Available online at http://www.journalcra.com

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

International Journal of Current Research
Vol. 17, Issue, 07, pp.33926-33940, July, 2025
DOI: https://doi.org/10.24941/ijcr.49150.07.2025

ISSN: 0975-833X

RESEARCH ARTICLE

NEEDS ASSESSMENT AND TECHNICAL SPECIFICATION FOR AN AI FEEDBACK SYSTEM FOR STUDIO-BASED LEARNING

*Henry Wanakuta, PhD

Stephen F. Austin State University, Texas, USA

ARTICLE INFO

Article History:
Received09[th] March, 2025
Received in revised form
21[st] April, 2025
Accepted 19[th] May, 2025
Publishedonline30[th] July ,2025

Keywords:

Studio-Based Learning, Feedback, AI-Enabled Feedback System, Technical Specification.

Corresponding author:
Henry Wanakuta

ABSTRACT

Studio- based learning (SBL) learning environments, typical of design and creative disciplines, rely on formative feedback and critique. Yet, it is difficult for educators to provide timely, substantive feedback to each student, and students report delay and brevity in comments. The current study fills that void by conducting needs assessment and conceptualizing an AI-enabled feedback system to complement studio-based learning. The study employed mixed-methods design: educator focus group interviewing (N=11) in Kenya, Saudi Arabia, and USA, and student survey (N=150) of studio classes. Findings indicate widespread frustrations about turn-around time to provide comments, comments depth, and grading workload. Kenyan, Saudi Arabian, and American instructors described feeling bogged down by high class enrollments and grading burden, as is also seen in studies on rising student-staff ratios and comments lag time. Students also described waiting weeks to receive feedback and that the comments largely lacked depth to enact change, as also seen in previous studies on discontent with feedback. Each of these groups also expressed optimism that an AI-enabled feedback system could provide more timely, individualized critiques, provided it is well-designed. The principal deliverable of this study is an elaborated technical specification sheet (see Appendix) of features and structure of conceptualized AI system. The study describes functional requirements (e.g., an AI engine to generate rubric-matched formative comments), technical design (integration into learning platforms, data safeguarding), and collaboration plans to engage a software developer and development cost estimate. Finally, the study proposes an orderly process of future development: sequential pilot testing, user feedback-led fine-tuning, and rollout to scale up. This project, financed by a $10,000 Stephen F. Austin State University JACK STARS grant, paves the ground for closing the pedagogical disparity of AI in SBL feedback.

Citation: Henry Wanakuta. 2025. "Needs assessment and technical specification for an ai feedback system for studio-based learningp". *International Journal of Current Research, 17, (07), 33926-33940.*

