Beyond artificial encouragement: Why pre-service teachers value critical over positive feedback

Eloise Thomson
Department of Arts, Education & Agritech, Melbourne Polytechnic, Australia
https://orcid.org/0000-0002-3752-4807

Hayden Park
Department of Arts, Education & Agritech, Melbourne Polytechnic, Australia
https://orcid.org/0000-0003-3287-2199

Rachel Chapman
Department of Arts, Education & Agritech, Melbourne Polytechnic, Australia
https://orcid.org/0000-0002-5606-7306

Sarah Nelson
Department of Arts, Education & Agritech, Melbourne Polytechnic, Australia
https://orcid.org/0009-0007-4400-9276

As generative artificial intelligence (GenAI) tools proliferate in higher education, critical questions arise about their effectiveness in providing assessment feedback, particularly in professional preparation contexts where feedback shapes future practice. This mixed methods study examined the quality and student experience of GenAI-generated versus human feedback among 15 first-year pre-service teachers. Using ChatGPT-4 to produce rubric-aligned feedback compared with expert human assessment, the results showed GenAI consistently awarded higher marks and revealed a "praise paradox", where its uniformly positive tone undermined educational value by appearing excessive and developmentally limited. Student surveys showed a strong preference for human feedback across all quality dimensions, especially accuracy, actionability and professional contextualisation. Key themes included superficial praise concealing learning gaps, reliance on pedagogical expertise over algorithmic judgement and confusion from score disparities. While GenAI demonstrated strengths in personalisation and positivity, it lacked evaluative judgement and contextual understanding essential for teacher education. The study concludes that GenAI should complement, not replace, human expertise, advocating a human-in-the-loop model that preserves the relational and developmental dimensions critical to effective feedback in professional preparation programmes.

Implications for practice or policy:

- Teacher education programmes should not use GenAI to replace human feedback in primary assessment.
- Teacher educators should retain responsibility for evaluative judgement, professional context and developmental guidance.
- Course leaders may use GenAI to support formative feedback drafting, but only with human review and refinement.
- Institutions should establish transparent policies and professional learning so GenAI supports, rather than substitutes, relational feedback practice.

Keywords: generative artificial intelligence (GenAI), assessment feedback, pre-service teachers, teacher education, educational technology, mixed methods

Introduction

The integration of generative artificial intelligence (GenAI) into educational contexts represents one of the most significant technological disruptions facing higher education today. Tools such as ChatGPT, Claude and Gemini demonstrate sophisticated capabilities in content generation and analysis, creating both opportunities and challenges for educational institutions (Al-Ali & Miles, 2025; Baidoo-Anu &Owusu Ansah, 2023; Chen et al., 2020; Xia et al., 2024). This disruption appears most pronounced in assessment practices, where two core functions – measuring achievement and providing feedback – face unprecedented transformation (Memarian & Doleck, 2024).

Within initial teacher education (ITE), assessment feedback holds particular significance, playing a central role in building academic content mastery, pedagogical knowledge and professional identity formation (Henderson et al., 2025). Pre-service teachers (PSTs) need feedback that is both technically accurate and personally relevant, helping them develop as reflective practitioners capable of complex pedagogical decisions. Traditional feedback approaches face increasing strain under conditions of large cohorts, heavy marking loads and demands for timely, personalised responses (Rose et al., 2024).

GenAI technologies appear to offer compelling solutions to these scalability challenges, with emerging research demonstrating their capacity to provide consistent, detailed and timely feedback at scale (Li et al., 2024; Weng et al., 2024). However, fundamental questions remain about the quality, appropriateness and impact of AI-generated feedback, particularly in ITE contexts, where feedback can be transformational rather than merely informational (Yu & Xie, 2025).

GenAI in educational assessment: Promise and limitations

Recent systematic reviews have reveal thated GenAI provides diverse feedback across various educational contexts, reducing instructor workload while enhancing communication and accessibility (Bond et al., 2024; Lee & Moore, 2024). While GenAI has shown significant promise in providing rapid, rubric-aligned feedback at scale, important limitations persist. Recent reviews have highlighted that such systems often struggle with contextual sensitivity and pedagogical nuance (Demirkan & Rasul, 2024; Zhang et al., 2025). Farrokhnia et al. (2024) have provided a useful synthesis through a strengths, weaknesses, opportunities and threats analysis of ChatGPT in educational contexts, demonstrating both its potential and its risks. Their analysis underscored that while ChatGPT's strengths include immediacy, personalisation and workload reduction for educators, its weaknesses stem from a lack of deep understanding and contextual awareness. These limitations mean that AI-generated feedback may appear polished but fail to engage critically with student learning needs. Farrokhnia et al. have cautioned that, without active human oversight, ChatGPT risks producing generic, misleading or surface-level comments that do not align with disciplinary or cohort-specific expectations. Such findings reinforce the need for human-in-the-loop approaches where educators filter, adapt and contextualise AI outputs to ensure developmental value for learners. Tools like ChatGPT can produce rubric-aligned comments and deliver feedback with remarkable speed and consistency (Li et al., 2024; Weng et al., 2024).

