Advancing Project-Based Learning Through Iterative AI Feedback: Strengthening Girls' Confidence and Agency in Grade 9 Religious Experience

Marianne Rule

Merion Mercy Academy, Merion Station, Pennsylvania, USA

Abstract

The widespread accessibility of generative artificial intelligence (AI) has raised questions for educators regarding ethical use, student dependence, and intellectual agency. This mixed-methods action research study examined how iterative, teacher-calibrated AI feedback fostered girls' confidence, feedback literacy, and intellectual agency during Grade 9 project-based learning in a Roman Catholic all-girls' school. Sixteen girls participated in a 16-week classroom intervention centered on the Teacher-Calibrated Iterative Feedback Framework, which integrated AI-generated feedback, teacher conferencing, metacognitive reflection, and scaffolding reduction across iterative revision cycles. Data included pre- and post-surveys, student artifacts, reflections, AI interaction logs, teacher conference logs, interviews, and field notes, supporting polyangulation across student perceptions, feedback interactions, and revisions. Quantitative data were analyzed descriptively, and qualitative data were coded thematically to identify patterns in confidence and revision behavior. Findings indicated that students distinguished between exploratory AI feedback and evaluative teacher feedback, using each strategically. Over time, this differentiation was associated with increased confidence in revision decisions and strengthened intellectual agency in evaluating and applying feedback. Iterative feedback cycles supported by reflection and teacher conferencing also contributed to more critical engagement with revision processes. The study contributes an adaptable framework for integrating AI-

supported feedback within project-based learning while preserving the central pedagogical role of the teacher. Findings suggest that calibrated feedback systems combining AI, reflection, and teacher expertise can support student agency and self-directed revision practices in secondary classrooms. These findings will inform continued implementation of the Teacher-Calibrated Iterative Feedback Framework in future project-based learning contexts.

Glossary

Confidence: Students' perceived belief in their ability to complete analytical tasks successfully, demonstrated through their willingness to take intellectual risks, seek feedback, revise independently, and engage in complex or open-ended problems.

Feedback Literacy: The ability to interpret, evaluate, and apply feedback effectively to improve one's work, including making informed revision decisions independently.

Flint: An educational artificial intelligence platform designed to support teaching and learning by enabling teachers to design structured, curriculum-aligned AI interactions and feedback processes.

Growth: The measurable improvement between baseline and post-intervention data, demonstrated through polyangulated changes in confidence scores and improvements in revision behavior across artifacts, logs, and reflections.

Intellectual Agency: Students' capacity to make independent learning decisions, critically evaluate feedback sources, and take ownership of revisions.

Iterative Revision: The process of revising written work multiple times in response to feedback from AI, teacher, and self-reflection to improve reasoning, clarity, and alignment with assessment criteria.

Project-Based Learning: An instructional approach in which students engage in sustained inquiry around authentic, open-ended problems and produce a product demonstrating understanding and skill application.

Strengthening: Evidence of improvement in student confidence, agency, or feedback literacy supported by multiple quantitative and qualitative data sources including surveys, writing artifacts, revision logs, and reflections.

Advancing Project-Based Learning Through Iterative AI Feedback: Strengthening Girls' Confidence and Agency in Grade 9 Religious Experience

The widespread adoption of generative AI has prompted educators to reconsider how feedback, critical thinking, and intellectual agency are supported in classroom learning. In classrooms where students engage in open-ended analytical tasks, concerns about over-reliance on AI-generated responses have heightened the need for approaches that strengthen confidence, discernment, and ethical AI use.

Merion Mercy Academy teachers across departments have observed that girls often struggle with analytical tasks where there is no single correct answer. This was also evident in my Grade 9 Religious Experience class, where project-based assessments require students to interpret concepts, engage with multiple perspectives, and construct evidence-based arguments. The open-ended nature of these tasks can make revision difficult for students who lack confidence in their reasoning. This uncertainty may lead students to rely uncritically on AI-generated responses. Because theological inquiry requires discernment, interpretation, and complex reasoning, cultivating student confidence became essential for supporting both meaningful revision and responsible AI use.

These observations led me to examine whether iterative AI-generated feedback could support the development of confidence and feedback literacy, rather than function primarily as a shortcut. To explore this question, I designed the Teacher-Calibrated Iterative Feedback Framework, a feedback process using the AI writing assistant Flint (powered by Claude). The tool provided rubric-aligned feedback and reflective prompts grounded in students' writing to support evaluative decision-making during revision. Across the study, students engaged in repeated cycles of drafting, feedback, reflection, and revision with increasing independence. The

study explored how these feedback processes shaped students' confidence and engagement with revision.

Guided by this focus, the study addressed the following research question: How does iterative AI-generated feedback foster girls' confidence during project-based learning assessments in Grade 9 Religious Experience?

The students completed five scaffolded projects with iterative AI- and teacher-supported feedback designed to strengthen confidence, agency, and independent revision decisions. By embedding calibrated AI feedback into classroom practice and inviting students to critique AI responses, this study positions AI as a tool that can support student agency within curriculum-aligned projects.

Action research provided an appropriate methodological framework for examining changes in confidence and revision behaviors while simultaneously refining the design of the intervention within an authentic classroom context. This approach aligned with the cyclical nature of the project, allowing the feedback process to be adjusted in response to student experiences and reflections.

Literature Review

The rapid expansion of generative artificial intelligence (Gen AI) large language models (LLMs) has created a "tsunami effect" in education, with ChatGPT achieving 100 million active users just two months after its launch in November 2022 (García-Peñalvo, 2023, as cited in Volante et al., 2023, p. 40). Serious concerns about academic integrity quickly emerged as students gained the ability through LLMs to use simple prompts to generate sophisticated essays, bypassing genuine learning and streamlining cognitive tasks (Chiang et al., 2024; Escalante et al., 2023; Kos'myna et al., 2025; Lan & Chen, 2024; Volante et al., 2023). While AI detection

tools exist, they are largely ineffective, incorrectly flagging human-written text and revealing biases against non-native writers (Aaron et al., 2024; Escalante et al., 2023; Volante et al., 2023). Ensuring authentic learning remains a challenge in an AI-rich environment.