INTRODUCTION

One of the cornerstones of effective learning, particularly in studio classes where student creative works evolve from successive rounds of critique, is feedback. There is research that suggests excellent feedback has the capacity to greatly enhance student outcomes (Hattie & Timperley, 2007). Studio-Based Learning (SBL) is used in as disparate fields as architectural design, product design, civil engineering, and performing and fine arts, and encourages "learning by doing" and producing a product or final performance. Faculty in studio classes normally involve students in critiques and one-on-one meetings, providing formative feedback on works-in-progress. The optimal type of feedback, of course, is timely, concrete, and encourages students to revise and refine works and to enhance skills. In fact, provision of timely individual feedback to every student is problematic. Shortage of instructional staff or high enrollments confront most SBL programs, which results in lateness and brevity of comments. Students rightly objected to delay in receipt of feedback following an assignment, for example, one student in one national survey, who complained, *"Why can work be handed back late, but you can't submit it late?"*, complaining of what is close to double standards. Once more, students also receive confusing or superficial comments, whereupon, as one student was heard to remark, students might be *"disappointed and frustrated when feedback is vague [or] too succinct"*. Such issues can interfere with studio model gains. Students, as is well appreciated, are more likely to respond to feedback if it is provided on time (Hanmore *et al*., 2023). Conversely, undue delay has the result of undermining timeliness of feedback to what is being worked on, and superficial, hasty comments might be powerless to enable improvement (Hanmore *et al*., 2023). Instructors, however, experience heavy workload to grade as class enrollments increase and support (e.g., instructional assistants) is not provided. Such workload is detrimental to timeliness as well as quality of feedback issued (Robinson *et al*., 2013; Hanmore *et al*., 2023). These instructional design issues in SBL contexts recognize an acute needs gap: *How do you provide students with swift, nuanced feedback without further burdening instructors?* One possible answer is to insert artificial intelligence (AI) into the feedback process. AI-enabled learning software has evolved quite strongly in recent years, and there is potential for automated assessment and custom-created feedback generation. For example, natural language processing algorithms can read written assignments and provide immediate formative feedback, as in certain intelligent tutoring systems. Preliminary studies suggest that AI-enabled feedback could help to deliver timeliness and detail students prefer while freeing up routine grading for instructors. Finally, it is to be noted that the goal is to support, not replace, instructors: AI does routine feedback and one-pass critiques, freeing instructors to focus on higher-order feedback and one-on-one mentoring. Yet, to implement an AI feedback system, planning needs to be conducted to ensure that it meets student and teacher needs. Too often, educational technologies break down because they were not user needs-driven or contextually embedded in instructional practices. An extensive needs analysis and technical specifications that address the actual context of SBL, therefore, need to be conducted before any AI system is designed. This paper reports on such an investigation. This initial study did not design or implement an AI tool but instead engaged in acquiring stakeholder needs and collating them into an accurate design specification. In determining what the system needs to accomplish and how it needs to accomplish it, the study hopes to establish a strong foundation for future development work to build upon. The research questions for this study were:

- What are the specific pain points and requirements related to feedback in studio-based learning, from the perspectives of instructors and students?
- What features and capabilities should an AI-driven feedback system have to effectively address those needs? and
- How might such a system be implemented and evaluated in practice (development considerations, pilot deployment, etc.)?

METHODS

Research design: The research used a mixed-methods sequential exploratory design (Creswell & Plano Clark, 2018) to conduct needs analysis for an AI feedback system. During phase one, qualitative data were collected using focus group interviews of instructors to gain an in-depth understanding of current feedback processes and issues in studio classes. During phase two, an undergraduate student questionnaire was used to quantify incidence of specific problems (e.g. timeliness of feedback, perceived quality of feedback) and to measure student receptiveness to possible AI solutions. Qualitative findings informed survey design, and quantitative findings were used to underpin and augment qualitative findings. This permitted data triangulation and more informed insight into needs. The research was approved by our university's Institutional Review Board, and informed consent was granted by participants.

Participants

Instructors (Focus groups): The study engaged 11 instructors (5 women, 6 men) who taught design disciplines through studio classes. For diversity of educational context, participants were recruited from three countries: Kenya (n=3), Saudi Arabia (n=4), and the United States (n=4). They were veterans of studio contexts (mean = 8.7 years, range 3–20 years). They came from graphic design, architecture, interior design, and studio art, where project assessment and repeated critique were core pedagogical practices. Purposive and snowball sampling were used to recruit instructors, leveraging professional contacts and following up using referrals. Statistical representativeness was not the study target but rich insightfulness in cross-institutional and cross-cultural contexts. The number of participants in the focus groups (11 total, divided into 3 regional groups for ease of discussion) was sufficient to reach saturation of general themes (Krueger & Casey, 2015).

Students (Survey): 150 students (58% female, 42% male, approximately 18–24 years old for undergraduates) participated in the survey. Students were enrolled in studio courses taught by or associated with the faculty members in the study semester. The study obtained an approximately balanced division by three regions: Kenya (45 students), Saudi Arabia (50), USA (55). The great majority were upper-class undergraduate students studying fine arts, design, or architectural programs. The sample was one of convenience (students volunteered or were invited by means of class announcements and emails), but it represented a broad spectrum of perspectives within the studio learning environment. No personal identifying data were collected in the survey, and answers were anonymized.