Despite these advances, significant limitations remain. While GenAI can replicate surface-level alignment with assessment criteria, it frequently lacks the evaluative depth, contextual sensitivity and professional judgement inherent in human feedback (Zhang et al., 2025). Key concerns include grade inflation (Li et al., 2024), generic or overly positive feedback lacking actionable detail (Escalante et al., 2023) and insufficient understanding of pedagogical contexts (Demirkan & Rasul, 2024). Recent research has suggested AI feedback varies in perceived usefulness depending on context and user trust, with learners finding it either too general or too complex (Burner et al., 2025).

Feedback as relational practice in teacher education

Rather than functioning merely as the transmission of information, feedback in higher education is shaped by the people involved and is ideally a relational and developmental process (Dawson et al., 2023;

Henderson et al., 2025). In ITE contexts, feedback plays a unique role in professional identity formation, helping PSTs integrate theoretical knowledge with practical pedagogical skills (Galindo-Domínguez et al., 2024).

Emerging research has emphasised feedback as a fundamentally relational and developmental process, whose impact can be significantly enhanced when delivered by humans (Playfoot et al., 2024). Human markers draw upon tacit knowledge of student progress, discipline-specific norms and pedagogical expectations to deepen the relevance and resonance of feedback. Students especially value dimensions such as care, encouragement and recognition of individual effort, which can strengthen engagement and professional growth (Henderson et al., 2025).

The distinctive demands of teacher preparation – fostering pedagogical reasoning, reflective practice and professional identity – mean that feedback in ITE contexts should aim to be more than generic academic commentary (Jeon & Lee, 2023). While GenAI systems are increasingly capable of simulating relational qualities in feedback, research has affirmed that technology alone cannot provide the relational qualities needed for transformative feedback in ITE, especially in professional preparation programmes, where human connection is vital for development (Payne et al., 2023).

GenAI in ITE: Current understanding

Educators hold conflicting views of GenAI. Some have highlighted its potential for fostering higher-order thinking (Baidoo-Anu & Owusu Ansah, 2023), while others fear erosion of authentic learning experiences (Estaiteyeh & McQuirter, 2024). From a PST perspective, generally positive attitudes have emerged towards GenAI as a learning buddy, particularly for accessing resources and clarifying concepts, yet persistent concerns remain regarding academic integrity and authenticity of learning experiences (Nyaaba et al., 2024).

Importantly, ITE students have expressed scepticism about whether GenAI can meaningfully support development of professional judgement and pedagogical nuance central to teaching practice (Zhang et al., 2025). While GenAI might supplement some feedback tasks, replacing human feedback in contexts demanding contextual sensitivity, relational connection and professional insight remains problematic (Burner et al., 2025; Henderson et al., 2025).

The literature has highlighted a critical tension: although GenAI tools can automate surface-level feedback, they fall short in replicating relational and developmental aspects essential in professional education. Studies have frequently called for human-in-the-loop models where GenAI assists rather than replaces human educators (Buckingham Shum et al., 2023; Li et al., 2024). Emerging research is beginning to investigate a critical consideration surrounding GenAI feedback: the student experience of receiving it.

Two recent studies of relevance have offered valuable insights into this space. Roe et al. (2026) have shown mixed student attitudes towards GenAI use for feedback, highlighting that feedback combining both GenAI and human educator input was more positively received than GenAI alone. The research by Henderson et al. (2025) have provided insight into an Australian higher education context. They examined how almost 7000 students perceive and engage with AI-generated feedback compared to human educator feedback. While students generally found teacher feedback more helpful and trustworthy, they valued GenAI for its accessibility, speed, volume and clarity. What appears scarce in the literature are empirical investigations directly comparing GenAI and human feedback within teacher education contexts.

This study addresses this perceived gap by exploring how GenAI-generated feedback compares to human feedback in ITE, examining both objective differences in feedback quality and perceptions among PSTs. It contributes nuanced insights into the appropriate role of GenAI in professional preparation programmes, where feedback serves as a cornerstone of professional identity formation. This study explored the use of GenAI-generated feedback in PST education through the following research questions:

How does GenAI-generated feedback compare with human feedback regarding content relevance, accuracy, supportive language and actionable guidance?
What perceptions do PSTs form about GenAI feedback quality, usefulness and implementation potential?

Methodology

This study examined the educational potential of GenAI in providing formative assessment feedback to PSTs. It employed a sequential explanatory mixed methods design (Toyon, 2021) to investigate the comparative quality and perceived value of feedback provided by GenAI versus human educators in an ITE context. The rationale for this design lies in its capacity to integrate quantitative and qualitative data to provide a comprehensive understanding of both the objective characteristics of feedback and the subjective experiences of PSTs (Creswell & Plano Clark, 2018). This methodological integration allows for comprehensive understanding through methodological triangulation (Meydan & Akkas, 2024; Vivek et al., 2023), particularly valuable when researching emerging educational technologies.

The research was conducted within a first-year education unit at an Australian university, involving 15 PSTs who submitted essays for a formal assessment task focused on comparative education. The data collection proceeded in three distinct phases. During the first phase, essays were assessed by experienced ITE educators through blind moderation using a standardised rubric. This ensured consistency and minimised potential bias in human grading. In the second phase, the same essays were marked using ChatGPT-4, prompted to generate rubric-aligned grades and provide 200-word formative feedback. Prompts were carefully designed to elicit strengths-based, supportive commentary while maintaining alignment with the assessment criteria (Table 1). The third and final phase involved the participants receiving both forms of feedback and completing a survey designed to capture their evaluations of each feedback type in terms of content relevance, accuracy, tone and actionability (Research question 1). Participants were also asked to comment on their perceptions of the overall quality and usefulness of each feedback type (Research question 1) and provide commentary on their views of the potential role of GenAI feedback in professional preparation (Research question 2).