Integrity concerns are further compounded by findings from Kos'myna et al. (2025) showing that using Gen AI may lead to cognitive debt and other neural and behavioral consequences, with LLM users displaying weaker brain connectivity and underperforming at neural, linguistic, and behavioral levels. While LLMs offer “immediate convenience,” such findings highlight potential cognitive costs of Gen AI use, including weaker engagement in critical thinking and conceptual understanding (Kos'myna et al., 2025). This raises particular concerns for adolescent girls, who often enter learning tasks with lower confidence in critical thinking and self-direction. In a validation study of the “Student Instrument for measuring Confidence in Key Skills” (SICKS), Bray et al. (2020) found a statistically significant gender gap in perceived critical thinking confidence, with male students reporting significantly higher confidence in critical thinking. Girls may be more likely to over-rely on AI tools, potentially amplifying disparities in confidence that already exist.

Because students may treat AI as a shortcut, risking superficial learning (Chiang et al., 2024; Escalante et al., 2023; Volante et al., 2023), educators must prioritize meaningful assessments that foster higher-order and critical thinking across all educational levels (Aaron et al., 2024; Bray et al., 2020; Lan & Chen, 2024; Loyens et al., 2023). As Gen AI reshapes education, teachers are called to reconsider evaluation practices and ethical AI use (Aaron et al., 2024; Chiang et al., 2024; Escalante et al., 2023; Lan & Chen, 2024; Volante et al., 2023).

One effective response to concerns about superficial AI-supported learning and the need for deeper cognitive engagement is project-based learning, which engages students in authentic

problems and requires the application of critical thinking skills (Aaron et al., 2024; Bray et al., 2020; Coffman et al., 2023; Loyens et al., 2023; Volante et al., 2023). In addition, scaffolding shifts assessment focus from outcome to process and reduces the temptation to bypass authentic work with AI-generated submissions (Chiang et al., 2024; Lan & Chen, 2024; Volante et al., 2023). Iterative revision in response to feedback also provides opportunities for students to develop the capacity to ethically use AI as a learning tool (Aaron et al., 2024; Volante et al., 2023). Gen AI can deliver immediate, personalized feedback during assignments, and pedagogical AI agents may be calibrated to implement the teacher's instructional goals, guiding students through learning processes, providing real-time support, and aligning AI activities with learning objectives (Chiang et al., 2024; Escalante et al., 2023; Lan & Chen, 2024).

The most promising feedback models advocate for human-AI collaboration (Aaron et al., 2024; Escalante et al., 2023; Lan & Chen, 2024), where AI supports tasks like providing immediate, personal, and highly detailed feedback and human teachers focus on pedagogical expertise, emotional intelligence, and mentorship (Escalante et al., 2023; Lan & Chen, 2024). A blended approach combines the strengths of AI and human feedback (Escalante et al., 2023; Lan & Chen, 2024) and enhances "feedback literacy" in students, teaching them to effectively interpret and utilize feedback for continuous improvement (Aaron et al., 2024, p. 69). For girls, who often undervalue their own abilities, this kind of guided practice may be especially important for building confidence (Bray et al., 2020).

Confidence is closely connected to feedback literacy (Aaron et al., 2024; Volante et al., 2023) and confidence in self-direction is a statistically significant predictor of students' aspirations and goals (Bray et al., 2020). One vital aspect of self-direction is a student's confidence in their ability to "use peer, teacher or expert feedback to change your work" (Bray et

al., 2020, p. 5). Effectively leveraging feedback supports students' educational journeys and strengthens their confidence, demonstrating the connection between feedback literacy and self-directed learning (Aaron et al., 2024; Chiang et al., 2024; Escalante et al., 2023; Bray et al., 2020).

However, when it comes to confidence, female students report lower levels in areas like critical thinking, creativity, and using technology for learning (Bray et al., 2020). *The Girls' Futures Report* (Girls' Day School Trust, 2022) emphasizes that girls experience a marked dip in confidence during adolescence, with decreased belief in their abilities and future opportunities. These findings highlight the urgency of educational approaches that deliberately strengthen girls' confidence, providing opportunities to develop skills, engage in risk-taking, and critically evaluate information (Aaron et al., 2024; Bray et al., 2020; Coffman et al., 2023; Lan & Chen, 2024; Loyens et al., 2023; Volante et al., 2023).

With these themes in mind, I designed a project-based learning environment for my Grade 9 Religious Experience class in which each assessment was scaffolded and supported by a teacher-calibrated AI agent providing iterative feedback. Students revised in response to feedback, reflected on their revisions before progressing to the next project stage, and engaged with AI as a structured support for learning rather than a replacement for cognitive effort. I anticipated that by the end of the study, this design would foster increased confidence in critical thinking, feedback literacy, and using technology for learning.

Research Context

Merion Mercy Academy is an independent, all-girls' Roman Catholic college preparatory high school (Grades 9–12) sponsored by the Sisters of Mercy. Located in Merion Station,

Pennsylvania, the school draws students from the greater Philadelphia region, including neighboring New Jersey.

The participants in this study were 16 girls enrolled in one section of Grade 9 Religious Experience, the first theology course that students (ages 14 to 15 years old) take upon entering the school. The course met an average of 3.5 hours per week over the 16-week study period. I selected this course because it represents students' initial engagement with the theology curriculum and the semester is fully project-based, aligning closely with the focus of this action research. The class size of 16 was also appropriate for action research (Mertler, 2024). Participants were selected through convenience sampling since they were members of my existing class.

