Jiang, B. et al. (Eds.) (2025). Proceedings of the 33rd International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Empowering, Not Replacing: Using Generative AI to Coach Educators in Providing Effective Feedback

Antonette SHIBANI[a*], Jayakrishnan WARRIEM[b], Lisa-Angelique LIM[c], Gnana BHARATHY[d] & Arup CHATTOPADHYAY[e]
[a]*Transdisciplinary School, University of Technology Sydney, Australia*
[b]*NPTEL, IIT Madras, India*
[c]*Connected Intelligence Centre, University of Technology Sydney, Australia*
[d]*Faculty of Engineering & Information Technology, University of Technology Sydney, Australia*
[e]*BS Program, IIT Madras, India*
*Antonette.shibani@uts.edu.au

Abstract: As generative AI (GenAI) technologies become increasingly integrated into educational contexts, the importance of supporting rather than replacing educators with AI is gaining prominence. This study introduces Feedback Tutor, a GenAI-powered tool designed to enhance teacher feedback literacy by coaching educators in crafting effective feedback, while preserving their agency. It explores how human-AI collaboration can empower educators through a teacher professional development intervention using Feedback Tutor. Eleven educators from two institutions participated in a workshop and follow-up interviews, with qualitative data analyzed thematically. Four key themes emerged: (1) *AI-supported professional growth*, where educators valued the tool's coaching capabilities for improving clarity and feedback quality; (2) *Human-AI Negotiation and Agency*, highlighting the tool's role as a collaborative partner rather than a replacement; (3) *Socio-Technical Integration*, emphasizing the importance of aligning AI tools with existing educational workflows and infrastructure; and (4) *Ethical and Pedagogical Considerations*, underscoring the need for balanced feedback that supports student autonomy and learning. Findings suggest that GenAI tools like Feedback Tutor can significantly enhance educators' feedback practices and professional growth when designed with pedagogical intent and ethical sensitivity. However, challenges such as lack of content ingestion, and usability constraints must be addressed to optimize adoption and impact. This study contributes to the growing discourse on AI in education by demonstrating how thoughtfully designed AI tools can augment, rather than diminish, the human elements of teaching and learning.

Keywords: educator, GenAI, feedback, feedback tutor, teacher feedback literacy, empowering, artificial intelligence, teacher professional development

1. Introduction

Generative AI (GenAI) has been touted as a powerful tool to transform educational practices. Across various contexts, GenAI has demonstrated the potential to enhance student learning and assist educators in the teaching process. This includes the provision of real-time feedback for personalized support that can help improve students' writing and save teachers' time by allowing them to focus on other aspects of teaching (Kasneci et al., 2023). Indeed, feedback is an area where there is an identified need for better support so students can reflect on their academic performance and improve over time. Sufficiently detailed, high-quality comments that are usable and tailored to students work can help them improve (Dawson et al., 2019), but is time-consuming for the educator to provide. Many students report issues on the feedback they receive, which include the lack of specificity, delayed delivery, limited guidance on how to apply it in future tasks, difficulty in interpretation, and potentially affecting students' self-perception and confidence (Carless & Boud, 2018; Winstone et al., 2017).

While it is important for students to learn to make good sense of feedback provided and act upon it (Student Feedback Literacy - (Carless & Boud, 2018)), and even, develop specific skills to critically engage with automated feedback (Automated feedback literacy - (Shibani et al., 2022)), it is equally important to support *teachers' feedback literacy* due to their intertwined nature (Carless & Winstone, 2020). This is because, for students to get value out of feedback, teachers must provide effective feedback. While there is no single definition of good feedback practice, commonly agreed upon principles guide them (Nicol & Macfarlane-Dick, 2006). Novice educators are often learning to provide feedback, have limited pedagogical training, and can miss many components of such effective feedback. This can lead to concerns navigating the design, relational, and pragmatic aspects of teacher feedback literacy (Carless & Winstone, 2020). Supporting educators to learn to provide good feedback for students is an important step towards fostering students' feedback literacy.

Technology mediated feedback has seen a steady growth in literature, allowing feedback to be a continuous, real-time part of student work. Automated systems and learning analytics tools have demonstrated impact at scale due to their ability to provide feedback to large cohorts at a personalized level (Knight, Shibani, et al., 2020; Pardo et al., 2019). With advances from GenAI, it is no surprise that feedback was a viable candidate for AI to offer support. Recent works explore the use of GenAI to provide feedback on many types of tasks, including on writing (Lee et al., 2024), computer programming (Balse et al., 2023), and content creation (Pozdniakov et al., 2025) often in par with humans, with tools such as Cogniti offering AI feedback tailored to assignment rubric set up by educators in higher education (Huynh et al., 2024). However, it also raises the question of whether AI will support educators or de-skill them in the future as they turn towards AI for writing their feedback messages, without sufficient experience and skill development of their own.