To ensure methodological rigour, the study incorporated several strategies designed to enhance credibility, dependability and confirmability. Blind moderation was employed during the human assessment phase to reduce potential bias and ensure consistency in grading (Prichard et al., 2025). At the same time, standardised prompting protocols were used when generating GenAI feedback, ensuring uniformity in the model's responses and minimising variability introduced by prompt design (Robertson et al., 2024). These protocols were carefully constructed to elicit rubric-aligned, strengths-based feedback while maintaining a supportive tone, thereby enabling a fair comparison between human and AI-generated feedback. Ethical approval was obtained through the university's Human Research Ethics Committee. Participation was voluntary, and all data was anonymised using pseudonyms to protect student confidentiality. The study's design reflects a commitment to methodological rigour and ethical integrity.

Data analysis

Data analysis involved separate quantitative and qualitative analyses, followed by integration to provide a comprehensive interpretation of the findings. This approach enabled us to examine both numerical trends and narrative insights without privileging one form of data over the other, ensuring the integrity of each data set. Through this process, the study achieved methodological triangulation, enhancing the validity of findings (Arias Valencia, 2022) and allowing for a more nuanced interpretation of how feedback type influenced both academic outcomes and student perceptions. The integration of these data sources facilitated a richer understanding of the relational, evaluative and developmental dimensions of feedback in ITE courses (Ahmed et al., 2024).

Quantitative data comprised assessment scores and survey responses. Descriptive statistics (Cooksey, 2020, Chapter 5) were used to compare grade distributions between human and GenAI markers, revealing a consistent pattern of grade inflation in GenAI-generated results. Frequency distributions from the

survey provided insight into how PSTs rated each feedback type across dimensions including accuracy, usefulness and actionability. Qualitative data was derived from open-ended survey responses and analysed using reflexive thematic analysis (Braun & Clarke, 2021). This process involved iterative coding and theme development, beginning with familiarisation and initial coding, followed by the identification and refinement of themes. We conducted thematic analysis collaboratively to enhance interpretive validity and reduce individual bias (Braun & Clarke, 2021).

Table 1
Assessment task description

Assessment component	Details
Assessment title	Assessment task 1: Essay on comparative education (35%)
Word limit	1400 words
Task description	Students write an analytical essay examining comparative education with specific focus on early childhood and care education. The essay addresses: Definition and scope of comparative education Benefits and limitations of comparative education approaches Application of comparative education in Australian early childhood education and care frameworks and policies Personal application for future teaching practice
Assessment criteria	Essays are evaluated using a 5-point scale (not satisfactory to high distinction) across five dimensions: understanding of comparative education concepts, critical analysis of benefits and limitations, application to Australian contexts, personal reflection on teaching practice and academic writing quality including referencing and structure.

Quantitative data, assessment scores and survey frequency distributions, were examined descriptively to compare grading patterns and perceived usefulness. Qualitative data from open-ended survey responses were analysed thematically (Braun & Clarke, 2021) to explore how participants experienced and valued each feedback type. Findings from both strands were integrated to provide a nuanced answer to the research questions.

Findings

Part 1: Researcher analysis of GenAI vs human grading

Differences in scoring

GenAI consistently awarded higher grades than human markers (Figures 1 and 2). ChatGPT frequently assigned high distinction or distinction grades where human markers awarded credit or pass scores. For example, one student receiving consistent pass scores from human markers received distinction-level feedback from GenAI. This systematic grade inflation suggests that while GenAI aligns broadly with assessment rubrics, it lacks the professional discernment to moderate borderline performances or penalise gaps in critical analysis and referencing standards.