I was the classroom teacher of this course and held a dual role as both instructor and researcher. Students, parents, and caregivers were informed about the study and invited to opt out if they did not wish to participate. To ensure confidentiality, the girls were assigned pseudonyms (letters of the alphabet) in all data collection, analysis, and reporting. No identifying information is included in this final report. All data were stored securely and accessed only by the researcher. Students and families were informed that choosing not to participate would have no impact on students' grades or standing in the course.

The Action

The Teacher-Calibrated Iterative Feedback Framework consisted of four interconnected components: targeted AI-generated feedback aligned to rubric criteria and delivered through open-ended critical thinking questions; access to one-on-one teacher conferencing; structured metacognitive reflection following brainstorming, drafting, and final project submission; and the reduction of scaffolding over time. Across each project, students moved through iterative cycles

of drafting, AI and teacher feedback, reflection, and revision. Each cycle informed subsequent revisions and increased student independence. Scaffolds such as project steps, sample prompts, and reminders to critically evaluate AI-generated feedback were gradually removed to shift greater responsibility for revision decisions to students.

As students were introduced to each project for the semester, they completed brainstorming and initial drafting before receiving access to AI feedback. For Unit 1, the girls were offered four project options, allowing for meaningful choice within the project-based framework. Each student received a checklist tailored to her project choice and organized into three stages: brainstorming, drafting, and feedback. This structure supported student agency within the revision process while ensuring alignment with project expectations.

During brainstorming and drafting, the girls were encouraged to seek clarification from the teacher or peers as needed. Once students had developed a core piece of their project work (which may not have been a complete draft), they received a link to access Flint for AI-generated feedback and a separate link to schedule one-on-one teacher conferences. Access to both AI and teacher feedback allowed students to decide how and when to seek support during revision. This structure shifted the revision process from primarily teacher-mediated feedback toward student-directed feedback cycles supported by AI-generated responses.

Students participated in an introductory lesson on Flint's interface, including how to upload their work and navigate the feedback chat. To scaffold effective use, the girls were provided with sample prompts aligned to rubric criteria to support productive engagement. I explicitly addressed the limitations of AI and instructed the girls to challenge feedback they disagreed with by clarifying the type of guidance they wanted. This approach positioned AI feedback as something students were expected to question, evaluate, and selectively apply.

Students were not required to use Flint a specific number of times. Instead, they were encouraged to return to the platform as needed, in or outside of class, prior to final project submission. This flexible structure supported individualized revision cycles.

I calibrated Flint's feedback using project rubrics and structured responses in three parts: a rubric-based comment, direct quotations from student work showing where the feedback applied, and open-ended questions designed to prompt reflection and revision decisions. Student reflections and field notes informed iterative refinement of the AI feedback structure, indicating that further calibration was needed to reduce feedback wordiness and enhance clarity and usability. In response, I recalibrated Flint from a balanced response style to a concise response style starting in Unit 2. This adjustment reduced the volume of feedback and improved clarity. Later student reflections suggested that the streamlined outputs were easier to interpret and apply. These refinements reflected the cyclical nature of the action research process, in which student reflections and field notes informed ongoing modifications to the intervention.

For Units 2 through 5, the iterative feedback structure remained consistent while scaffolding was gradually reduced to increase student independence. Rather than providing a detailed checklist of steps, students were asked to establish their own goals for the brainstorming and drafting stages of their projects. The girls continued to receive access to feedback through Flint and the option to schedule one-on-one teacher conferences. By Unit 3, students were independently directing much of their feedback process, using Flint and teacher conferences strategically based on their goals and time management. For Units 4 and 5, I no longer provided sample prompts to promote independent interpretation of rubric criteria.

For all five units, the girls also completed structured metacognitive and evaluative reflections after brainstorming, drafting, and final project submission. These reflections were

designed to strengthen confidence in reasoning and revision decisions throughout project-based learning. The primary difference across units was the gradual removal of teacher-provided scaffolds and the transfer of goal-setting responsibility to students. As a result, students engaged more independently with drafting and AI feedback. Detailed calibration protocols representative of Units 2 through 5, along with examples of calibrated feedback, are provided in Appendix A and Appendix B.

Data Collection

This action research project examined how iterative AI-generated feedback fostered girls' confidence during project-based learning in Grade 9 Religious Experience. Following Mertler's (2024) cyclical action research model, the study employed mixed-methods to capture measurable outcomes and descriptive accounts of students' confidence, feedback literacy, critical thinking, and revision practices. This design emphasized qualitative indicators of student experience while also considering quantitative measures of growth. Data were collected through writing artifacts, surveys, reflections, interviews, AI interaction logs, and field notes.

At the beginning of the study, the girls completed an analytical writing prompt and a pre-survey adapted from SICKS measuring confidence in critical thinking and creative iteration. These four confidence items were measured on a six-point Likert scale (1 being "Not Confident at All" and 6 being "Totally Confident") adapted from the original five-point version (Bray et al., 2020) to better align with the study's operational definition of confidence. Items were selected to capture confidence in open-ended reasoning and iterative improvement, central components of feedback literacy. Responses were averaged to generate a composite baseline and post-study confidence score for each participant.

Throughout the semester, students completed five scaffolded, project-based learning assessments where they engaged with AI-generated feedback through Flint. Written reflections were collected at three points during each project: after brainstorming, after draft revisions, and after project submission. In these reflections, the girls considered their confidence, their revision decisions in response to feedback from multiple sources, and their development as learners.

