Pilots Completion Report

Document details

Document owner	Elizabeth Lewis
Document author	Elizabeth Lewis
Current version	1
Issue date	22/09/2025

Version control

Date	Version	Change details

Title of the Pilot	Exploring AI-Assisted Feedback to Enhance Flexibility in Assessment
Pilot ID Number	Pilot-ID76

Pilot Completion Report Template

Report Category	Report Requirement
Overall Rating - ***Partially Delivered,***	
Summary	The purpose of this pilot was to explore how AI-assisted feedback might enhance the quality, efficiency, and flexibility of assessment feedback in higher education. Students consistently identify feedback as one of the most problematic aspects of their learning, describing it as unclear, inconsistent, or difficult to act upon. At the same time, staff face significant workload pressures in producing timely, personalised, and high-quality feedback at scale. This project set out to evaluate whether AI could provide meaningful support in addressing these challenges, without undermining academic standards or the relational aspects of feedback. The pilot considered three feedback models. Human-only feedback – the traditional approach where comments are produced entirely by the marker. AI-only feedback – where feedback is generated independently by AI. And Human+AI feedback – the model trialled in this project, where markers draft feedback and a custom AI tool provides *coaching questions* to improve clarity, structure, and actionability. The project aimed to: 1. Enhance staff capacity by building confidence and competence in AI for feedback. 2. Increase efficiency and consistency in feedback processes. 3. Explore student engagement with human-only, AI-only, and human+AI feedback. 4. Develop and share best-practice resources. 5. Promote long-term adoption, aligned with Flexible Learning Programme (FLP) priorities. **Enhancing Staff Capacity, Efficiency, and Consistency** A staff workshop with 12 participants, unit staff and Graduate Teaching Assistants, provided hands-on experience of the AI feedback assistant. Survey results showed strong outcomes: 100% of staff reported increased understanding of AI tools, 80% greater confidence, and 90% an improved ability to apply them in marking. These findings exceeded the project's 70% confidence benchmark and demonstrate that the pilot met its first aim of enhancing staff capacity. Staff valued the way coaching prompts encouraged them to clarify and expand their feedback. Unit coordinators saw particular potential in large teaching teams, where AI could help ensure a consistent standard of comments and support moderation. New markers were identified as the group most likely to benefit, as the tool provided scaffolding at the outset of their practice. Efficiency gains were less immediate. Some participants felt the process took longer initially as they learned how to prompt

	effectively, and experienced GTAs reported a sense of losing their personal style. Audio input was highlighted as a time-saving feature, allowing staff to dictate feedback quickly. Overall, the tool was viewed as reliable and helpful, with moderation benefits especially valued, though efficiency improvements will depend on staff training and familiarity. **Exploring and Improving Student Engagement** Three focus groups with 14 psychology undergraduates (balanced by year group and attainment) provided insight into student perceptions. Students strongly rejected the idea of AI-only feedback. Many described their own experiences of using ChatGPT to create feedback for themselves, noting that outputs could be inflated, generic, or inaccurate. This reinforced scepticism that AI could replace expert judgement. However, when presented with examples of AI feedback without knowing its source, some responded positively to its clarity and level of detail. This suggests that AI-only models may be acceptable if combined with human interpretation and support, but not as stand-alone feedback. The *human element* of feedback was consistently valued, with students emphasising the importance of learning from experts and feeling seen and supported. At the same time, they recognised limitations in current human-only feedback, including inconsistency between markers and vague or unhelpful phrasing. The human+AI approach was widely seen as a promising middle way. Students felt it preserved human judgement while improving clarity, actionability, and consistency. They welcomed features such as embedded hyperlinks to resources, and coaching prompts that encouraged markers to adopt an empathetic stance and avoid common pitfalls (e.g. suggesting “add more detail” when essays were already at the word limit). Concerns remained that feedback could “read like AI,” or that staff might rely too heavily on the tool. Students stressed the need for visible human quality assurance and opportunities for dialogue. Seeing how the AI-assistant worked, the safeguards in place and understanding the custom tool (as distinct from generic chat GPT was essential for acceptance. Perceptions were shaped by student experience: Year 3 students who described higher AI literacy were more open to the model, while Year 2 students, who did not use AI tools, were more sceptical and raised issues such as environmental impact. This suggests suitability may vary by year group and context. Students also identified peer review as an area where the tool could add value, providing structure and support in an activity they often found difficult.