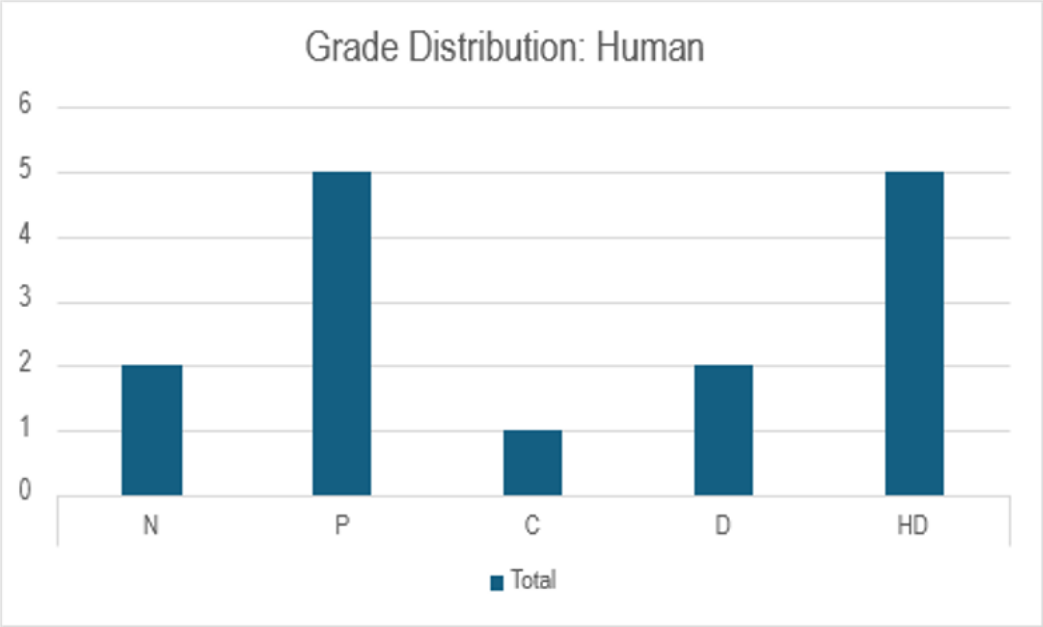

Figure 1. Human grade distribution

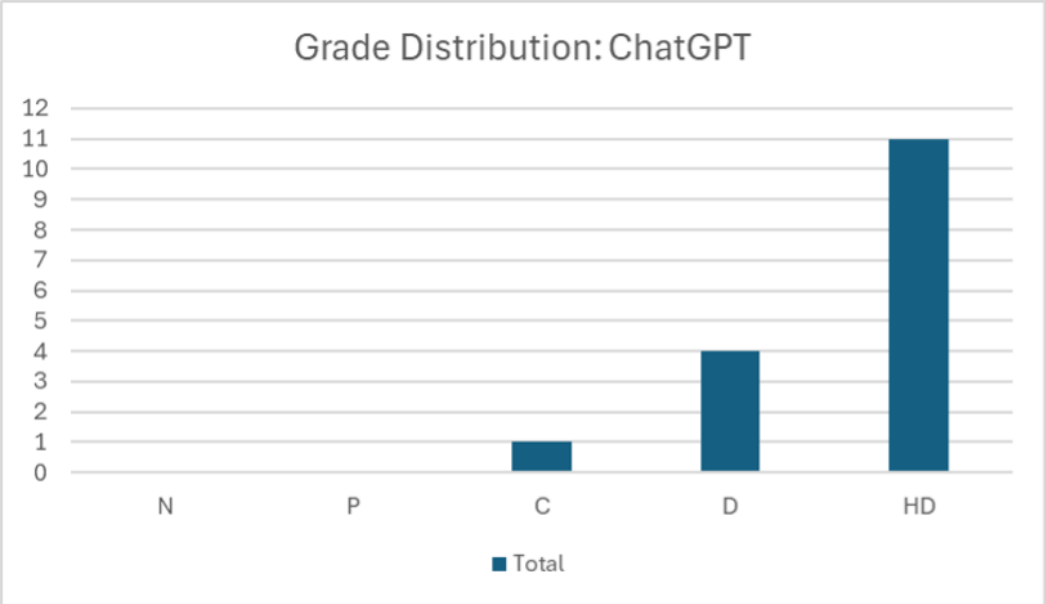

Figure 2. GenAI grade distribution

Differences in feedback tone and content

GenAI feedback was uniformly positive, affirming and strengths based, often beginning with phrases such as "Your essay demonstrates a strong understanding" or "This was a well-structured and thoughtful response". However, as outlined in Table 2, this positivity came at the expense of critical depth and actionable suggestions.

In contrast, human feedback was more nuanced, often affirming where appropriate but also identifying specific weaknesses, such as lack of comparative analysis, insufficient use of scholarly sources or underdeveloped personal reflections. Human comments also referred to student growth and previous work, situating feedback within a relational and developmental context absent from GenAI outputs.

Three notable limitations emerged from our analysis:

- Surface alignment with rubrics: GenAI appeared to match rubric language without critically engaging with the student's interpretation of task requirements or the nuances of academic writing quality.
- Overgeneralisation: Feedback often repeated generic praise without targeting specific criteria, even when essays had visible weaknesses in argumentation or evidence use.
- Lack of contextual judgement: GenAI lacked awareness of broader course expectations, cohort norms and individual student trajectories, all of which inform professional educator feedback.

Table 2
Human and GenAI feedback

PST	Human grade	AI grade	Feedback comparison
Amber	Unsatisfactory	Credit	AI feedback: "Your essay demonstrates a solid understanding of comparative education, effectively exploring its definition, benefits, limitations, and application in early childhood education and care." Human feedback: "It is evident that you have some understanding of what comparative education is and its impact on the education sector broadly. The use of an image in this section detracts from your writing and is not supported by your own commentary."
Grace	Pass	Distinction	AI feedback: "Your essay on comparative education demonstrates a commendable understanding of the subject and its application in early childhood education and care."
Lyn	Pass	High distinction	AI feedback: "Your examination of Australia's use of comparative education in its early childhood education and care policies is detailed and well-researched. You have successfully highlighted how international best practices have influenced Australia's frameworks, providing specific examples and references to support your analysis."

Part 2: Student perceptions of GenAI and human feedback

Nine participants completed the survey after receiving both forms of feedback, providing insights into perceived value, limitations and experiences with each feedback type. Analysis revealed four key themes that emerged from both survey responses and open-ended comments (Table 3).