Flint recorded student interactions, capturing the frequency, timing, and content of feedback requests. These logs were used to polyangulate the girls' self-reported confidence and revision practices by providing evidence of how they engaged with AI-generated feedback across iterative cycles. Teacher conference logs documented when students sought feedback during projects and the purposes of those interactions, including asking questions, seeking reassurance about strengths, and requesting critical feedback to guide revisions. Field notes further provided contextual insight into how students balanced teacher and AI feedback sources during the revision process.

At the beginning of the study, each girl recorded a 5 to 10-minute video responding to open-ended questions about her comfort using AI as a feedback resource, confidence in completing projects, and what she valued in her teacher's role supporting project-based learning. Following the intervention, students completed the same analytical writing prompt and post-survey administered at baseline, allowing for direct comparison of changes in confidence over time. They also participated in post-study exit interviews reflecting on their experiences with AI-supported revision, feedback use, and the project-based process.

Data Analysis

Following data collection, analysis focused on identifying growth in both measurable outcomes and student perceptions. Quantitative data from surveys and rubric scores were

analyzed descriptively to compare mean changes in confidence and feedback literacy. Agency was assessed through observable behavioral indicators across multiple data sources. Rubric data from baseline and end-of-study writing samples were compared to evaluate critical thinking growth. Flint logs were reviewed to identify engagement patterns, such as consultation frequency and revision iterations.

All data were de-identified, and qualitative data were coded iteratively using both deductive categories aligned with research constructs (confidence, feedback literacy, agency, and revision behavior) and inductive themes emerging from the data. Conflicting data were examined to refine interpretations and ensure an authentic representation of student voice.

Polyangulation across surveys, interviews, classroom observations, artifacts, teacher conferencing forms, and Flint interaction logs contributed to credibility by identifying consistent patterns. Guided by Mertler's approach (2024), the integration of data sources yielded four key themes related to confidence, metacognition, agency, and intellectual responsibility within an all-girls setting.

Discussion of Findings

The Pattern of Girls' Confidence Depends on the Nature of Feedback

Analysis identified patterns in girls' confidence associated with different feedback sources, reflecting a blended approach that integrated both human and AI collaboration. Flint interaction logs showed that the girls used AI during drafting to check mechanics, clarify ideas, and identify missing details. All students sought AI feedback before consulting the teacher during the five assessments, demonstrating an emerging capacity to revise independently. Classroom observations confirmed this trend during each project: even when the teacher was available, participants self-initiated AI feedback and brought polished drafts to conferences.

Most girls perceived teacher feedback as authoritative and evaluative, closely aligned with rubric requirements and grading criteria. As Student E explained, “the teacher, who will be ultimately assessing your work, knows what is being looked for in that work more than an AI software.” The girls consistently requested instructor input later in their drafting, often after completing a major section or a near-final draft. During the study, all but one teacher conference focused on final rubric-aligned feedback right before project submission. Observations indicated that all participants consistently arrived prepared to discuss their work and sought teacher confirmation before submitting assignments. These patterns are consistent with research emphasizing human-AI collaboration in feedback models (Aaron et al., 2024; Escalante et al., 2023; Lan & Chen, 2024). Escalante et al. (2023) and Lan and Chen (2024) describe this division where AI supports immediate, highly detailed feedback and human instructors contribute pedagogical expertise and mentorship.

Compared to pre-action responses, 12 of 16 girls self-reported higher confidence in using feedback from a teacher, peer, or AI independently on the post-action Likert scale, particularly in understanding and applying suggestions to make their ideas stronger. In the post-study responses, these twelve girls reported themselves as “very confident” (5) or “totally confident” (6) in this area.

These results align with the study’s initial expectation that confidence in critical thinking, feedback literacy, and technology would increase over time. This confidence in independence coexisted with continued reliance on teacher expertise: post-action survey results showed that ten participants preferred teacher feedback, while six trusted AI and teacher feedback, valuing instructor guidance for evaluative confirmation before project submission. Together, these observations and survey responses illustrate how AI and human guidance functioned in distinct

but complementary ways across drafting, revision, and submission. This pattern suggests a redistribution of feedback authority, where students differentiated between exploratory support from AI and evaluative validation from the teacher.

Targeted, Open-Ended AI Feedback Supports Girls' Metacognition and Agency

Throughout each project, the calibrated AI feedback structure shaped how students approached revision. Because Flint provided a single, focused point of feedback anchored in direct quotes from the girls' work and also posed open-ended critical thinking questions, students were prompted to decide how to revise. This structured format remained consistent across all five units. Appendix B illustrates the three-part model in a representative interaction, demonstrating how each AI output followed the same design.

The girls emphasized that Flint "doesn't give you the answers" (Student B) but instead allowed them to "bounce off" its feedback (Student J), reflecting heightened metacognitive engagement. These findings align with Lan and Chen's (2024) findings that AI-powered feedback enables real-time iteration and personalized support for self-regulated learning.

Flint interaction logs further suggested that many participants engaged in intellectual agency by prompting for clarification, selectively revising their work, and requesting rechecks after changes. AI functioned as a thought partner rather than an authority, helping most girls identify gaps in reasoning or missing perspectives. For example, during the Unit 2 project, Student O asked Flint, "What's missing that would help a viewer understand why people find meaning in this tradition," and Student M asked, "Is this information free of stereotypes." This iterative process reinforced student ownership and positioned AI as a support for learning rather than a substitute for intellectual agency.

The Evaluation of AI Feedback Reflects Girls' Developing Intellectual Responsibility

In baseline reflections, the girls acknowledged AI's limitations, including statements from Students K, G, and D, such as "oftentimes it's wrong," "[it] is often faulty," and "[it] always doesn't know." Students responded to calibrated feedback by critically evaluating suggestions they perceived to be inaccurate or incomplete. When AI feedback did not align with their reasoning, students challenged AI suggestions and requested clarification. This pattern demonstrated emerging responsibility in evaluating automated feedback, aligning with recommendations that students should be explicitly taught to interrogate AI outputs (Chiang et al., 2024; Escalante et al., 2023; Volante et al., 2023).

