, D. (2021). Artificial intelligence for education: knowledge and its assessment in AI-enabled learning ecologies. *Educ. Philos. Theory* 53, 1229–1245. doi: 10.1080/00131857.2020.1728732

Davies, A. A., von Mislevy, R. J., and Hao, J. eds. (2021). Introduction to computational psychometrics: towards a principled integration of data science and machine learning techniques into psychometrics. *Computational psychometrics: new methodologies for a new generation of digital learning and assessment*. Boston: Springer.

Davies, M., and Barnett, R. (2015). *The Palgrave handbook of critical thinking in higher education*. Springer.

Dewey, J. (1910). *How we think*. Boston: Dover Publications.

Dignath, C., Büttner, G., and Langfeldt, H. (2008). How can primary school students learn self-regulated learning strategies most effectively? A meta-analysis on self-regulation training programmes. *Educ. Res. Rev.* 3, 101–129. doi: 10.1016/j.edurev.2008.02.003

Double, K. S., McGrane, J. A., and Hopfenbeck, T. N. (2020). The impact of peer assessment on academic performance: a meta-analysis of control group studies. *Educ. Psychol. Rev.* 32, 481–509. doi: 10.1007/s10648-019-09510-3

- Durlak, J. A., and Weissberg, R. P. (2011). Promoting social and emotional development is an essential part of students' education. *Hum. Dev.*, 54, 1–3. doi: 10.1159/000324337
- Engeness, I. (2021). Developing teachers' digital identity: towards the pedagogic design principles of digital environments to enhance students' learning in the 21st century. *Eur. J. Teach. Educ.* 44, 96–114. doi: 10.1080/02619768.2020.1849129
- English, N., Robertson, P., Gillis, S., and Graham, L. (2022). Rubrics and formative assessment in K-12 education: a scoping review of literature. *Int. J. Educ. Res.* 113:101964. doi: 10.1016/j.ijer.2022.101964
- Ennis, R., Millman, J., and Tomko, T. (2005). *Cornell critical thinking tests level X & Level Z Manual*. The Critical Thinking Co.
- Er, E., Dimitriadis, Y., and Gašević, D. (2021). A collaborative learning approach to dialogic peer feedback: a theoretical framework. *Assess. Eval. High. Educ.* 46, 586–600. doi: 10.1080/02602938.2020.1786497
- Gamlem, S. M., and Vattoy, K.-D. (2023). "Feedback and classroom practice" in *International encyclopedia of education*. eds. R. J. Tierney, F. Rizvi and K. Ercikan, vol. 13. 4th ed (Elsevier), 89–95.
- Gardner, J., O'Leary, M., and Yuan, L. (2021). Artificial intelligence in educational assessment: 'breakthrough? Or buncombe and ballyhoo?'. *J. Comput. Assist. Learn.* 37, 1207–1216. doi: 10.1111/jcal.12577
- Gillani, N., Eynon, R., Chiabaut, C., and Finkel, K. (2023). Unpacking the "Black Box" of AI in Education. *J Educ Techno Soc*, 26, 99–111.
- Griffin, P., and Care, E. (2014). *Assessment and teaching of 21st century skills: methods and approach*. Springer Dordrecht: Springer.
- Gutierrez, F., and Atkinson, J. (2011). Adaptive feedback selection for intelligent tutoring systems. *Expert Syst. Appl.* 38, 6146–6152. doi: 10.1016/j.eswa.2010.11.058
- Haenlein, M., and Kaplan, A. (2019). A brief history of artificial intelligence: on the past, present, and future of artificial intelligence. *Calif. Manag. Rev.* 61, 5–14. doi: 10.1177/0008125619864925
- Halai, A., Sarungi, V., and Hopfenbeck, T. N. (2022). Teachers' perspectives and practice of assessment for learning in classrooms in Tanzania. *Int. Encycl. Educ.* 63-72. doi: 10.1016/B978-0-12-818630-5.09039-4
- Hattie, J. A. (2009). *Visible learning: a synthesis of over 800 meta-analyses relating to achievement*. London and New York: Routledge.
- Hill, P. W., and Barber, M. (2014). *Preparing for a renaissance in assessment* Pearson.
- Hopfenbeck, T. N., and Stobart, G. (2015). Large-scale implementation of assessment for learning. *Assess. Educ.* 22, 1–2. doi: 10.1080/0969594X.2014.1001566