Teacher Professional Development (TPD) for AI thus becomes increasingly important as educators are now required to work alongside hybrid intelligent systems (Molenaar, 2022) and the TPD should help in getting educators AI ready (Luckin et al., 2022). Long and Magerko's (2020) framework for AI literacy emphasizes that educators need opportunities to develop critical evaluation skills and collaborative approaches to working alongside AI tools. Similarly, the Ethical AI readiness framework (Luckin et al., 2022) underscores the importance of TPD that empowers teachers to make informed decisions about AI implementation rather than merely training them as end-users. These approaches recognize that educators must maintain agency in human-AI partnerships, positioning them as active designers and critical implementers rather than passive consumers of AI technologies.

The integration of GenAI thus presents a unique opportunity to enhance, rather than replace, the critical role of educators in providing feedback. While concerns about AI potentially diminishing the human element in education are valid, this technology can be harnessed to empower educators by augmenting their ability to deliver more effective, rubric-aligned feedback, ensuring that students receive the support they need to improve their work. Our study introduces a GenAI-support tool, 'Feedback Tutor', along with an intervention that helps educators improve their written feedback by providing guiding feedback on their written feedback, while also empowering their own knowledge with principles of good feedback practice and human-AI collaboration. We explore the research question: *What affordances of Feedback Tutor can empower educators to improve their feedback practices, and influence changes?*

2. Background

2.1 Feedback

Feedback has long been acknowledged as a critical factor in improving students' learning and performance. In this paper, we define feedback as a dialogical process, "whereby learners make sense of information from various sources and use it to enhance their work or learning strategies" (Carless & Boud, 2018) (p.1). Effective feedback needs to follow established principles; namely: 1) specify clear performance standards; 2) facilitate self-assessment; 3) communicate high-quality information for students to bridge learning gaps; 4)

encourage dialogue; 5) support motivation and foster self-efficacy; 6) provide opportunities to close the gap between current and expected performance (Nicol & Macfarlane-Dick, 2006). To this end, feedback needs to be timely, personalized, actionable and directed at students' self-regulated learning, not just the outcome or performance (Hattie & Timperley, 2007). The provision of such effective feedback, however, remains challenging in higher education contexts. Scaling personalized feedback to accommodate large cohorts presents a significant burden on faculty members who must simultaneously balance substantial teaching responsibilities with administrative obligations and research requirements (Henderson et al., 2019; Sembey et al., 2024). Consequentially, empirical research points to students' low engagement and satisfaction with their feedback (e.g., (Winstone et al., 2017)).

Furthermore, given the process-oriented nature of feedback, feedback literacy has become increasingly recognized as important – both teacher feedback literacy and student feedback literacy, which are intertwined (Carless & Winstone, 2020), as both teachers and students are jointly responsible for feedback to be effective. While a number of frameworks exist, student feedback literacy broadly refers to students' understandings and appreciation of feedback, their ability to evaluate the work of themselves as well as of others based on feedback, their ability to manage emotional responses to feedback, and drawing on a range of strategies to act on feedback (Carless & Boud, 2018). On the other hand, teacher feedback literacy refers to the "knowledge, expertise, and dispositions to design feedback processes in ways which enable student uptake of feedback and seed the development of student feedback literacy" (Carless & Winstone, 2020) (p. 4). Based on previous work involving interviews and focus groups with teachers about their assessment and feedback practices, Boud & Dawson (2023) further identified what feedback literate teachers actually do with respect to three distinct levels. The *macro* level relates to how feedback is built into programme design and development; the *meso* level relates to how feedback is purposely designed in specific modules or units; while the *micro* level relates distinctly to feedback practices involved in students' work. In this study, we focus on teacher feedback literacy at the micro level, with a view to developing teacher feedback literacy, because teacher feedback literacy is fundamental for fostering student feedback literacy within educational contexts.

2.2 Human-AI interaction paradigm

The humanistic approach to AI in education advocates for AI being used to facilitate higher order thinking, human interaction, and human values rather than diminishing them (Carvalho et al., 2022). This implies that the goal should be to build hybrid systems that augment rather than replace human intelligence, leveraging human strength and compensating for its weaknesses (Molenaar, 2022). Such hybrid collaborations will cover a spectrum of use-cases with varying contributions and role of AI, with the extremes of AI being the complete driver to AI being a mere assistant (Banihashem et al., 2025).