Data Collection

Focus Group Procedure: Three online video conferencing focus groups (one for each country grouping) were conducted in total. Each session took approximately 60–75 minutes to complete. A semi-structured moderator's guide was used to facilitate comparability of groups as well as allow participants scope to discuss issues of interest to them. The questions used in the guide were: "What problems do you find in giving feedback to students in your studio classes?" ; "How does timeliness and detail of feedback affect student learning, in your experience?"; "How do you manage workload in terms of grading and feedback?"; and "How do you react to using technology or AI features to support feedback?" The follow-up questions requested participants to give specific examples or incidents (e.g. an occurrence where delay of feedback created an issue, or experience of using automated grading tools). All videoconferencing took place in English, that was a second language for most of the participants but one in which participants were proficient at speaking for such purposes. The videoconferencing was recorded and transcribed verbatim subsequently, by agreement of participants, from which participants were allocated pseudonyms or codes for anonymization purposes to allow for confidentiality.

Survey Instrument: Based on issues that emerged in the focus groups, the study created an online student survey (Qualtrics platform) comprising two parts: (1) Likert-scale items to measure students' agreement or disagreement with statements about feedback in studio classes, and (2) open-ended questions to facilitate further comment. The Likert items (5-point scale: Strongly Agree to Strongly Disagree) were, for example: "I wait too long to receive feedback on work," "The feedback I get on assignments is detailed and helpful" (reverse-coded to measure lack of detail), "I get frustrated at how slow it is to receive feedback," "My teachers aren't able to keep up with the workload of grading and writing comments," and "I'd like to receive feedback generated by an AI program, as an adjunct to teacher's feedback." The study designed items to measure qualitatively stated concerns (timeliness, detail, workload for instructors, openness to AI). The study also asked students to report an average number of days to receive feedback in class and to rate, on 1–10 scale, how satisfied they were with feedback, to add context. The open-ended questions requested that students describe best feedback experience they've had and worst feedback experience they've had in the program, and what changes to feedback, above everything, they'd like to be made in its provision. The two independent educational researchers outside of the study confirmed it for face validity, and it was piloted for clarity to 5 students. Anonymity to students was guaranteed, and it was sent using an anonymous link. Students participated voluntarily and were permitted to skip questions or withdraw at any time. Data was collected over 3 weeks near term end (when students would have accrued several experiences of feedback within courses).

Data Analysis

Qualitative analysis: Thematic analysis (Braun & Clarke, 2006) of focus group data was used. The study used an intensive coding process: two researchers read transcripts to familiarize themselves with data for an initial time, subsequently open-coding line by line, drawing out meaningful data that were salient to our research questions. Initial codes (such as "delay in return of work," "students not reading feedback," "language problem in feedback," "excessive number of students," "reluctance to automated aid") were created. The researchers discussed and integrated these codes into a codebook, resolving disagreement by discussion and reference to data. With agreed codebook, re-coding of transcripts was performed, and codes were subsequently compiled into higher-order themes. Four prominent themes were created (described in the Results section): Issues of Timeliness of Feedback, Detail/Quality of Feedback, Instructors' Workload, and Acceptance of AI Solutions. Divergent opinions or context-limited specific issues (such as where instructors in one country but not another) were also made to note them to observe them. To enhance trustworthiness, the study undertook member checking: participants were sent an outline of thematized findings to receive feedback or validation. Instructors generally concurred that it reflected their perspectives to an extent, with minor clarifications provided. The study also maintained an audit trail of coding decisions made.