When the girls first encountered inaccurate or questionable AI feedback, many frequently sought teacher guidance. These moments reinforced critical evaluation of AI output rather than deferring to it. Across projects, participants independently evaluated AI feedback and revised their work selectively. These patterns suggest emerging judgment and responsibility in AI-supported, project-based environments. In one instance, Flint incorrectly assumed that a "Faith in a Box" project designed to teach four- to five-year-olds required explanations appropriate for that audience. Student P responded by clarifying, "It's supposed to be explained to high school students," correcting the AI's misunderstanding of project expectations. This exchange illustrates students actively evaluating and adjusting AI feedback rather than accepting it uncritically, taking responsibility for ensuring alignment with project goals, and reflecting emerging intellectual autonomy.

Girls' Confidence in Complex, Open-Ended Tasks Aligns with Iterative AI Feedback

Engagement with calibrated, iterative AI feedback corresponded with students' self-reported confidence and their willingness to engage with open-ended tasks without a single

correct answer. For girls entering the study with high self-reported confidence, AI feedback supported iterative approaches to learning. As Student N reflected, “I am very confident in my work and I work hard to make it better with feedback and suggestions.” Flint logs indicated that many students frequently engaged in iterative revisions, refining ideas and maintaining ownership of their work. This pattern aligns with research showing that confidence in critical thinking and self-direction is a statistically significant predictor of students’ active engagement with learning and their sense of purpose (Bray et al., 2020).

Patterns in AI use also aligned with broader confidence trends. Most participants reported increases in confidence over the course of the study in responding to complex questions, experimenting with ideas, and proposing solutions to challenging problems. Eleven students reported measurable gains, and the four girls among the most frequent Flint users reported the greatest gains. Student C’s exit interview captured this shift, “My confidence definitely went up because usually I would be like … there’s only one right answer and I'd put all my effort into it, but now I want to express myself in my projects and my thoughts.” Although variation existed among participants, the overall upward shift suggests that structured AI-supported revision cycles may have supported greater confidence when engaging with complex tasks.

Conclusions

Iterative, teacher-calibrated AI feedback integrated into project-based learning was associated with growth in girls’ confidence, metacognitive awareness, and agency. This study contributes the Teacher-Calibrated Iterative Feedback Framework, a replicable system integrating rubric-aligned AI feedback, teacher conferencing, iterative revision, structured metacognitive reflection, and scaffold reduction to position AI as an embedded pedagogical support. This framework addresses the need across disciplines for structured AI integration.

Across five units, the girls engaged in repeated cycles of drafting, receiving feedback from Flint and the teacher, and revising their work. Scaffolding was gradually reduced, and students assumed greater responsibility for setting revision goals and determining when and how to seek feedback. Data from surveys, interviews, writing artifacts, and AI interaction logs indicate that most participants developed stronger confidence in open-ended tasks and greater independence in revision.

These findings suggest that blended feedback models combining human expertise with calibrated AI support can function as complementary resources. The model shifted students from passive recipients of feedback to active participants in interpreting, evaluating, and applying feedback throughout revision. The girls generally used AI for immediate drafting support while positioning teacher feedback as evaluative guidance prior to submission. While most participants reported increases in confidence with open-ended tasks without a single correct answer, three students reported decreases in post-study Likert-scale responses. These students required reminders to complete reflections, and after the baseline and post-study analytical writing prompts, reported only spending part of the allotted time on their paragraph before shifting focus to work for another class. Limited engagement with both reflection and sustained task focus may have constrained confidence gains.

For educators integrating AI tools into project-based learning, intentional scaffolding is essential. Explicit modeling, calibrated prompts, and clear expectations support productive use of the tool. Gradual reduction of structured supports such as checklists and prescribed prompts may promote ownership of learning and strengthen students' ability to independently interpret and apply feedback. Teachers should also explicitly teach students how to evaluate AI-generated suggestions critically and position AI as a revision partner rather than an authority. Providing

opportunities for reflection after feedback cycles can deepen metacognition and reinforce intellectual responsibility. Importantly, blended feedback approaches were associated with stronger patterns of confidence and agency within this implementation context.

Implications extend beyond the context of this study. The calibrated feedback model provides a replicable framework for integrating AI into project-based learning while preserving pedagogical oversight. Teachers can adapt rubric-aligned AI calibration, iterative revision cycles, and reflection prompts to their contexts. The model will continue to be offered through future cohorts of Religious Experience and expanded into an additional Semester 2 class by other teachers with project-based assessments. Ongoing refinement will continue based on student feedback and classroom implementation.

This study was context bound and its time constraints limited examination of long-term impact, as data collection reflected one semester of implementation. Future research could explore variations in scaffold levels, compare AI-supported and teacher-only feedback models, and expand implementation to additional subject areas to test transferability.

Reflection Statement

My work with AI began on November 30, 2022, when I first explored large language models to understand their instructional potential and how to guide students in using them ethically and responsibly. This action research project allowed me to analyze how AI reshapes feedback, agency, and authentic learning environments. Working alongside my students as they navigated iterative AI-supported revision allowed me to witness how structured calibration and reflective practice can transform feedback into a tool for intellectual growth. Student questions and concerns about AI, especially regarding privacy, future employment, and ethical implications, pushed me to think more carefully about how technology is framed within

pedagogy. The project shifted my understanding of AI feedback from a supplemental tool to a feedback system directly embedded into the pedagogical structure of a class.