Table 3
Themes and evidence in student perception data

Theme	Key findings	Survey data	Student quotes
Theme 1: The "praise paradox": when positivity undermines learning	Students recognised excessive positivity as counterproductive to authentic learning, creating a disconnect between supportive tone and perceived value.	8/9 agreed GenAI feedback was supportive in tone Only 2/9 found it valuable overall	"How will anyone really learn if feedback is always positive and vague?" "It gave me top marks—but said nothing useful about how to do better next time." “Learning shouldn't be about flattery, it's about growth."
Theme 2: Relational	Students valued feedback rooted in	9/9 rated human feedback higher	"My teacher knows my work and how far I've come. AI can't see that."

Theme	Key findings	Survey data	Student quotes
nature of feedback and professional trust	relational connection and professional expertise, emphasising the importance of educators who understand their learning trajectory and professional context.	for trustworthiness	"AI hasn't worked with children... My teacher knows what to look for in a future educator." "I feel like I pay for the human professional with experience to help me learn—even if AI were better—I think I would prefer the human."
Theme 3: Developmental usefulness vs surface-level praise	While GenAI feedback appeared comprehensive, students found it lacking in actionable guidance for improvement, preferring critical engagement over mere encouragement.	Only 3/9 found AI feedback actionable 9/9 found human feedback actionable	"The AI feedback was encouraging, but I couldn't see what specifically I needed to improve." "It was nice, but I didn't know what to fix." "It felt like artificial encouragement... learning shouldn't be about flattery—it's about growth."
Theme 4: Confusion and conflict from score disparities	Students experienced confusion when confronted with discrepancies between human and GenAI assessments, highlighting emerging challenges in digital assessment environments and the need for AI literacy support.	Multiple instances of conflicting scores between human and AI feedback	"I don't know which one is right—they were really different." "The AI passed me and the human failed me... Which one is right?" "I prefer the GenAI feedback because it passed me. But I also know that's not really fair."

Discussion

This study set out to examine GenAI-generated feedback in ITE through a comparative analysis of feedback quality characteristics and an exploration of PSTs' perceptions regarding its usefulness and implementation potential. The findings provide important insights into how GenAI functions in professional education contexts where feedback serves not only as an academic tool but as a critical driver of professional identity formation. This aligns with Li et al. (2024) and Sætre (2025), highlighting GenAI's limited ability to replicate the evaluative judgement central to professional assessment practice.

The illusion of alignment: Surface-level consistency without depth

Our analysis reveals that while GenAI feedback often mirrors rubric terminology and presents a polished, professional tone, it lacks the evaluative depth and contextual insight required for meaningful assessment in ITE. Consistent with prior research (Zhan et al., 2025; Zhang et al., 2025), GenAI appears adept at replicating the form of academic feedback but falls short in delivering its substance.

The systematic grade inflation observed in this study, with AI consistently awarding higher grades than human assessors, highlights GenAI's limitations in nuanced performance evaluation. Rather than differentiating between borderline performances, GenAI tended to award higher scores, failing to penalise gaps in critical analysis or insufficient referencing. These results align with concerns raised by Li et al. (2024) and Bond et al. (2024) that while GenAI demonstrates surface alignment with assessment

criteria, it often lacks the discriminatory capacity essential for authentic assessment in professional education. In ITE, feedback that is direct, contextual and pedagogically sound is crucial.

Feedback as a relational and developmental practice

A critical dimension emerging from this study is the importance of relational feedback in teacher education. Students consistently valued human feedback for its authenticity, contextual relevance and developmental usefulness. Unlike GenAI, human markers have the innate ability to connect feedback to students' prior work, individual learning trajectories and the professional expectations of teaching practice – an element central to effective ITE (Dai et al., 2025; Galindo-Domínguez et al., 2024; Henderson et al., 2025).

Participants' comments underscored that feedback is not simply informational but fundamentally relational. Trust and professional growth were linked to an educator's ability to recognise individual progress and situate feedback within the realities of classroom practice. This finding reinforces recent literature highlighting the role of feedback in shaping professional identity and promoting reflective practice (Playfoot et al., 2024; Rose et al., 2024).

The praise paradox: When positivity undermines learning

A novel and significant contribution of this study is the identification of what we term the "praise paradox". While GenAI feedback was uniformly positive and affirming, students perceived this as superficial and, paradoxically, counterproductive. Despite acknowledging the encouraging tone, many participants found GenAI's feedback lacking in actionable guidance and critical depth essential for improvement.

This challenges the common assumption that students invariably prefer positive feedback. Instead, our findings reveal that PSTs value honest, constructive critique over generic praise, particularly in contexts where professional competence and identity are being formed. As one participant observed, "Learning shouldn't be about flattery – it's about growth". This insight contributes to growing evidence that effective feedback in higher education must balance supportiveness with critical engagement to be genuinely developmental and effective (Dawson et al., 2023; Henderson et al., 2025).

Navigating conflicting scores: Trust and AI literacy

A further challenge illuminated by our study is the confusion generated by discrepancies between human and GenAI assessments. Several participants expressed uncertainty about whose feedback to trust when grades and comments diverged substantially. This highlights an emerging issue in digital assessment environments: the need for AI feedback literacy among students. Trust in feedback is also shaped by transparency. Luo (2024) found that students' confidence in teacher–student relationships was weakened when AI use in assessment was not openly communicated. Students were uneasy about declaring their own AI use while unsure whether teachers were also relying on it, creating a sense of unfairness and low trust. In teacher education, where feedback plays a central role in professional growth, such gaps highlight the importance of openness about how AI is used and the continued presence of human oversight.