	Building the Evidence Base The focus groups highlighted the range of student reactions, the factors that shape those reactions, and potential barriers and facilitators to adoption. However, the small sample size limits the generalisability of the findings. To build on this, the project will conduct a quantitative survey with a larger student cohort. This will establish the prevalence of key viewpoints, segment responses by year group, attainment, and AI literacy, and provide stronger evidence to guide decisions around the adoption of AI feedback approaches. **Communicating and Embedding Findings** Insights were presented at a national Flexible Assessment event (Over 80 attendees from across the UK and internationally), shared with the Faculty of Humanities Associate Dean for Realising Student Potential, and discussed with unit leads on the BSc Psychology. Plans are in place to trial the tool in a low-stakes peer review activity, giving students direct experience of AI-assisted feedback. There are also ongoing discussions about integrating the tool into the Academic Development Programme as part of training in effective feedback. This positions the pilot not only as an evaluation, but as the foundation for longer-term adoption. **Resource Development** A staff-facing decision and implementation resource is being developed to support wider adoption. This will help colleagues assess whether the human+AI model is suitable for their local context and provide practical guidance on how to implement it. The resource is being shaped by the pilot evidence and will include: • Criteria for assessing suitability. • Implementation steps and examples. • Strategies for communicating with staff and students. • Guidance on training and moderation. **Outcomes and Next Steps** Overall, the pilot: • **Met Objective 1 (Staff capacity)** – staff reported significant gains in confidence and competence. • **Partially met Objective 2 (Efficiency and consistency)** – strong moderation benefits but mixed short-term efficiency. • **Met Objective 3 (Student engagement)** – students rejected AI-only, valued human+AI, and identified concerns and enablers. • **Progressed Objective 4 (Resources)** – draft staff resource under development. • **Progressed Objective 5 (Adoption)** – dissemination underway, with early integration routes identified.
Deliverables	The pilot delivered the following key activities and outputs:

	Staff development ○ One workshop with 12 participants (mix of unit staff and GTAs), providing hands-on experience of the AI feedback assistant. ○ Post-workshop survey collected to assess changes in confidence, competence, and perceptions of efficiency and consistency. **Student engagement** ○ Three focus groups with 14 psychology undergraduates, balanced across year groups (50% entering Year 2, 50% entering Year 3) and attainment levels. ○ Collection of qualitative data on student perceptions of human-only, AI-only, and human+AI feedback approaches. **Resources under development** ○ A draft staff-facing decision and implementation resource, designed to help colleagues assess local suitability and adopt the human+AI model effectively. **Dissemination and communication** ○ Presentation of pilot insights at the Flexible Assessment Workshop. ○ Discussions with the Faculty of Humanities Associate Dean for Realising Student Potential regarding adoption potential. ○ Engagement with unit leads on the BSc Psychology to explore integration, including a planned low-stakes peer review activity. ○ Ongoing discussions about integration into the Academic Development Programme as part of staff training. **Next steps** Planned a large-scale quantitative survey to establish the prevalence of student views and the role of AI literacy in shaping acceptance.
Relevance	**Students** valued the human+AI model for producing clearer, more actionable feedback while preserving human judgement and empathy. They also raised concerns around over-reliance, environmental impact, and the need for opportunities for dialogue. Valued the opportunity to participate in discussions around the future of feedback and AI, and the potential to directly shape the feedback they receive. **Staff** reported increased confidence and competence in using AI tools, and saw particular benefits for moderation and consistency across teams. New markers were especially positive, while some

	experienced GTAs highlighted challenges in adapting established routines. **Leadership and programme leads** recognised the potential to save time in moderation, improve quality assurance, and embed the tool into training programmes. The pilot aligns strongly with FLP objectives by: • Addressing **technology-enhanced assessment** through a practical trial of AI tools. • Responding to **student experience priorities** by tackling dissatisfaction with clarity and actionability of feedback. • Supporting **staff development** and capacity-building to reduce workload pressures. • Contributing to **policy and strategy discussions** on the ethical and cultural considerations of AI in assessment.
Efficiency	The pilot explored whether AI-assisted tools could make assessment processes more efficient. Evidence indicates limited short-term workload savings but clear potential for efficiency at scale. **Efficiency Gains** **Moderation support**: Coaching questions acted as "on the spot" moderation. Coordinators noted that access to chat logs could allow earlier resolution of issues, reducing time spent in formal moderation. **Support for new markers**: The tool provided structured prompts that reduced the need for intensive oversight of GTAs and early-career staff. **Audio input**: Staff valued the ability to dictate feedback, capturing thoughts quickly while the AI refined structure and clarity. **Challenges** **Initial time costs**: Some experienced GTAs found the tool slower than established routines, as they needed to learn effective prompting. **Personal voice**: Staff were concerned that over-reliance on AI could dilute the individuality of their feedback. **Mixed perceptions**: Not all staff believed workload was reduced in the short term; training and familiarity were seen as prerequisites.
Effectiveness	The pilot demonstrated the clear effectiveness of the human+AI approach in improving clarity, actionability, and consistency of feedback, benefiting both staff and students. **Improvements in Effectiveness** **Clarity and structure**: Staff survey data showed significant confidence gains, with staff highlighting how the tool helped refine feedback into structured formats ("what went well/what to improve"). Students similarly praised the increased detail: *"the*