I am grateful to ICGS and the GARC community for fostering research that connects purpose in girls' education with practice in meaningful and impactful ways. I also appreciate the mentorship and institutional support that made this work possible. My research advisor AnnMarie Zigrossi has been a tremendous support along with my in-school mentor Philip Vinogradov. I am especially grateful to my President Marianne Grace for approving my participation and to my former Assistant Head of School for Academics Valerie Prucnal who shared the opportunity with me. The overwhelming support of my students and their families has truly made this journey remarkable.

This study marks the beginning of an ongoing cycle of refinement as I continue to explore how blended feedback models can strengthen girls' confidence and agency in meaningful ways.

References

Aaron, L., Abbate, S., Allain, N. M., Almas, B., Fallon, B., Gavin, D., Gordon, C., Jadamec, M., Merlino, A., Pierie, L., Solano, G., & Wolf, D. (2024). Student evaluation practices and assessment strategies. In Optimizing AI in higher education: SUNY FACT guide (pp. 69–72). SUNY Press. https://doi.org/10.2307/jj.20522984.35

Bray, A., Byrne, P., & O'Kelly, M. (2020). A short instrument for measuring students' confidence with "key skills" (SICKS): Development, validation and initial results. *Thinking Skills and Creativity*, *37*, 100700. https://doi.org/10.1016/j.tsc.2020.100700

Chiang, Y. V., Chang, M., & Chen, N.-S. (2024). Can generative AI help realize the shift from an outcome-oriented to a process-outcome-balanced educational practice? *Educational Technology & Society, 27*(2), 347–385. https://doi.org/10.30191/ETS.202404_27(2).TP04

Coffman, S., Iommi, M., & Morrow, K. (2023). Scaffolding as active learning in nursing education. *Teaching and Learning in Nursing, 18*, 232–237. https://doi.org/10.1016/j.teln.2022.09.012

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*, 57. https://doi.org/10.1186/s41239-023-00425-2

Kos'myna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X. H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing tasks. arXiv. https://arxiv.org/abs/2506.08872

Girls' Day School Trust. (2022). *The girls' futures report: Shaping skills, driving confidence and creating leaders.* https://www.gdst.net/wp-content/uploads/2023/09/Girls-Futures Research-updated-01.08.23.pdf

Lan, Y.-J., & Chen, N.-S. (2024). Teachers' agency in the era of LLM and generative AI: Designing pedagogical AI agents. *Educational Technology & Society, 27*(1), i–xviii. https://doi.org/10.30191/ets.202401_27(1).pp01

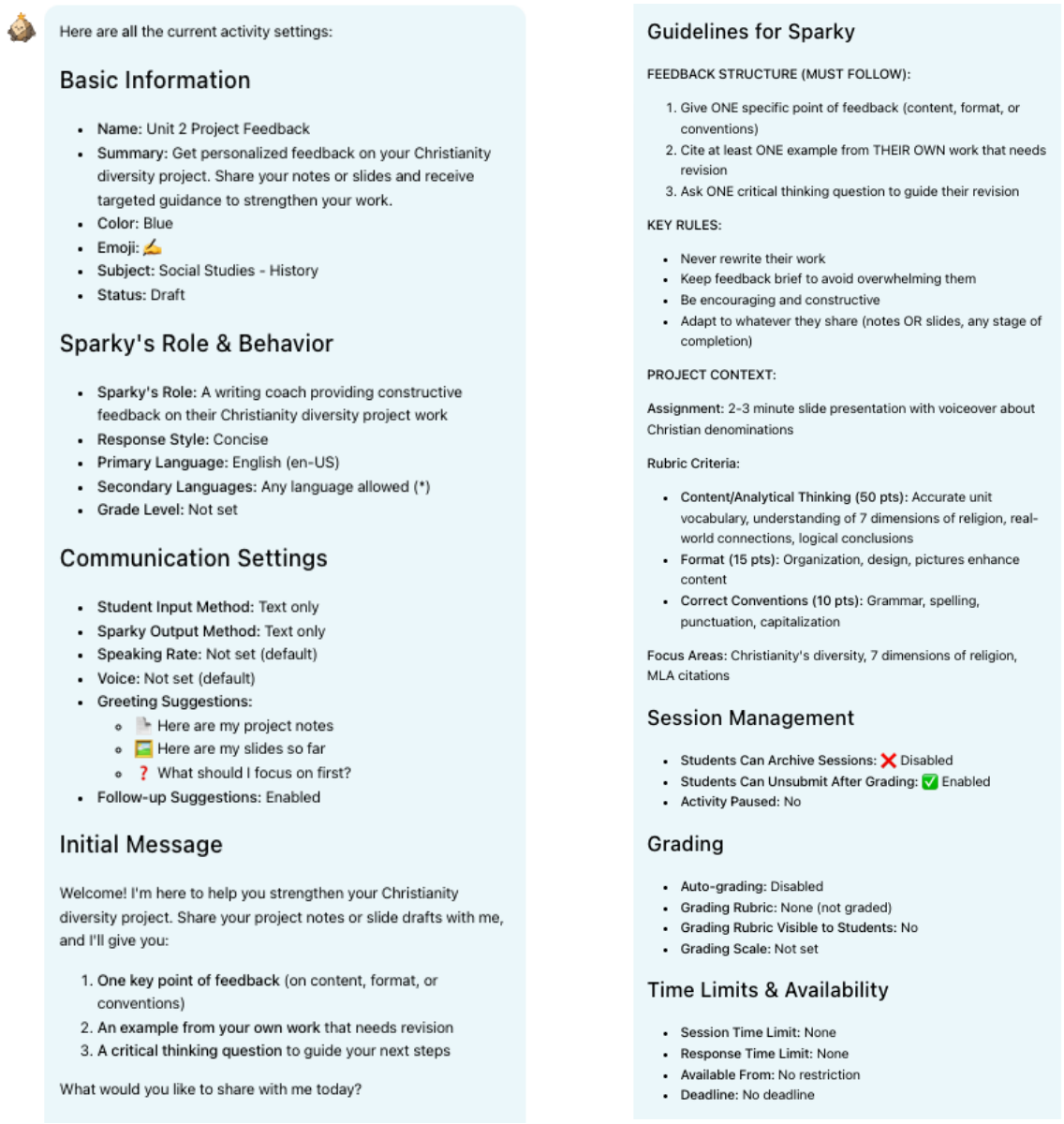
Loyens, S. M. M., van Meerten, J. E., Schaap, L., & Wijnia, L. (2023). Situating higher-order, critical, and critical-analytic thinking in problem- and project-based learning environments: A systematic review. *Educational Psychology Review, 35*, 39. https://doi.org/10.1007/s10648-023-09757-x