Without guidance on interpreting AI-generated feedback and understanding its limitations, students risk being misled by superficially positive but educationally limited responses. This finding aligns with concerns raised by Heiser et al. (2024) and Nikolic et al. (2024) about the potential for AI to erode trust in assessment systems if its role and limitations are not transparently communicated. Addressing this gap will be crucial for institutions integrating AI tools into assessment processes.

Human-in-the-loop: A path forward

Taken together, the findings suggest that while GenAI has potential as a supplementary tool for providing formative feedback, it cannot replace the critical, relational and pedagogical functions of human feedback

in ITE. GenAI may effectively generate initial feedback drafts, identify surface-level strengths or offer rapid responses in large cohorts. Recent work with PSTs has also shown that AI feedback is most effective when paired with human guidance. Wang et al. (2025) found that students were initially anxious about receiving AI-generated analyses of their teaching, but when these were discussed in supportive group settings, they became valuable for reflection and improvement. This highlights that AI feedback gains meaning when educators help students interpret and apply it, reinforcing the need for human-in-the-loop approaches in teacher education. However, its limitations in critical judgement, contextual sensitivity and professional insight underscore the importance of maintaining human-in-the-loop models (Buckingham Shum et al., 2023; Li et al., 2024).

Successful integration of GenAI into teacher education requires more than technical solutions. Institutions must provide professional development to enhance educators' and students' AI literacy, establish clear ethical guidelines and develop policies ensuring that GenAI serves as a support rather than a substitute for human expertise. This hybrid approach offers a promising pathway for leveraging GenAI's efficiencies while preserving the human relational dimensions essential to effective professional learning.

Considerations for ITE

Teacher education programmes must develop comprehensive AI literacy training for both faculty and students, emphasising critical evaluation of AI-generated feedback and positioning GenAI as supplementary rather than a replacement for human expertise. Programmes should establish transparent communication protocols about AI tool usage and engage in co-design processes with developers to ensure feedback tools align with pedagogical priorities and relational practices essential to teacher preparation. Future research should explore fine-tuned models and hybrid systems designed specifically for teacher education contexts, alongside longitudinal studies examining how repeated GenAI feedback exposure influences student perceptions and professional development across diverse institutional contexts.

Why might you use AI to give feedback?

AI feedback systems offer compelling advantages including quick implementation and significant workload reduction while maintaining assessment practices at scale (Bond et al., 2024). These systems enable more frequent and personalised feedback than current time constraints allow (Lindsay et al., 2025), while providing perceived consistency with reduced subjective bias risk (Henderson et al., 2025). Li et al. (2024) and Weng et al. (2024) have highlighted AI's ability to deliver detailed, rubric-aligned feedback with remarkable speed and consistency. Additionally, GenAI enables educators to adopt a more curatorial role, concentrating on high-level, subject-specific feedback rather than routine marking, potentially improving other teaching areas through decreased workload (Lee & Moore, 2024).

Limitations and considerations

The findings may reflect the specific GenAI tool (ChatGPT-4), version and prompt design used. Our rubric-aligned, strengths-based prompts likely encouraged surface alignment with criteria and contributed to the systematic grade inflation observed, limiting critical depth and penalisation of weaknesses. The "praise paradox" identified uniform positivity with low developmental utility may have been amplified by instructions to maintain a supportive tone, reducing actionable guidance.

Furthermore, the GenAI system's lack of access to students' prior work, learning trajectory, and course context restricted its ability to deliver relational and professionally grounded feedback – a dimension consistently valued in teacher education. As this study involved a small, single-institution sample, results may not generalise to other cohorts, disciplines or GenAI configurations. Future research should examine whether alternative prompting, fine-tuned models or hybrid human-in-the-loop approaches can address these limitations.

Conclusion

This investigation provides empirical evidence of GenAI's potential and limitations in teacher education assessment. While GenAI demonstrates capabilities in providing personalised and timely feedback, concerns about grade inflation, surface-level engagement and limited developmental utility challenge its suitability for primary feedback provision. Student experiences revealed sophisticated understanding of feedback's role, with consistent preferences for human expertise. The identified "praise paradox", where excessive AI positivity undermines learning, represents a novel contribution to understanding AI feedback effectiveness.

For teacher education programmes, human expertise remains irreplaceable for primary assessment feedback, though GenAI may offer valuable supplementary support with appropriate oversight. The future lies in thoughtful augmentation that preserves the mentorship, professional wisdom and contextual understanding central to preparing effective educators.

Disclosure of GenAI usage

ChatGPT-4 (OpenAI) was used in January and July–August 2025 to generate rubric-aligned grades and formative feedback for the study's comparative analysis, and for minor editorial refinements to the manuscript. Standardised prompts instructed the model to provide a grade and approximately 200 words of supportive, rubric-aligned feedback highlighting strengths and areas for improvement. All outputs were reviewed, edited and verified by the authors, who takes full responsibility for the final content.

Author contributions

Eloise Thomson: Conceptualisation, Investigation, Writing – original draft, Writing – review and editing; **Hayden Park:** Data curation, Investigation, Formal analysis, Writing – review and editing; **Rachel Chapman:** Writing – review and editing; **Sarah Nelson:** Writing – review and editing.