	feedback that [the AI feedback assistant] gives is much more detailed and that is what… students are wanting." **Actionability**: Students valued embedded hyperlinks and concrete examples, which gave clearer routes for improvement. Coaching prompts also reminded markers to provide constructive suggestions. **Consistency**: Unit coordinators saw benefits in levelling the quality of feedback across large marking teams, while new markers reported feeling more supported. **Flexibility**: Audio input allowed staff to dictate feedback quickly, while students appreciated that feedback could be presented in a more consistent and supportive way. Students valued the potential for them to shape coaching questions for markers so their feedback needs were represented during feedback creation. **Alternative Approaches** AI-only feedback was strongly critiqued by students in principle, largely due to concerns about trust, accuracy, and the absence of human contact. However, when students were shown examples of AI-generated feedback without knowing how it had been produced, they responded positively, praising the level of detail and clarity. This suggests that AI-only feedback could be received favourably if it were accurate and well-designed, but the lack of "human-ness" and expert engagement remains a barrier. A potential alternative balance is a model in which students receive AI-generated feedback (produced by a carefully designed system underpinned by expert criteria), combined with one-to-one human support to help interpret and apply that feedback. This could preserve the efficiency benefits of automation while retaining meaningful human dialogue, which students consistently emphasised as essential.
Outcome	The pilot achieved or made strong progress toward most objectives, with some still in progress. **Objective 1 – Staff capacity**: **Met**. Staff confidence and competence improved significantly. **Objective 2 – Efficiency and consistency**: **Partially met**. Consistency gains were strong, efficiency gains less immediate. **Objective 3 – Student engagement**: **Met (within scope)**. Students rejected AI-only, valued human+AI, and highlighted concerns and facilitators. A quantitative survey will test prevalence across the wider population.

	Objective 4 – Resources: **In progress**. A staff-facing implementation resource is being developed. **Objective 5 – Adoption**: **In progress**. Dissemination underway, with early pathways identified (e.g. staff training, peer review activity, faculty engagement). **Improved Outcomes** • **Staff**: More confident and competent in producing clear, consistent feedback. • **Students**: Better understanding of the role of AI; recognition that human+AI feedback is clearer and more actionable. • **Leadership**: Engaged in discussions on adoption and training integration. **Unintended Outcomes** • **Positive**: Staff identified new uses (training, moderation); students suggested peer feedback applications. • **Negative**: Experienced GTAs reported loss of personal voice; students raised concerns about over-reliance, environmental impacts, and “AI-like” tone in feedback.
Sustainability	The pilot highlighted strong potential for long-term adoption if key conditions are met. **Potential for Change** • Staff confidence and moderation benefits suggest sustainability in training and large teaching teams. • Students valued human+AI as a “middle way,” with potential extension into peer review. • Leadership interest is evident, with discussions at the faculty level and integration into the training under consideration. **Conditions for Sustainability** • **Training and support**: Staff need ongoing development in prompting skills to achieve efficiency. • **Clear communication with students**: Transparency that AI supports but does not replace staff, with dialogue opportunities preserved. • **Policy and governance**: Alignment with institutional AI strategies; attention to ethical and environmental concerns. • **Evaluation at scale**: Quantitative survey needed to inform decision-making across programmes. The pilot was delivered using ChatGPT Team licences, which provided a secure and reliable way to test the AI assistant with