Mertler, C. A. (2024). *Action research: Improving schools and empowering educators* (6th ed.*)*. SAGE.

Volante, L., DeLuca, C., & Klinger, D. A. (2023). Leveraging AI to enhance learning. *Phi Delta Kappan, 105*(1), 40–45. https://doi.org/10.1177/00317217231197475

Appendix A

Sample Flint Project Calibration and Student Welcome Message



Here are all the current activity settings:

Basic Information

- **Name:** Unit 2 Project Feedback
- **Summary:** Get personalized feedback on your Christianity diversity project. Share your notes or slides and receive targeted guidance to strengthen your work.
- **Color:** Blue
- **Emoji:** ✍️
- **Subject:** Social Studies - History
- **Status:** Draft

Sparky's Role & Behavior

- **Sparky's Role:** A writing coach providing constructive feedback on their Christianity diversity project work
- **Response Style:** Concise
- **Primary Language:** English (en-US)
- **Secondary Languages:** Any language allowed (*)
- **Grade Level:** Not set

Communication Settings

- **Student Input Method:** Text only
- **Sparky Output Method:** Text only
- **Speaking Rate:** Not set (default)
- **Voice:** Not set (default)
- **Greeting Suggestions:**
  - 📄 Here are my project notes
  - 🖼️ Here are my slides so far
  - ❓ What should I focus on first?
- **Follow-up Suggestions:** Enabled

Initial Message

Welcome! I'm here to help you strengthen your Christianity diversity project. Share your project notes or slide drafts with me, and I'll give you:

1. **One key point of feedback** (on content, format, or conventions)
2. **An example from your own work** that needs revision
3. **A critical thinking question** to guide your next steps

What would you like to share with me today?

Guidelines for Sparky

FEEDBACK STRUCTURE (MUST FOLLOW):

1. Give ONE specific point of feedback (content, format, or conventions)
2. Cite at least ONE example from THEIR OWN work that needs revision
3. Ask ONE critical thinking question to guide their revision

KEY RULES:

- Never rewrite their work
- Keep feedback brief to avoid overwhelming them
- Be encouraging and constructive
- Adapt to whatever they share (notes OR slides, any stage of completion)

PROJECT CONTEXT:

Assignment: 2-3 minute slide presentation with voiceover about Christian denominations

Rubric Criteria:

- **Content/Analytical Thinking (50 pts):** Accurate unit vocabulary, understanding of 7 dimensions of religion, real-world connections, logical conclusions
- **Format (15 pts):** Organization, design, pictures enhance content
- **Correct Conventions (10 pts):** Grammar, spelling, punctuation, capitalization

Focus Areas: Christianity's diversity, 7 dimensions of religion, MLA citations

Session Management

- **Students Can Archive Sessions:** ❌ Disabled
- **Students Can Unsubmit After Grading:** ✅ Enabled
- **Activity Paused:** No

Grading

- **Auto-grading:** Disabled
- **Grading Rubric:** None (not graded)
- **Grading Rubric Visible to Students:** No
- **Grading Scale:** Not set

Time Limits & Availability

- **Session Time Limit:** None
- **Response Time Limit:** None
- **Available From:** No restriction
- **Deadline:** No deadline

Figure A.1

Flint Unit 2 Feedback Calibration for Christian Ecumenism Project

Note. Screenshot of the Flint interface used in this study. Reproduced with permission from Flint.