There have been numerous examples of AI systems being used for supporting teachers with data driven decisions like dashboards, providing recommendations on instructional strategies and curriculum, and assessment support. With GenAI tools maturing, there is an increase in use cases of conversational AI systems (such as chat bots) that support teaching-learning scenarios including providing direct feedback to students and assisting in teacher professional development using a variety of pedagogical techniques (Tan et al., 2025).

These technological developments directly inform the evolving needs within TPD, where empowering educators to effectively collaborate with AI has become essential. Two key concepts have emerged in this context: AI Literacy and AI readiness. Long and Magerko (2020) had defined AI literacy as a competency required for individuals to critically evaluate AI technology, communicate and collaborate with AI systems, and use AI as a tool across professional and personal spaces. Complementarily, AI readiness refers to the ability of the teacher to understand, in non-technical terms, the capabilities of AI - enabling them to make informed decisions about the use and procurement of AI in educational settings (Luckin et al., 2022). TPD will be essential to fundamentally address these competencies and ensure the effectiveness and sustainability of training benefits.

To systematically develop these competencies among educators, several frameworks have been proposed that integrate the theoretical understanding about TPD along with its practical implementation strategies. The i-TPACK framework (Dogan et al., 2025) extends the TPACK model to incorporate AI specific knowledge domains. The ETHICAL AI readiness framework (Luckin et al., 2022) provides a 7-step structure for building an AI training and apply AI readiness thinking. In the specific context of AI supported feedback, the pedagogical framework for hybrid intelligent feedback by Banihashem et al (2025) outlines a structured process for integrating human and GenAI contributions in providing feedback.

A review of the research highlights a lack of attention to the critical role of teachers and their professional development needs in the context of AI integration, especially in enhancing teachers' AI literacy (Tan et al., 2025). This gap becomes more pronounced when considering hybrid Human-AI systems, that remain in early stages of research (Molenaar, 2022). There is a clear need for pedagogical guidelines for both AI systems and educators to follow while implementing AI feedback in their teaching-learning settings. To align with the humanistic approach, it is important to explore how roles can dynamically shift between human and AI based on task complexity, what educators find beneficial in these contexts, and identify areas where human strengths can be meaningfully leveraged - forming the foundation of our work.

2.3 Contextualised interventions

Automated tools and AI technologies can help enhance teaching and learning processes by offering personalized support at scale to learners and educators. However, their adoption is often determined by factors beyond just tool efficiency. For maximum impact, the technical aspects must be well aligned with the pedagogical aspects (Knight, Gibson, et al., 2020), which include how the tool is implemented in authentic classrooms, and how it is tied to learning outcomes and assessment practices of existing curricula for student facing tools (Shibani et al., 2019). For TPD, educators require support for adopting technologies in their teaching practices (Nelson et al., 2019). Importantly, the effective use of automated feedback tools requires teachers to demonstrate teacher feedback literacy competencies, including those at the micro-level (Buckingham Shum et al., 2023). Thus, the adoption of teaching and learning technologies is not just a technical endeavour, but a socio-technical one.

A review of intervention studies found that technologies for specific instructional purposes were more frequently adopted than general purpose technologies by educators (Wu et al., 2013). Tools and practices contextualized for specific settings also provide alignment between the technical and pedagogical aims (Shibani et al., 2019) and can be extended to the affordances of GenAI. While general purpose GenAI tools such as ChatGPT can assist educators in optimising standard responsibilities (Hashem et al., 2024), interventions designed using GenAI capabilities for specific purposes can offer targeted support. Our work introduces a specialised GenAI powered tool 'Feedback Tutor' and intervention that can empower educators by supporting effective feedback practices and fostering teacher feedback literacy. Such interventions can offer professional development opportunities for educators by not merely functioning as tools but also acting as catalysts for reflective teaching and improved teaching practices. In the following section, we introduce the design of Feedback Tutor, its underlying technology, and implementation of the tool in the intervention.