Quantitative Analysis: The study used descriptive statistics to analyze survey data on Likert items. It calculated percent agreeing or disagreeing by statement and means and standard deviation by Likert item (coding of Strongly Agree =5 to Strongly Disagree =1). For example, the study calculated proportion who agreed (Somewhat or Strongly) that wait times for feedback were too long. Also calculated was the mean student self-reported feedback turn-around time and correlated it with satisfaction scores. With the sample (N=150) and largely descriptive intent, the study did not perform sophisticated inferential analyses but did perform simple subgroup contrasts (e.g., by region or by gender) to decide whether to look for any glaring differences. They were few - perceptions of problems were very uniform by region, so report results. The open-ended comments were content analyzed using straightforward content analysis: the study determined prevalent words or attitudes (e.g., several students echoed theme of "waiting" and "brief comments" as negatives, and commended examples of "detailed critiques" as positives). These qualitative student comments were used to illustrate the quantitative trends and to ensure the study captured any issue the Likert items might have missed. Mixed methods integration was accomplished by comparing results at interpretation time. The study pursued convergence (does student data support what instructors observed?) and complementarity (was one data set revealing something not revealed by another?). On underlying issues, strong convergence was reached and each data set added to richer insight: instructors could describe why these issues occur, and students could describe how it affects them. Taken in combination, these informed us about AI system design requirements. Findings were synthesized into user needs and functional requirements that further guided development of technical specification sheet creation.

RESULTS

Instructor Focus Groups: The instructors' focus group discussions provided a sobering though not surprising picture: studio class feedback practices get constrained by workload and time pressures at the expense of students and instructors alike. Despite geographical and institutional diversity, Kenyan, Saudi, and U.S. instructors voiced very similar concerns. We summarize the prominent topics that emerged, using representative quotes (Table 1).

Delayed Feedback and Its Impacts: All focus group participants pinpointed timeliness as an issue of significant concern. Instructors confirmed that it takes considerably more time than they'd like to get back to student feedback on project work. Often, larger project assignments or design portfolios won't get detailed comments for weeks after submission. One U.S.-based teacher added, *"By the time I get back to them, weeks having lapsed since submission. The feedback is too late to be used on next project."* Several instructors said that feedback best needs to be in process or shortly after final submission, but workload leads to backlog, which causes delay. In Kenya, instructors said that structural issues (high student-to-teacher ratios, as an example, in state universities) compound delay. One Kenyan lecturer said that sometimes, they needed to *"rush through grading 100 design plates, so some students got feedback a month later,"* admitting that it diminished value of comments. The instructors strongly felt that delay annoyed students and squashed potential for students to iterate upon that feedback. Indeed, in isolation, they referenced student complaints in research (students complaining that submitting an item past due is punished yet return of feedback can be "late" with impunity). There was consensus that delay ruins the feedback loop, and that some of them were concerned that students simply lose hope of useful feedback if it never shows up timely for work on next task.

Depth and Quality of Feedback: The instructors took pride in giving quality feedback but acknowledged that time constraints cause them to abbreviate their comments. They discussed resorting to *"bare-bones critique"* or resorting to curt written comments when workload is heavy. A Kenyan professor explained, *"I wish I could write more detailed comments, but with that number of students, my comments get brief and not very individualized."* Others concurred, saying that they resort to one-liner comments or to generic comments when rushed, though acknowledging that this is pedagogically short of best. A Saudi professor observed that in critiques of studio art, it's possible to provide rich oral feedback in class, but on written comments on assignments turned in *"I write two or three lines only because I simply can't write an entire paragraph to each of 50 students."* Instructors acknowledged, however, that brevity leaves students uncertain about how to address areas of weakness. One said that students sometimes corner her to get clarification of curt written comments that she scribbled quickly. This thread of feedback quality sacrificed to time pressures is in line with earlier documented research that students resent feedback that is too meager or unclear. Indeed, instructors bemoaned that, at times, *"the feedback is so superficial, I wonder if the student even reads it,"* feeling frustrated in their role when not able to supply detailed guidance that students need.