- Hopfenbeck, T. N., Flórez Petour, M. T., and Tolo, A. (2015). Balancing tensions in educational policy reforms: large-scale implementation of assessment for learning in Norway. *Assess. Educ.* 22, 44–60. doi: 10.1080/0969594X.2014.996524
- Humphry, S. M., and Heldsinger, S. A. (2014). Common structural design features of rubrics may represent a threat to validity. *Educ. Res.* 43, 253–263. doi: 10.3102/0013189X14542154
- Johnston, S.-K. (2023). *Privacy considerations of using social robots in education: policy recommendations for learning environments*. United Nations, Department of Economics and Social Affairs, Sustainable Development.
- Jönsson, A., and Panadero, E. (2016). "The use and Design of Rubrics to support assessment for learning" in *Scaling up assessment for learning in higher education*. eds. D. Carless, S. M. Bridges, C. K. Y. Chan and R. Golfcheski (York: Springer), 99–111. (<https://www.pearson.com/content/dam/one-dot-com/one-dot-com/global/Files/about-pearson/innovation/open-ideas/PreparingforaRenaissanceinAssessment.pdf>)
- Ke, Z., and Ng, V. (2019). Automated Essay Scoring: A Survey of the State of the Art. *Paper presented at the 28th International Joint Conference on Artificial Intelligence*.
- King, P. M., and Kitchener, K. S. (1994). *Developing reflective judgment: understanding and promoting intellectual growth and critical thinking in adolescents and adults*. San Francisco: Jossey-Bass.
- Kuhn, D., and Weinstock, M. (2004). "What is epistemological thinking and why does it matter?" in *Personal epistemology: the psychology of beliefs about knowledge and knowing*. eds. B. K. Hofer and P. R. Pintrich (New York: Routledge), 121–144.
- Lee, A. V. Y. (2021). Determining quality and distribution of ideas in online classroom talk using learning analytics and machine learning. *Educ. Technol. Soc.* 24, 236–249.
- Lee, A. V. Y. (2023). Supporting students' generation of feedback in large-scale online course with artificial intelligence-enabled evaluation. *Stud. Educ. Eval.* 77:101250. doi: 10.1016/j.stueduc.2023.101250
- Li, C., Xing, W., and Leite, W. (2022). Using fair AI to predict students' math learning outcomes in an online platform. *Interact. Learn. Environ.* 1-20, 1–20. doi: 10.1080/10494820.2022.2115076
- Li, T., Reigh, E., He, P., and Adah Miller, E. (2023). Can we and should we use artificial intelligence for formative assessment in science? *J. Res. Sci. Teach.* 60, 1385–1389. doi: 10.1002/tea.21867
- van der Linden, W. J., and Glas, C. A. (2010). *Elements of adaptive testing*. New York: Springer.
- Linn, M. C., Gerard, L., Ryoo, K., McElhaney, K., Liu, O. L., and Rafferty, A. N. (2014). Computer-guided inquiry to improve science learning. *Science* 344, 155–156. doi: 10.1126/science.1245980
- Lipnevich, A. A., and Smith, J. K. (2018). *The Cambridge handbook on instructional feedback* Cambridge University Press.
- Liu, N. F., and Carless, D. (2006). Peer feedback: the learning element of peer assessment. *Teach. High. Educ.* 11, 279–290. doi: 10.1080/13562510600680582
- Liu, O. L., Rios, J. A., Heilman, M., Gerard, L., and Linn, M. C. (2016). Validation of automated scoring of science assessments. *J. Res. Sci. Teach.* 53, 215–233. doi: 10.1002/tea.21299
- Luaces, O., Díez, J., and Bahamonde, A. (2018). A peer assessment method to provide feedback, consistent grading and reduce students' burden in massive teaching settings. *Comput. Educ.* 126, 283–293. doi: 10.1016/j.compedu.2018.07.016
- Luckin, R., and Holmes, W. (2016). *Intelligence unleashed: an argument for AI in education*. Available at: <https://discovery.ucl.ac.uk/id/eprint/1475756>
- Ma, W., Adesope, O. O., Nesbit, J. C., and Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: a meta-analysis. *J. Educ. Psychol.* 106, 901–918. doi: 10.1037/A0037123
- McMillan, J. H. (2013). *SAGE handbook of research on classroom assessment*. Los Angeles: Sage.