3. Feedback Tutor Tool and Intervention for Teacher Feedback Literacy

3.1 Tool features

Feedback Tutor was designed to have an intuitive chat interface where educators can type in their feedback and receive suggestions for improvement. They can select an option from two types of assignment at the start: A. Writing assignment and B. Programming assignment, reflective of the most common types of feedback provided in the two institutional contexts. The tool first asks the educator to describe the context of the assignment including rubric details if available. Then educators are asked to type their draft feedback. The tool responds by suggesting specific improvements for the written feedback, also including code if

needed for programming assignments. Importantly, in addition to suggesting revisions, the tool provides coaching to the educator, empowering them to write better feedback by themselves in the future. Coaching includes the rationale for the suggestions and an explanation of why it would be helpful to students in improving their work, grounded in effective feedback principles from literature (Nicol & Macfarlane-Dick, 2006). Each chat also has features to regenerate feedback if needed, like/ dislike, and users can download a copy of the feedback as a PDF file. Users can initiate any number of chat threads as needed, which are saved in their log for later retrieval, as required. Figure 1 provides a screenshot showcasing a sample chat in Feedback Tutor interface. Additionally, the tool has functionality for admin users to view analytics and logs of the tool usage in a dashboard (these admin features are not intended for educators and hence are not the focus of the current study).

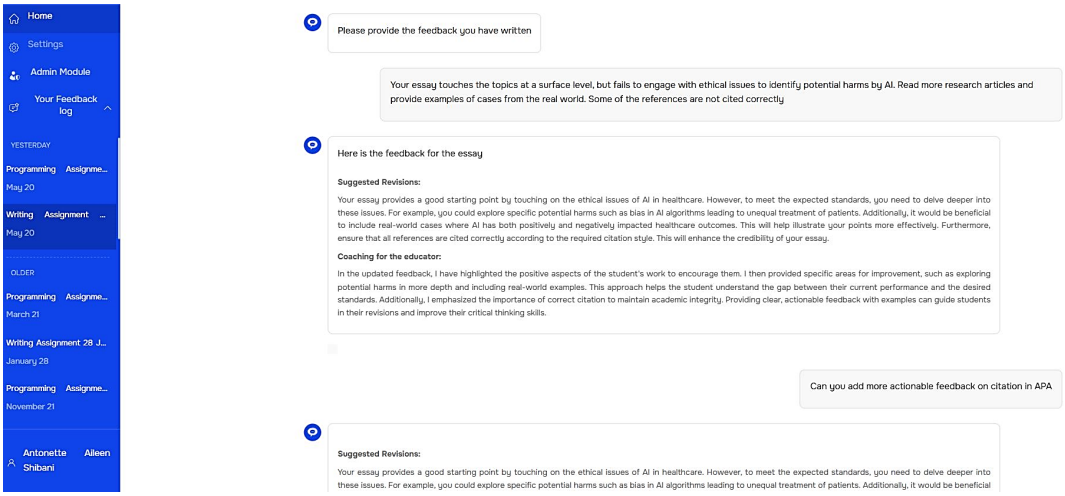

Figure 1. Screenshot of the Feedback Tutor interface for educators with a sample chat

3.2 Technical implementation

Feedback Tutor was implemented using Microsoft Azure tools and could be accessed via a URL[1]. It was developed as a React app and used the GPT-4 model for providing feedback which was cost-effective while delivering good results. The tool was deployed using Docker on Azure cloud services. Iterative testing and changes were made to the prompts provided to the large language model (LLM) to guide its feedback which specifically included instructions to preserve the original voice of the feedback creator to maintain authenticity and to provide coaching in addition to suggesting improvements to feedback. The full source code of the tool has been made available as an open-source resource on Github[2]. This includes the LLM prompts used for the writing and programming assignment feedback sections, and the detailed technical architecture of the tool.

3.3 Intervention design

In line with the socio-technical nature of AI technologies, the intervention was designed and delivered as a two-hour online professional development workshop for educators from two institutions located in different countries. The workshop consisted of introductory topics including the definitions of feedback (from academic and practical viewpoints), summative and formative feedback mechanisms, strategies to deliver them effectively such as the sandwich approach, and knowledge sharing activities among the participants using short polls designed in Mentimeter[3], facilitated by the research team in the first part. Participants were introduced

[1] https://feedbacktutorweb.azurewebsites.net
[2] https://github.com/jkmpod/feedback-tutor
[3] https://www.mentimeter.com/

to Feedback Tutor's main features, and learned how the tool could be used to improve their feedback in the second part. All participants were previously instructed to bring 2-3 samples of draft feedback they had written to try out the tool. They were provided with instructions for stimulated recall and recording as they explored the tool independently in the last part.