Grading Workload and Burnout: The most discussed topic of conversation was heavy grading and feedback in studio classes. Normally two or more classes of 20–30 students are taught by instructors, and while one quiz or one examination is graded in no time, each student's project requires attention and remark. *"Grading and returning 60 students every week is tiring. There is no way one can catch up and still be able to write quality comments to every one of them,"* was what one of two sections of an architectural studio's professor informed the research. An American lecturer painted the picture of end-of-semester rush as *"a mountain of portfolios to get through; I work all night and still feel like I'm short-changing some students."* The Kenyan lecturers further informed me that no co-instructors nor teaching assistants were normally present, and one person graded it alone. Such heavy workload directly brings about issues one and two (lateness and brevity of comments). All our interviewees reported stress, pangs of guilt, and burnout of having to struggle to keep up with feedback expectations. One of them summed it up as follows: *"We're always playing catch-up at marking. There's never time, and it gnaws at you as you know students deserve better."* All these findings confirm larger trends in higher education: higher numbers of students and fewer resources mean that lecturers feel increasing pressures. Findings from focus groups confirm that unless change is brought about, the cycle of high workload → delayed/brief feedback → student discontent is likely to repeat itself.

Openness to AI and Technological Assistance: Despite issues outlined, educators were excited when brainstorming solutions. Introduced to the idea of an AI-powered feedback assistant, most were excited, seeing it as game-changing. Few, if any, of them used such tools (apart possibly from plagiarism detectors or grammar-checkers), but many saw several ways that AI might be used. *"If an AI assistant might be able to create an opening draft of feedback or catch glaring errors, it would save me so much time. I'd still read it, but it could save so much time,"* one U.S. educator explained. Others agreed, seeing AI taking on more routine or repetitive aspects of feedback, such as finding technical errors or checking basic criteria, leaving them to work on more high-order inquiry and one-on-one mentoring. One interesting perspective was expressed by one Saudi educator, explaining that, as a non-native speaker of English, sometimes it is difficult for her to phrase feedback: *"an AI tool might help me to phrase feedback more accurately in English,"* reiterating findings that educators who struggle with language perceive AI as helping them to achieve accuracy and clarity of comments. Indeed, more than half of educators in focus groups identified that they would welcome help to word or grammar of feedback, to assist them in communicating criticism. There were, of course, caveats and concerns. Two of the instructors reacted instantly to AI with healthy suspicion: one questioned whether AI can safely be relied upon to grasp creative work, and one feared that AI-generated feedback might be generic or impersonal. But as it was discussed, most of them decided that, properly done (the AI well-trained on the task prompts, working with a predetermined fed rubric and final judgment by the instructor) an AI feedback system could be an enormous aid. An instructor from Kenya put it best: *"I'm not looking to lose my job to a machine, but if it can help me; provide me with an appropriate starting point for each pupil, I'm happy to try it."* This is in accord with recent qualitative studies in which educators perceive AI to be able to produce detailed, customized feedback in timely fashion, especially in large classes. Table 1 synthesizes these focus group ideas in representative quotes.

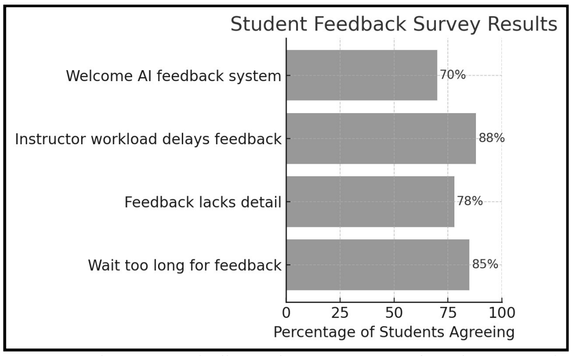
Table 1. Focus Group Themes and Sample Instructor Responses