	staff and students. While this approach was effective for piloting, it is relatively cost-intensive and would limit feasibility for large-scale adoption. As part of the project, an alternative route was explored through hosting custom GPTs via a WordPress-based website. This model would allow multiple users to access the system through a single ChatGPT licence, supplemented by hosting and usage costs, delivering the same pedagogical outcomes at a fraction of the cost. It also creates a flexible "sandbox" environment where unit leads can set up, test, and deploy custom GPTs tailored to their needs. Importantly, the sandbox extends beyond assessment practices, offering opportunities to explore uses such as personalised learning, student support, and discipline-specific applications of AI. In doing so, it provides staff with a safe space to build expertise in AI-assisted teaching and learning design, fostering innovation at both unit and programme level. This technical model offers a scalable and financially viable pathway for adoption, directly supporting the Flexible Learning Programme's priorities of innovation, flexibility, and staff development. In parallel, further capacity to sustain and extend this work has been secured through a two-year Institute of Teaching and Learning Fellowship (0.4 FTE) focused on AI in education. While not tied exclusively to this pilot, the fellowship aligns closely with its aims and will provide a dedicated mechanism to continue developing, testing, and embedding the human+AI feedback model. It ensures continuity beyond the life of the pilot, creates opportunities to integrate findings into staff development initiatives, and strengthens institutional capacity to scale innovative AI practices.
Financial	There were some changes from the anticipated spend profile. Fewer ChatGPT Team licences were required than originally budgeted, and workshop expenses were lower than expected. The remaining funds were reallocated to support the development of a low-cost, high-access version of the tool via a WordPress-hosted custom GPT system. This approach not only made best use of the available funds but also provided a more scalable and financially sustainable route for future adoption.
Lessons Identified / Learned	**Systems and Process** • *Enablers*: Coaching prompts provided "on the spot" moderation; potential use of chat logs for early quality assurance. • *Inhibitors*: Lack of integration with existing workflows risks inefficiency for experienced staff. **Incentives and Capacity** • *Enablers*: Valuable scaffolding for new markers, reducing training and oversight demands. • *Inhibitors*: Experienced GTAs are resistant due to loss of personal style; incentives are needed to encourage

	adoption. **Policy and Strategy** • *Enablers*: Alignment with FLP priorities supports the strategic case for adoption. • *Inhibitors*: Student concerns (sustainability, perceptions of staff disengagement) must be addressed through policy and communication. **Student Experience** • *Enablers*: Students valued clarity, detail, and empathetic prompts; human+AI preferred to AI-only. • *Inhibitors*: Preservation of human contact and dialogue is essential; AI literacy and year group shaped receptiveness. **Culture** • *Enablers*: Openness to experimentation among staff and students. • *Inhibitors*: Fear that over-reliance could undermine the "human-ness" of feedback; cautious attitudes from staff with entrenched routines.
Materials or publications	• **Workshop materials** – slides and activity guides from the staff workshop (July 2025). Planned updates: student perspectives, getting the most from the tool (prompt training), maintaining your voice. • **Survey instruments** – staff confidence/competence survey, student focus group schedule. • **Student focus group data** – transcripts and thematic analysis notes (Year 2 and Year 3 participants). • **Draft staff-facing resource** – *Decision and Implementation Guide for AI-Supported Feedback* (in development). **Dissemination activities**: • Presentation at the Flexible Assessment Workshop (September 2025). • Briefing shared with the Faculty of Humanities Associate Dean for Realising Student Potential (September 2025). • Engagement with unit leads on the BSc Psychology to trial low-stakes peer review with AI-supported feedback. **Planned Publications and Dissemination** **Academic article** – focus group data will be written up and submitted to a peer-reviewed journal in the field of higher education. **Conference presentation** – dissemination at the *Assessment in Higher Education Conference 2026*, hosted in Manchester.

	Staff development – integration into the Academic Development Programme (2026 onwards). **Quantitative student survey** – planned for November 2026 to expand on qualitative insights.
Report approval and comments	*To be completed by a delegated person agreed by the workstream governance group.*

Cost Type	Description	Costs and Total
Staff	*Insert staff name and grade, post, FTE in the quarter in this quarter and the post.*	*Please add the cost of each post and the total claim for staff, this can then be added to the retrospective costs in the forecast form.*
Non-Staff	*Please add in any actuals (received and receipted paid in the quarter).* *Non staff can include any other approved cost category including:* • *Goods* • *Services (inc. Consultancy)* • *Travel* • *Fees*	*Please add the cost of each post and the total claim for staff, this can then be added to the actuals in the forecast sheet.*
Adjustments	*Please add any adjustments from previous quarters. This will be added or subtracted from your quarterly request for payment.* *Please give details of the original cost and the reason for the adjustment.*	*Please give the adjustment amount.*
Final reconciliation	*Please calculate the total costs of the pilot and the total income to ensure the claim has covered all eligible costs.*	
Request for payment	*The payment in the next box should include all staff costs, non-staff costs and adjustments total.* *By submitting this form, you are a confirming that the figures are correct to the best of your knowledge and the correct processes for recruitment procurement and selection have been followed.*	*Please provide:* • *total staff costs* • *total non-staff costs* • *total adjustments* • *final reconciliation* • *Total request for payment*