4. Methodology

The study employs qualitative methods to understand educators' perspectives of the Feedback Tutor tool and intervention. A total of eleven participants volunteered to participate in the online workshop and in the follow-up interview and were provided with a monetary incentive for contributing to the study. The follow-up one-on-one stimulated recall interview was conducted with each participant within 2-4 weeks after the workshop, which was guided by the recording from their trial of the tool. Each interview lasted between 45 minutes and 1 hour. Analysis for the current study uses interview data from the eleven participants - P1 to P8 from Institution 1, and P9 to P11 from Institution 2. While all of them had provided feedback and assessment in their institutional contexts, they had varied levels of teaching experience.

The transcripts were cleaned and anonymized for inductive thematic analysis. The analysis was performed by two large language model (LLM) based tools and three human coders using a hybrid approach (Barany et al., 2024). The following prompt was used in Google NotebookLM and Microsoft Co-pilot (with institutional access for secure data analysis): *"Analyse transcript.docx - do a thematic analysis of the interview with P[]. Extract the themes that emerge and also provide the supporting quotes (verbatim) for the same. Where multiple quotes exist, provide all of them".* All LLM outputs were verified for accuracy and evaluated by two researchers. They were further compared with a third researcher's own analysis to add missing quotes and re-word headings. Neither the researchers nor the LLM adopted a positionality; instead, they aimed to present all key themes that emerged from the data.

5. Findings and discussion

To answer our research question, we examined how the affordances of Human-AI collaborative tools such as Feedback Tutor can empower educators by enhancing their ability to deliver effective feedback. Our thematic analysis identified several factors, which were organized into four main themes that influence educator feedback practices when mediated through the given GenAI tool.

Theme 1: AI-supported professional growth

The theme around educators valuing professional growth using AI support emerged prominently, alongside other beneficial features of the tool. Participants consistently valued the "Coaching for Educators" aspect embedded within the tool, emphasizing its significant role in improving their feedback practices. This finding aligns with the work of Graßmann & Schermuly (2021), who emphasize that AI coaching is most effective when used in conjunction with human oversight, allowing feedback to be more precisely tailored to individual educational needs.

P3 notably highlighted this as impactful, stating, *"the coaching for educator part… helped me improve my feedback clearly and concisely"*. This capability enabled educators to explicitly visualize and refine feedback principles, thereby supporting their professional growth. P8 appreciated the tool's ability to highlight their own missteps, remarking, *"tells what I did wrong. So as an educator I also did something wrong"*. This meta-feedback aspect was consistently highlighted as a distinctive feature – setting the tool apart from general LLMs, which are often used to provide direct evaluations and feedback rather than supporting educator reflection.

Educators found elements such as language enhancement and error identification particularly beneficial. Participants appreciated controlled augmentations - for example, *"did some wording changes to make it more coherent"* (P10), *"… grammatical error .., missing something …"* (P5) - highlighting the utility of the tool in improving clarity and coherence in feedback delivery (Graßmann & Schermuly, 2021).

Educators also appreciated how the tool improved the structure and organization of their feedback, making it more accessible and digestible for students. For instance, P2 noted, *"… the flow of the feedback to make it more... understandable to the students"*. Similarly, P9 highlighted the benefit of structured formatting compared to typical paragraph style used by instructors: *"…. the feedback [was] … divided between themes …, so it's easier to read…. It streamlines your text [whereas] … I tend to be a very verbose… summarizes better what I wanted to say in the feedback*".

In addition, participants valued the tool's ability to provide clear examples, which made their feedback more concrete and actionable. As stated by P2, *"It actually gave examples in a clear manner, …"* and P5 *"… it gives more example, … giving correct code also"*.

However, some educators raised concerns about potential overreach by AI, as the tool occasionally provided overly explicit solutions in its feedback, inadvertently limiting students' opportunities for critical thinking (P4). Thus, while clearly beneficial, careful balance is necessary to ensure educators' agency and pedagogical intentions remain intact in AI-assisted feedback. Indeed, the ability to recognize and navigate these concerns is a key part of developing teacher feedback literacy (Buckingham Shum et al., 2023) and AI literacy (Long & Magerko, 2020). The intervention supported this development by enabling educators to learn how to engage effectively in hybrid human-AI settings.

Theme 2: Human-AI Negotiation and Agency

The Human-AI Negotiation and Agency theme reflected the dynamic partnership educators desired to have with AI – favoring active collaboration over passive consumption of AI-generated feedback. Educators appreciated the tool's ability to enhance their quality of feedback without diminishing their authority or personal pedagogical style, a design consideration explicitly programmed in the LLM prompt of the tool as discussed in Section 3.1. As P8 observed, *"The tool evolves my way to provide feedback… without replacing my role".* This sentiment was echoed by many participants, who viewed (and would like to view) the technology as a collaborator rather than a substitute (Graßmann & Schermuly, 2021).