References

Ahmed, A., Pereira, L., & Jane, K. (2024). *Mixed methods research: Combining both qualitative and quantitative approaches*. ResearchGate. https://www.researchgate.net/publication/384402328_Mixed_Methods_Research_Combining_both_qualitative_and_quantitative_approaches

Al-Ali, S., & Miles, R. (2025). Upskilling teachers to use generative artificial intelligence: The TPTP approach for sustainable teacher support and development. *Australasian Journal of Educational Technology*, *41*(1), 88–106. https://doi.org/10.14742/ajet.9652

Arias Valencia, M. M. (2022). Principles, scope, and limitations of the methodological triangulation. *Investigación y Educación en Enfermería*, *40*(2). https://doi.org/10.17533/udea.iee.v40n2e03.

Baidoo-Anu, D., & Owusu Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, *7*(1), 52–62. https://doi.org/10.61969/jai.1337500

Bond, M., Ajjawi, R., Bearman, M., Dracup, M., Handel, M., Bennett, S., Cherry, E., & Tai, J. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education*, *21*, Article 4. https://doi.org/10.1186/s41239-023-00436-z

Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. SAGE Publications.

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), Article 40. https://doi.org/10.1186/s41239-023-00410-9

Burner, T., Lindvig, Y., & Wærness, J. I. (2025). We should not be like a dinosaur—Using AI technologies to provide formative feedback to students. *Education Sciences*, *15*(1), Article 58. https://doi.org/10.3390/educsci15010058

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, *8*, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510

Cooksey, R. W. (2020). *Finding meaning in quantitative data*. Springer.

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.

Dai, W., Tsai, Y-S., Gašević, D., & Chen, G. (2025). Designing relational feedback: A rapid review and qualitative synthesis. *Assessment & Evaluation in Higher Education*, *50*(1), 16–30. https://doi.org/10.1080/02602938.2024.2361166

Dawson, P., Henderson, M., Mahoney, P., Phillips, M., Ryan, T., Boud, D., & Molloy, E. (2023). What makes for effective feedback: Staff and student perspectives. *Assessment & Evaluation in Higher Education*, *48*(1), 25–43. https://doi.org/10.1080/02602938.2018.1467877

Demirkan, O., & Rasul, T. (2024). Exploring the use of artificial intelligence (AI) in the delivery of effective feedback. *Assessment & Evaluation in Higher Education*, *49*(8), 1654–1672. https://doi.org/10.1080/02602938.2024.2415649

Escalante, J., Pack, A., & Barrett, A. (2023). AI-generated feedback on writing: Insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education*, *20*(1), Article 57. https://doi.org/10.1186/s41239-023-00425-2

Estaiteyeh, M., & McQuirter, R. (2024). Generative or degenerative?!: Implications of AI tools in preservice teacher education and reflections on instructors' professional development. *Brock Education Journal*, *33*(3), 75–98. https://doi.org/10.26522/BROCKED.V33I3.1176

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, *61*(3), 460–474. https://doi.org/10.1080/14703297.2023.2195846

Galindo-Domínguez, H., Bezanilla, M. J., & Fernández-Nogueira, D. (2024). Artificial intelligence for higher education: Benefits and challenges for preservice teachers. *Frontiers in Education*, Aritcle *9*. https://doi.org/10.3389/feduc.2024.1501819

Heiser, R. E., Palalas, A., & Gollert, A. (2024). Digital well-being begins with inclusion: A systematic review of videoconferencing guidelines for equitable learning. *Australasian Journal of Educational Technology*, *41*(1), 18–26. https://doi.org/10.14742/ajet.9552

Henderson, M., Bearman, M., Chung, J., Fawns, T., Buckingham Shum, S., Matthews, K. E., & de Mello Heredia, J. (2025). Comparing generative AI and teacher feedback: Student perceptions of usefulness and trustworthiness. *Assessment & Evaluation in Higher Education*, 1–16. https://doi.org/10.1080/02602938.2025.2502582

Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies, 28*, 15873–15892. https://doi.org/10.1007/s10639-023-11834-1

Lee, S. S., & Moore, R. L. (2024). Harnessing generative AI (GenAI) for automated feedback in higher education. *Online Learning Journal*, *28*(3), Article 4593. https://doi.org/10.24059/olj.v28i3.4593

Li, J., Jangamreddy, N. K., Hisamoto, R., Bhansali, R., Dyda, A., Zaphir, L., & Glencross, M. (2024). AI-assisted marking: Functionality and limitations of ChatGPT in written assessment evaluation. *Australasian Journal of Educational Technology*, *40*(4), 56–72. https://doi.org/10.14742/ajet.9463

Lindsay, E.D., Zhang, M., Johri, A., & Bjerva, J. (2023). *The responsible development of automated student feedback with generative AI*. arXiv. https://doi.org/10.48550/arXiv.2308.15334

Luo, J. (2024). How does GenAI affect trust in teacher-student relationships? Insights from students' assessment experiences. *Teaching in Higher Education*, *30*(4), 991–1006. https://doi.org/10.1080/13562517.2024.2341005

Memarian, B., & Doleck, T. (2024). A review of assessment for learning with artificial intelligence. *Computers and Education: Artificial Intelligence*, *2*, Article 100040. https://doi.org/10.1016/j.chbah.2023.100040