- Milano, S., McGrane, J. A., and Leonelli, S. (2023). Large language models challenge the future of higher education. *Nat. Mach. Intell.* 5, 333–334. doi: 10.1038/s42256-023-00644-2
- Milligan, S. K., and Griffin, P. (2016). Understanding learning and learning design in MOOCs: a measurement-based interpretation. *J. Learn. Analyt.* 3, 88–115. doi: 10.18608/jla.2016.3.25
- Mousavinasab, E., Zarifsanaiy, N., Niakan Kalhori, S. R., Rakhshan, M., Keikha, L., and Ghazi Saeedi, M. (2021). Intelligent tutoring systems: a systematic review of characteristics, applications, and evaluation methods. *Interact. Learn. Environ.* 29, 142–163. doi: 10.1080/10494820.2018.1558257
- Murphy, R. F. (2019). Artificial intelligence applications to support K-12 teachers and teaching. *Rand Corp.* 10, 1–20. doi: 10.7249/PE315
- Nicol, D., and MacFarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Stud. High. Educ.* 31, 199–218. doi: 10.1080/03075070600572090
- Panadero, E., and Jönsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: a review. *Educ. Res. Rev.* 9, 129–144. doi: 10.1016/j.edurev.2013.01.002
- Panadero, E., Andrade, H., and Brookhart, S. (2018). Fusing self-regulated learning and formative assessment: a roadmap of where we are, how we got here, and where we are going. *Aust. Educ. Res.* 45, 13–31. doi: 10.1007/s13384-018-0258-y
- Perry, N., Lisaingo, S., Yee, N., Parent, N., Wan, X., and Muis, K. (2020). Collaborating with teachers to design and implement assessments for self-regulated learning in the context of authentic classroom writing tasks. *Assess. Educ.* 27, 416–443. doi: 10.1080/0969594X.2020.1801576
- Popham, W. J. (1997). What's wrong - and what's right - with rubrics. *Educ. Leadersh.* 55, 72–75.
- Ratnam-Lim, C. T. L., and Tan, K. H. K. (2015). Large-scale implementation of formative assessment practices in an examination-oriented culture. *Assess. Educ.* 22, 61–78. doi: 10.1080/0969594X.2014.1001319
- Ruiz-Primo, M., and Brookhart, S. (2018). *Using feedback to improve learning*. Routledge.
- Scalise, K., Wilson, M., and Gochyyev, P. (2021). A taxonomy of critical dimensions at the intersection of learning analytics and educational measurement. *Front. Educ.* 6:656525. doi: 10.3389/feduc.2021.656525
- Schwartz, S. (2020). Flood of online learning resources overwhelms teachers. *Educ. Week*. Available at: March 25, 2020: <https://www.edweek.org/teaching-learning/flood-of-online-learning-resources-overwhelms-teachers/2020/03>
- Shin, J., Chen, F., Lu, C., and Bulut, O. (2022). Analyzing students' performance in computerized formative assessments to optimize teachers' test administration decisions using deep learning frameworks. *Journal of Computers in Education* 9, 71–91. doi: 10.1007/s40692-021-00196-7
- Skeat, J., and Ziebell, N. (2023). University students are using AI, but not how you think. Available at: <https://pursuit.unimelb.edu.au/articles/university-students-are-using-ai-but-not-how-you-think>
- Stobart, G., and Hopfenbeck, T. (2014). "Assessment for learning and formative assessment" in *State of the field review assessment and learning*. eds. J.-A. Baird, T. Hopfenbeck, P. Newton, G. Stobart and A. Steen-Uthheim (Oxford: Norwegian Knowledge Centre for Education).
- Sun, S. Z. (2021). Epistemological beliefs: the key to enhance critical thinking for higher education students in the east. Paper presented at the 2021 American Educational Research Association (AERA) annual meeting.
- Sun, Z. S. (2023). *Developing and validating an operationalisable model for critical thinking assessment in different cultures*. The University of Melbourne.
- Swiecki, Z., Khosravi, H., Chen, G., Martinez-Maldonado, R., Lodge, J. M., Milligan, S., et al. (2022). Assessment in the age of artificial intelligence. *Comp. Educ.* 3:100075. doi: 10.1016/j.caeai.2022.100075