However, concerns about the extent of augmentation were also expressed. For instance, P2 noted "*Sometimes … I felt that it actually kind of removed … my personal elements to the feedback",* while P11 remarked *"…sometimes adding some extra word[s] that was not expected. …".*

Several participants described engaging in an active, iterative process of evaluating and selectively adopting AI-generated suggestions. For example, P8 shared, *"I'll regenerate two three times and I'll pick the best one"* and P11 added, *"… I have to write 4 to 5 words. And, Gen AI will provide between 30, 40 words in a 3, 4 sentence that could be".* This selective engagement illustrates the educators' desire for control and intentionality in shaping AI-augmented feedback. It also highlights a new affordance offered by generative AI – the ability to engage in a dialogic process - effectively utilized by educators to co-construct meaningful feedback relevant to their contexts.

Nonetheless, challenges remain – particularly regarding the lack of AI's conversational memory, particularly the absence of session memory or iterative dialogue functionalities, which hindered continuity in interactions and forced repetitive engagement, leading to frustrations that P1 articulated as: "*Lack of conversational memory forces repeated manual interactions; it's cumbersome*". These frustrations point to a critical need for improved conversational continuity to support more seamless and effective human-AI collaboration, which was later developed and offered as a feature of the tool (explained in the next section).

Theme 3: Socio-Technical Integration

The need for Socio-Technical Integration was consistently highlighted, emphasizing that the effectiveness of GenAI tools depends significantly on their compatibility with existing educational workflows, practices and technological infrastructures. The potential for integration with Learning Management Systems (LMS) (P9) and platforms like Discourse (P1, P3, P5) was seen as a critical facilitator for adoption and usability, promising substantial time savings - an aspect also emphasized in Hashem et al (2024). The potential for improved workflow through alignment with pre-existing systems supports broader arguments by Knight

et al. (2020), who asserted that successful technology adoption in education hinges on its integration into familiar pedagogical settings.

However, educators raised concerns over limitations related to content ingestion in the form of documents or images, retention, and the tool's ability to remember previous interactions - all perceived as obstacles to usability. The inability to upload student work limited the tool's capacity to tailor feedback to specific assignments. Similarly, the absence of conversation history prevented users from continuing discussions on related topics. For instance, P1 noted, *"The main part is memory. So, we don't have to copy paste when we want any improvement in chat GPT. Here, we have to do that"*. P10 wanted to *"ask more questions"* after getting a revision and wished for *"the ability to continue the chat"* to refine irrelevant content. While at the time of this trial run, the memory feature was not available. However, to respond to multiple educators' concerns regarding the lack of conversation history, this feature is implemented in the current version of Feedback Tutor. Now, users can extend the conversation further by asking follow-up questions on the feedback, which are answered by the LLM's knowledge base.

P10 also identified a main limitation as the inability to *"continue same chat for the current feedbacks belonging to same assessment"* which meant having to *"start another chat to test the other feedback".* This process of going back to "home" for each new feedback was *"not really user friendly"*. Participants explicitly noted the inconvenience caused by interface designs requiring extensive tab-switching, underscoring that a more seamless integration would significantly enhance educators' workflow efficiency, a key aspect of effective edtech integration (Nelson et al., 2019).

Theme 4: Ethical and Pedagogical Considerations

The analysis revealed an essential dimension concerning Ethical and Pedagogical Considerations in the use of GenAI tools for educators. Educators appreciated the tool's capacity to promote structured, clear, and pedagogically sound feedback. This kind of feedback was seen as fostering student self-regulation and agency– a balanced approach to feedback aligned with effective formative feedback principles (Nicol & Macfarlane-Dick, 2006). As P4 reflected, *"The suggestions improved how I guide students without giving direct answers"*.

However, concerns were raised about the accuracy and validity of AI-generated feedback. For example, P6 questioned the reliability of outputs asking, *"How are we validating that the responses generated are always correct?"* Such concerns point to the need for transparent validation mechanisms and ongoing oversight when integrating GenAI tools into pedagogical contexts given the inconsistent nature of LLM outputs.

Participants also raised ethical concerns around the level of detail in AI-generated feedback, fearing it could undermine student autonomy and reduce their engagement with the learning material. They expressed caution about overly prescriptive responses, which may discourage independent thinking. As P4 noted, *"The complete code snippet is given…. …but we don't tell them how to do it"*, illustrating the risk of bypassing the learning process in favor of efficiency.