Theme	Example Quote
Feedback Timeliness	"By the time I return their assignments, weeks have passed since submission. The feedback comes too late to be useful in the next project." — USA instructor
Feedback Detail	"I wish I could give more detailed feedback, but with so many students, my comments tend to be short and not very individualized." — Kenyan instructor
Grading Workload	"Grading and giving feedback to 60 students every week is exhausting. It's hard to keep up and still give quality feedback to each one." — Saudi instructor
Openness to AI Assistance	"If an AI assistant could generate an initial feedback draft or catch obvious issues, it would greatly reduce my workload. I'd still review it, but it could save a lot of time." — USA instructor

In summary, the instructor focus groups validated that there is an intervention needed: they wish to provide timely and detailed feedback, but logistical pressures interfere. Most importantly, they are receptive to an AI-backed solution that could enhance what they can do. All these findings informed directly of the specification of the proposed system, responding to pain points revealed.

STUDENT SURVEY RESULTS

The results of student surveys reinforced many of the instructors' findings, highlighting the learners' prioritization of issues of feedback. Figure 1 is an aggregation of salient quantitative findings of the questionnaire, expressed as percentage agreement on statements about issues of feedback by students.



Note: N=150 students. Bars indicate the percentage of students agreeing with each statement about their current feedback experience.

Figure 1. Student Feedback Survey Results.

As is evident in Figure 1, a total 85% of students concurred (48% of them "Somewhat Agreeing," while 37% of them "Strongly Agreeing") that *"I often have to wait too long for feedback on my work."* This verifies slow feedback as an ongoing student issue. On average, students took 10–14 days to receive feedback on main studio assignments, while exceptional examples took an entire month to receive final project feedback. Notably, most programs only allow two weeks for feedback, yet our results suggest that this standard is frequently not being upheld, and that students feel this delay. In open-ended statements, one student observed: *"By the time that I received comments back on my design, I'd completed two other assignments. Not that useful then,"* which well illustrates lost value of slow feedback. This student feeling is supported by earlier research indicating that timely feedback is necessary to allow students to act upon feedback in subsequent assignments. In terms of feedback quality and detail, slightly more than 22% of students concurred that *"The feedback I receive is detailed and sufficient."* On the other hand, nearly 78% indicated that assignments receive insufficient comments in terms of detail (operationalized as strong disagreement using the positive-worded stem, or merely agreement using one of the negative-worded items). Students experience comments as superficial: in terms of comments, that is, one of the survey items, *"My instructor's comments on assignments are thorough and helpful,"* received an average rating of just 2.3, where agreement is on a 5-point scale (3 =neutral). Qualitative comments corroborate most of what students said about comments received used words like "generic," "brief," or "vague." One student from Saudi Arabia, for example, exclaimed, *"Too often, the remark is simply 'good work' or 'needs more detail' and no explanation is provided. I find myself thinking, what needs more detail?"* One U.S. student indicated, *"I get lots of 'OK' or check marks but not much written comments, so I don't know what to correct."* Such comments confirm analogous difficulties that befall instructors and is what's prophesied in writing by Literature which assures that students become frustrated when feedback is imprecise or superficial. Lack of detailed action items is, it appears, keeping students from being able to learn from mistakes.

The issue of instructor workload filtering down to students' experience was also there. Around 88% of students agreed that *"I think my instructors have too many students/works, which slows down getting feedback."* Students might not be fully aware of just how heavy instructors' workload is, but several perceived that instructors were busy or that they were spread too thin. A few comments showed understanding: e.g., "*I know my professor has like 3 classes, so I know why it takes long, but it's frustrating as a student,"* one student speculated. A second student said, *"It just seems like the instructors get overloaded, and that's why we don't get as much feedback."* Student judgments of instructors' workload mirror instructors' reports that workload influences delay in feedback. They also suggest that students might be relatively compassionate if, in their mind, delays are structural, not professor sponsored. In any case, to the student's eye, it still means slow or no feedback, which slows them down.