- Tashu, T. M., and Horvath, T. (2019). Semantic-based feedback recommendation for automatic essay evaluation. *Proceedings of SAI Intelligent Systems Conference* 334–346. London: Springer.
- Tomasik, M. J., Berger, S., and Moser, U. (2018). On the development of a computer-based tool for formative student assessment: epistemological, methodological, and practical issues. *Front. Psychol.* 9:2245. doi: 10.3389/fpsyg.2018.02245
- Verma, M. (2018). Artificial intelligence and its scope in different areas with special reference to the field of education. *Int. J. Adv. Educ. Res.* 3, 5–10.
- Vincent-Lamarre, P., and Larivière, V. (2021). Textual analysis of artificial intelligence manuscripts reveals features associated with peer review outcome. *Quant. Sci. Stud.* 2, 662–677. doi: 10.1162/qss_a_00125
- von Davier, A. A., Mislevy, R. J., and Hao, J. (2022). *Computational psychometrics: New methodologies for a new generation of digital learning and assessment: With examples in R and Python*. Springer.
- Webb, M., Gibson, D., and Forkosh-Baruch, A. (2013). Challenges for information technology supporting educational assessment. *J. Comput. Assist. Learn.* 29, 451–462. doi: 10.1111/jcal.12033
- Wen, H., Liu, Y., and Zhao, N. (2020). Longitudinal cognitive diagnostic assessment based on the HMM/ANN model. *Front. Psychol.* 11:2145. doi: 10.3389/fpsyg.2020.02145
- Wertz, M. H. (2019). *Epistemological developmental level and critical thinking skill level in undergraduate university students*. University of South Florida.
- Wiliam, D. (2011). What is assessment for learning? *Stud. Educ. Eval.* 37, 3–14. doi: 10.1016/j.stueduc.2011.03.001
- Wiliam, D., Lee, C., Harrison, C., and Black, P. (2004). Teachers developing assessment for learning: impact on student achievement. *Assess. Educ.* 11, 49–65. doi: 10.1080/0969594042000208994
- Wylie, C. E., and Lyon, C. J. (2015). The fidelity of formative assessment implementation: issues of breadth and quality. *Assess. Educ.* 22, 140–160. doi: 10.1080/0969594X.2014.990416
- Wollenschläger, M., Hattie, J., Machts, N., Möller, J., and Harms, U. (2016). What makes rubrics effective in teacher-feedback? Transparency of learning goals is not enough. *Contemp. Educ. Psychol.* 44–45, 1–11. doi: 10.1016/j.cedpsych.2015.11.003
- Zhai, X., and Nehm, R. H. (2023). AI and formative assessment: the train has left the station. *J. Res. Sci. Teach.* 60, 1390–1398. doi: 10.1002/tea.21885