Additionally, a few participants discussed constraints linked to content uploads and contextual continuity – issues connected to broader socio-technical and ethical considerations of GenAI tools. While the ability to enhance conversational context by uploading assignments was widely recognized as a desirable feature, its absence was not due to technical shortcomings, but rather deliberate design decisions based on two key factors: First, ethics approval did not permit the uploading of student work, as students had not consented to having their assignments or work shared with AI systems. Secondly, the tool was intentionally designed to encourage educators to develop their capacity to write feedback independent of AI support, reinforcing the research objective of empowering their feedback practices through TPD. Timesaving was not the tool's primary objective. This underscores a fundamental tension between usability demands and ethical boundaries and highlights a key insight: efficiency using GenAI tools must not come at the cost of critical engagement or educational integrity.

6. Conclusion

In this study, we presented Feedback Tutor, a GenAI-powered tool supporting TPD intervention designed to develop teachers' feedback literacy. The tool and intervention were specifically created to empower educators with the autonomy to write independent feedback while being coached to enhance their practices, evaluated by 11 educators from two institutional contexts. Findings indicate that human-AI collaboration in feedback practices can positively transform educational pedagogy, support professional learning and improve feedback quality. However, this positive impact depends on addressing key usability issues, maintaining a balance of educator autonomy, ensuring robust socio-technical integration, and articulating the ethical implications of AI-enhanced practices. Future work will explore how teacher agency and reflections of educators evolved after the initial exposure by engaging with the tool over time. It will also unpack in more depth the tensions between efficiency and ethical design of GenAI tools, and trade-offs influencing educator perspectives on use and usefulness of tools. We hope such thoughtful integration of AI can support the development of educator expertise while augmenting, rather than replacing, their professional judgment.

Acknowledgements

We would like to acknowledge the UTS Key Technology Partner (KTP) Funding Programme 2024 for funding this research collaboration between UTS and IITM. Thanks also to the team at Jumpstart Ninja for their rapid tool development and deployment support.

References

Balse, R., Valaboju, B., Singhal, S., Warriem, J. M., & Prasad, P. (2023). Investigating the potential of gpt-3 in providing feedback for programming assessments. Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V. 1,

Banihashem, S. K., Noroozi, O., Khosravi, H., Schunn, C. D., & Drachsler, H. (2025). Pedagogical framework for hybrid intelligent feedback, Innovations. *Education and Teaching International*, 1-17.

Barany, A., Nasiar, N., Porter, C., Zambrano, A. F., Andres, A. L., Bright, D., . . . Baker, R. S. (2024). ChatGPT for Education Research: Exploring the Potential of Large Language Models for Qualitative Codebook Development. In A. M. Olney, I.-A. Chounta, Z. Liu, O. C. Santos, & I. I. Bittencourt, *Artificial Intelligence in Education* Cham.

Boud, D., & Dawson, P. (2023). What feedback literate teachers do: an empirically-derived competency framework. *Assessment & Evaluation in Higher Education*, *48*(2), 158-171.

Buckingham Shum, S., Lim, L.-A., Boud, D., Bearman, M., & Dawson, P. (2023). A comparative analysis of the skilled use of automated feedback tools through the lens of teacher feedback literacy. *International Journal of Educational Technology in Higher Education*, *20*(1), 40.

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

Carless, D., & Winstone, N. (2020). 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

Carvalho, L., Martinez-Maldonaldo, R., Tsai, Y.-S., Markauskaite, L., & De Laat, M. (2022). How can we design for learning in an AI world? *Computers and Education: Artificial Intelligence*, *3*.

Dawson, P., Henderson, M., Mahoney, P., Phillips, M., Ryan, T., Boud, D., & Molloy, E. (2019). What makes for effective feedback: staff and student perspectives. *Assessment & Evaluation in Higher Education*, *44*(1), 25-36. https://doi.org/10.1080/02602938.2018.1467877

Dogan, S., Nalbantoglu, U. Y., Celik, I., & Dogan , N. A. (2025). Artificial intelligence professional development: a systematic review of TPACK, designs, and effects for teacher learning. *Professional Development in Education*, *51*(3), 519-546.

Graßmann, C., & Schermuly, C. C. (2021). Coaching with artificial intelligence: Concepts and capabilities. *Human Resource Development Review*, *20*(1), 106-126.