Of most interest in the aims of our research was student reaction to AI-generated feedback's forward-looking statement. Students were asked if they would accept AI program-provided feedback. About 70% were favorably disposed (and 30% of which were "Somewhat Agree" while 40% were "Strongly Agree") while 15% were not, and 15% opposed it. This shows that most students accept, or at least don't reject AI-provided feedback, especially if it isn't framed as in lieu of, but as an adjunct to, professor feedback. Several students clarified that any prompt feedback is preferable to nothing, otherwise waiting for professor feedback. One student explained, *"Even if a computer provided me with tips upon submitting, that'd be great, at least I'd know something to tinker about while waiting on my professor,"* in relishing AI feedback as preliminary or supplementary. Several students provided caveats, wishing to be assured that AI feedback is correct or filtered and based on a human imposed rubric: *"I'd use AI feedback if I could be certain that my professor verified it or it's been proven to be trustworthy. I don't want to be steered down that path,"* one student indicated. This is in line with instructors' caution and reminds that an AI system needs to be directed by a human and framed as an aid, not as final word or absolute judgment.

Other quantitative findings include mean student satisfaction rating of feedback (on 10-point scale) of 5.8, where there is clear negative correlation between wait time (reported as usual) and student satisfaction ratings ($r \approx -0.45$, moderate correlation, where larger wait times correlate with lower satisfaction ratings). And 92% of students agreed that *"Getting faster and more detailed feedback would improve my learning in this course."* Such widespread agreement is supportive of pedagogic case for faster feedback systems; students feel (theory predicts) that faster, richer feedback enables them to revise and perform better. This is as predicted by educational principles that timely, specific formative feedback enhances learning and performance.

To emphasize human contribution, the study notes one contrast of open-ended findings: Students described instances where an educator gave individual meeting or in-studio real-time feedback, and received it as motivating and clear, when asked to report best feedback experience. Students most frequently answered *"no feedback at all"* or *"feedback that came too late to be of any use"* when queried about worst feedback experience. One student simply stated: *"Worst: submitted a project, never heard anything about it, just received a grade. still don't know what I did or did not do wrong*"; an unhelpful outcome an AI system could avoid by ensuring that formative feedback is always provided. In summary, three main issues were confirmed: feedback is typically delayed, feedback is typically not specific, and these issues relate to instructors being workload-overwhelmed, which students feel. There is also cautious positive reception of AI-powered feedback, provided it is positioned as supplementing the role of the instructor. Such findings provide a mandate to an intervention design: students clearly desire faster, better feedback, and are receptive to novel means of achieving it, setting the stage for an AI-powered solution.

DISCUSSION

The goal of this project was to outline in a methodical way what such a system would need to be of value to studio-based learning. The findings of our needs analysis shed light on the pedagogic deficit and lead to a vision for an AI feedback instrument to fill it. This study discusses implications of the findings, contextualize them within the literature, and set out specifications and roadmap for the imagined AI system. It also mentions constraints and detailing next steps, such as pilot testing and upscaling.

Addressing the Pedagogical Gap in Studio-Based Feedback: The needs assessment confirms that there is widespread pedagogical deficit: students enrolled in studio courses aren't being provided with timely and specific feedback, and instructors aren't able to manage the workload. This is not a new deficiency; it's reflecting widespread anxieties within higher education systems. An issue of student discontent (Williams & Kane, 2012) and, to be frank, within mass-education systems, has been feedback for years. The value of this study lies in putting that issue within an SBL framework in different cultural contexts and using that directly to inform technical solutioning. Consistency of outcomes among student and teacher cohorts, as well as countries, is noteworthy. It is evident that delay and brevity of feedback issues are widespread problems to studio pedagogy, independent of institutions. Such prevalence is congruent with research that as class sizes increase and resources become fixed, even conscientious instructors will struggle to maintain feedback quality (Robinson *et al*., 2013; Carless & Winstone, 2020). Moreover, teacher participants were also open to using novel tools if it supports them to resolve these issues; an encouraging finding as teacher uptake is one of the primary reasons that educational technologies do not gain traction (Howard *et al*., 2021). That teachers brought up AI as an idea and reacted generally favorably is an encouraging sign towards uptake in the future.