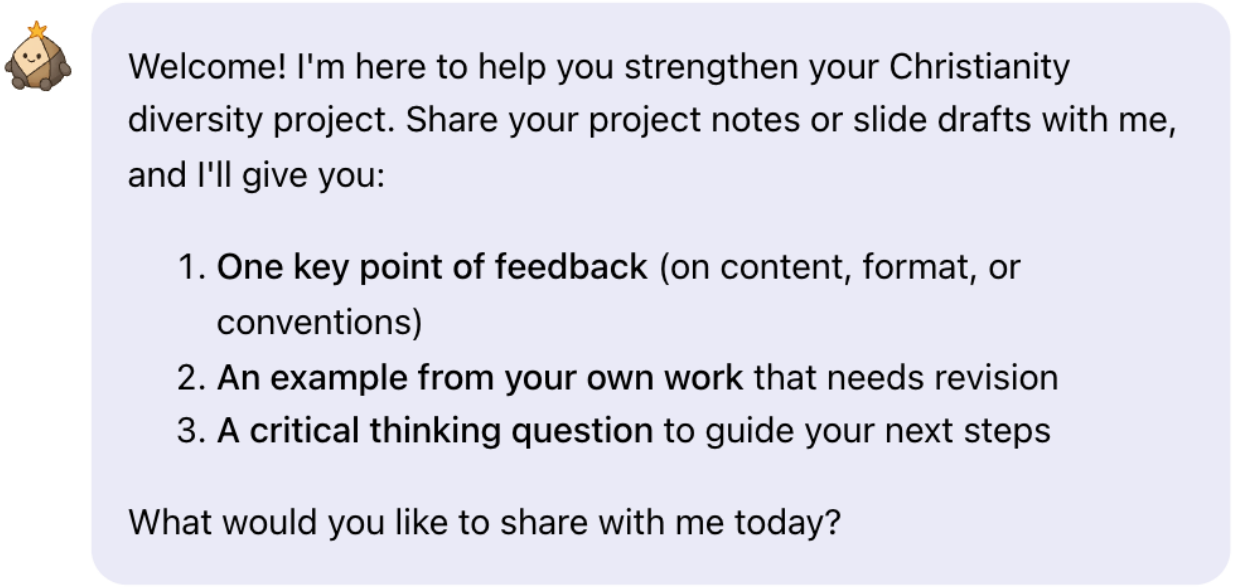


Figure A.2

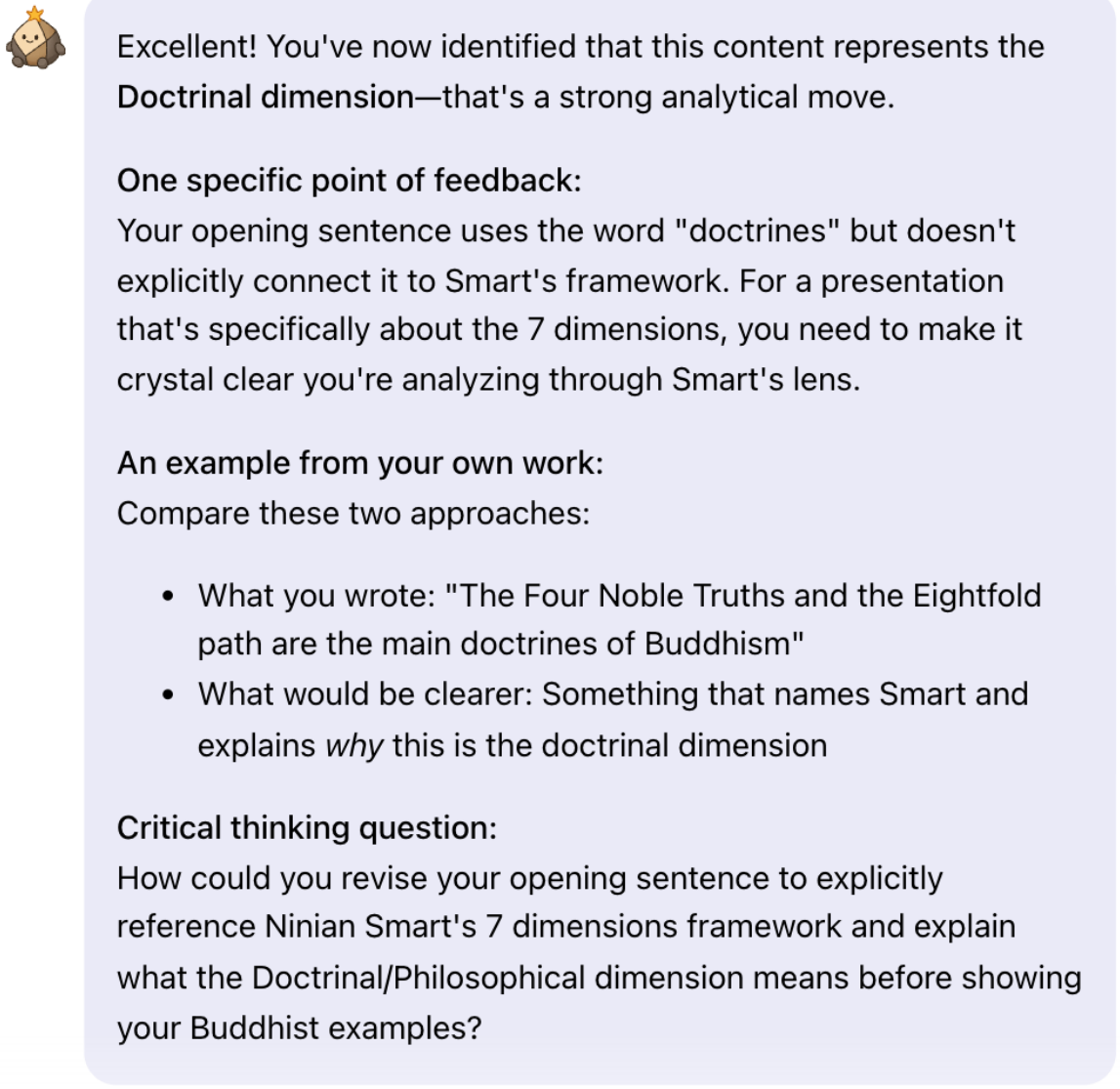
Flint Student Welcome Message for Unit 2 Project Feedback

Note. Screenshot of the Flint interface used in this study. Reproduced with permission from Flint.

Appendix B

Sample of Teacher-Calibrated Flint Feedback

Unit 3 Flint feedback provided to Student N after revision and resubmission.



Excellent! You've now identified that this content represents the **Doctrinal dimension**—that's a strong analytical move.

One specific point of feedback:
Your opening sentence uses the word "doctrines" but doesn't explicitly connect it to Smart's framework. For a presentation that's specifically about the 7 dimensions, you need to make it crystal clear you're analyzing through Smart's lens.

An example from your own work:
Compare these two approaches:

- What you wrote: "The Four Noble Truths and the Eightfold path are the main doctrines of Buddhism"
- What would be clearer: Something that names Smart and explains *why* this is the doctrinal dimension

Critical thinking question:
How could you revise your opening sentence to explicitly reference Ninian Smart's 7 dimensions framework and explain what the Doctrinal/Philosophical dimension means before showing your Buddhist examples?

Note. Screenshot of the Flint interface used in this study. Reproduced with permission from Flint.