Meydan, C., & Akkas, H. (2024). The role of triangulation in qualitative research. In A. Elhami, A. Roshan, & H. Chandan (Eds.), *Converging perspectives* (pp. 101–132). https://doi.org/10.4018/979-8-3693-3306-8.ch006

Nikolic, S., Wentworth, I., Sheridan, L., Moss, S., Duursma, E., Jones, R. A., Ros, M., & Middleton, R. (2024). A systematic literature review of attitudes, intentions and behaviours of teaching academics pertaining to AI and generative AI (GenAI) in higher education: An analysis of GenAI adoption using the UTAUT framework. *Australasian Journal of Educational Technology*, *40*(6), 56–75. https://doi.org/10.14742/ajet.9643

Nyaaba, M., Shi, L., Nabang, M., Zhai, X., Kyeremeh, P., Ayoberd, S. A., & Akanzire, B. N. (2024). *Generative AI as a learning buddy and teaching assistant: Preservice teachers' uses and attitudes.* arXiv. https://doi.org/10.48550/arXiv.2407.11983

Payne, A. L., Ajjawi, R., & Holloway, J. (2023). Humanising feedback encounters: A qualitative study of relational literacies for teachers engaging in technology-enhanced feedback. *Assessment & Evaluation in Higher Education*, *48*(7), 903–914. https://doi.org/10.1080/02602938.2022.2155610

Playfoot, D., Horry, R., & Pink, A. E. (2024). What's the use of being nice? Characteristics of feedback comments that students intend to use in improving their work. *Assessment & Evaluation in Higher Education*, *50*(2), 187–198. https://doi.org/10.1080/02602938.2024.2373799

Prichard, R., Peet, J., El Haddad, M., Chen, Y., & Lin, F. (2025). Assessment moderation in higher education: Guiding practice with evidence - an integrative review. *Nurse Education Today*, *146*, Article 106512. https://doi.org/10.1016/j.nedt.2024.106512

Robertson, J., Ferreira, C., Botha, E., & Oosthuizen, K. (2024). Game changers: A generative AI prompt protocol to enhance human–AI knowledge co-construction. *Business Horizons*, *67*(5), 499–510. https://doi.org/10.1016/j.bushor.2024.04.008

Roe, J., Perkins, M., & Ruelle, D. (2026). Is GenAI the future of feedback? Understanding student and staff perspectives on AI in assessment. *Intelligent Technologies in Education*, *1*(1), Article 7. https://doi.org/10.53761/ITED/1.7

Rose, S. E., Taylor, L., & Johnson, K. (2024). Perceptions of feedback and engagement with feedback among undergraduates: An educational identities approach. *Assessment & Evaluation in Higher Education*, *50*(2), 266–278. https://doi.org/10.1080/02602938.2024.2390933

Sætre, C. (2025). Behind the grades: Co-constructing fairness to reach agreement in evaluative judgement. *Assessment & Evaluation in Higher Education*, *50*(2), 236–249. https://doi.org/10.1080/02602938.2024.2373789

Toyon, M. (2021). Explanatory sequential design of mixed methods research: Phases and challenges. *International Journal of Research in Business and Social Science*, *10*(5), 253–260. https://doi.org/10.20525/ijrbs.v10i5.1262

Vivek, R., Nanthagopan, Y., & Piriyatharshan, S. (2023). Beyond methods: Theoretical underpinnings of triangulation in qualitative and multi-method studies. *SEEU Review*, *18*(2), 105–122. https://doi.org/10.2478/seeur-2023-0088

Wang, M., Chen, Z., Xu, Y., Maheshi, B., & Gašević, D. (2025). Intelligent teaching analytics for collaborative reflection: Investigating preservice teachers' perceptions, experiences and shared regulation processes. *International Journal of Educational Technology in Higher Education, 22*, Article 45. https://doi.org/10.1186/s41239-025-00538-w

Weng, X., Xia, Q., Gu, M., Rajaram, K., & Chiu, T. K. F. (2024). Assessment and learning outcomes for generative AI in higher education: A scoping review on current research status and trends. *Australasian Journal of Educational Technology*, *40*(6), 37–55. https://doi.org/10.14742/ajet.9540

Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education*, *21*, Article 40. https://doi.org/10.1186/s41239-024-00468-z

Yu, H., & Xie, Q. (2025). Generative AI vs. teachers: Feedback quality, feedback uptake, and revision. *Language Teaching Research Quarterly, 47*, 113–137. https://doi.org/10.32038/ltrq.2025.47.07

Zhan, Y., Boud, D., Dawson, P. & Yan, Z. (2025). Generative artificial intelligence as an enabler of student feedback engagement: A framework. *Higher Education Research & Development*, *44*(5), 1289–1304. https://doi.org/10.1080/07294360.2025.2476513

Zhang, A., Gao, Y., Suraworachet, W., Nazaretsky, T., & Cukurova, M. (2025). *Evaluating trust in AI, human, and co-produced feedback among undergraduate students.* arXiv. https://arxiv.org/abs/2504.10961

Corresponding author: Eloise Thomson, eloisethomson@melbournepolytechnic.edu.au

Please cite as: Thomson, E., Park, H., Chapman, R., & Nelson, S. (2026). Beyond artificial encouragement: Why pre-service teachers value critical over positive feedback. *Australasian Journal of Educational Technology*, *42*(4), 24–37. https://doi.org/10.14742/ajet.11550