Hashem, R., Ali, N., El Zein, F., Fidalgo, P., & Khurma, O. A. (2024). AI to the rescue: Exploring the potential of ChatGPT as a teacher ally for workload relief and burnout prevention. *Research & Practice in Technology Enhanced Learning*, *19*.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, *77*(1), 81-112.

Henderson, M., Molloy, E., Ajjawi, R., & Boud, D. (2019). Designing feedback for impact. In *The impact of feedback in higher education: Improving assessment outcomes for learners* (pp. 267-285). Springer.

Huynh, M., Van Den Berg, F., Van Ogtrop, F., Pozza, L., & Kite, J. (2024). *How generative AI can make personalised feedback at scale more consistent and efficient*. The University of Sydney. Retrieved 10 May 2025 from https://educational-innovation.sydney.edu.au/teaching@sydney/how-generative-ai-can-make-personalised-feedback-at-scale-more-consistent-and-efficient/

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., . . . Hüllermeier, E. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, 102274.

Knight, S., Gibson, A., & Shibani, A. (2020). Implementing learning analytics for learning impact: Taking tools to task. *The Internet and Higher Education*, *45*, 100729. https://doi.org/https://doi.org/10.1016/j.iheduc.2020.100729

Knight, S., Shibani, A., Abel, S., Gibson, A., Ryan, P., Sutton, N., . . . Buckingham Shum, S. (2020). AcaWriter: A Learning Analytics Tool for Formative Feedback on Academic Writing. *Journal of Writing Research*. https://doi.org/10.17239/jowr-2020.12.01.06

Lee, M., Gero, K. I., Chung, J. J. Y., Buckingham Shum, S., Raheja, V., Shen, H., . . . Siangliulue, P. (2024). *A Design Space for Intelligent and Interactive Writing Assistants* Proceedings of the CHI24 conference on human factors in computing systems, Honolulu, Hawai'i. arXiv:2403.14117.

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. Proceedings of the 2020 CHI conference on human factors in computing systems, Honolulu, Hawai'i, USA.

Luckin, R., Cukurova, M., Kent, C., & du Boulay, B. (2022). Empowering educators to be AI-ready. *Computers and Education: Artificial Intelligence*, *3*(22).

Molenaar, I. (2022). Towards hybrid human-AI learning technologies. *European Journal of Education*, *57*(4), 632-645.

Nelson, M. J., Voithofer, R., & Cheng, S.-L. (2019). Mediating factors that influence the technology integration practices of teacher educators. *Computers & Education*, *128*, 330-344.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in higher education*, *31*(2), 199-218.

Pardo, A., Jovanovic, J., Dawson, S., Gašević, D., & Mirriahi, N. (2019). Using learning analytics to scale the provision of personalised feedback. *British Journal of Educational Technology*, *50*(1), 128-138.

Pozdniakov, S., Brazil, J., Mohammadi, M., Dollinger, M., Sadiq, S., & Khosravi, H. (2025). AI-assisted co-creation: Bridging skill gaps in student-generated content. *Journal of Learning Analytics*, *12*(1), 129-151.

Sembey, R., Hoda, R., & Grundy, J. (2024). Emerging technologies in higher education assessment and feedback practices: A systematic literature review. *Journal of Systems and Software*, *211*, 111988.

Shibani, A., Knight, S., & Buckingham Shum, S. (2019). *Contextualizable Learning Analytics Design: A Generic Model, and Writing Analytics Evaluations* 9th International Conference on Learning Analytics and Knowledge, Tempe, Arizona.

Shibani, A., Knight, S., & Shum, S. B. (2022). Questioning learning analytics? Cultivating critical engagement as student automated feedback literacy. Proceedings of the 12th international conference on learning analytics & knowledge (LAK22), Online, USA. https://doi.org/10.1145/3506860.3506912.

Tan, X., Cheng, G., & Ling, M. H. (2025). Artificial intelligence in teaching and teacher professional development: A systematic review. *Computers and Education: Artificial Intelligence*, *8*.

Winstone, N. E., Nash, R. A., Rowntree, J., & Parker, M. (2017). 'It'd be useful, but I wouldn't use it': barriers to university students' feedback seeking and recipience. *Studies in higher education*, *42*(11), 2026-2041.

Wu, Y.-T., Hou, H.-T., Hwang, F.-K., Lee, M.-H., Lai, C.-H., Chiou, G.-L., . . . Chen, N.-S. (2013). A review of intervention studies on technology-assisted instruction from 2005-2010. *Journal of Educational Technology & Society*, *16*(3), 191-203.