Theoretically, formative and timely feedback is crucial to finishing loops of experiential learning cycle. In studio learning, every project is an opportunity for students to enact theory and approaches, be critiqued, and continue developing work. When that process is decelerated, however, that loop is lost and opportunity for learning is lost. The study of students confirms that is what is taking place on a day-by-day, student-by-student level, which most likely means students aren't getting full value from studio method. And superficial feedback (little or no remark, just a grade) does not engage students in reflection, one of experiential learning cycle's fundamental processes, as theory suggests by Kolb. Briefly, best practice for studio class formative assessment isn't what's currently carried out in most studio classes. For instance, guideline of the UK Quality Code that feedback be provided "in sufficient time to enable students to build on what has been learned and to improve their performance in subsequent assessment" isn't being provided in most students' books. The needs analysis also revealed an intriguing subtlety: instructor workload and burnout. This means that any solution must be not just effective for students, but also practicable and bearable for instructors. Teachers are gatekeepers of an innovation's adoption or rejection. If using the solution increases workload or complication, it won't be adopted in the long run. One of our design specifications for an AI feedback system is, therefore, to reduce instructors' net workload, not add to it. The AI system designed needs to assume time-consuming labor (like an inaugural round of comments or identification of errors most frequently made on studio projects) releasing instructors' time and effort, thereby. Well-designed, AI may even add to quality, as well as efficiency, of feedback, by affording instructors breathing time to focus on more individualized mentoring and higher-order feedback that AI is still not capable of (like higher aesthetic judgment or deep conceptual guidance). Such complementarity is in harmony with AI as an "assistant" or co-pilot, not replacement (Holmes *et al*., 2019). The instructors' receptiveness to using AI is contingent upon their keeping final control of feedback through imposed rubrics; as one of them put it, they wish to *"be in control of final grading."* A human-in-the-loop design is therefore indicated. The AI system needs to create draft feedback or suggestions as instructed, which can be viewed, edited, or accepted by instructors simply. Such design exploits AI's efficiency while ensuring quality and trust by means of human presence.

PROPOSED AI FEEDBACK SYSTEM: TECHNICAL SPECIFICATIONS AND RATIONALE

Based on the identified needs, the research developed a comprehensive technical specification for an AI Feedback System tailored to studio-based learning environments (see Appendix for the full specification sheet). ChatGPT was utilized in this process due to its advanced natural language processing capabilities, which enabled the articulation, refinement, and organization of complex system requirements in a clear and structured format. The system's key features are outlined below; each aligned with the specific challenges and needs highlighted during the needs assessment phase.

- **Rapid, Automated Feedback Generation:** The technology will use natural language processing (NLP) and machine learning capabilities to offer formative feedback on student work almost in real time after submission. Upon submission of text (e.g., design reports, conception statements), a large language model (e.g., a fine-tuned or prompted GPT-4) can create condensed feedback, highlight strengths, and make recommendations. Upon submission of a visual or design artifact (as in the case with SBL), the technology can use computer vision to review images or CAD files. The general aim here is to have a student get at least partial feedback within minutes to a few hours, rather than days. This addresses directly the issue of timeliness, giving a student actionable feedback while they are still thinking about it.
- **Feedback Aligned with Criteria and Personalized:** To address generic feedback, the AI will be guided by project-specific rubrics or criteria supplied by instructors. Instructors input or select major criteria for each project at setup (e.g., in an architecture studio: originality of idea, accuracy of technical drawing, aesthetic value, etc.). The AI then structures its comments through these criteria to ensure it addresses applicable aspects of work. In addition, the system can store each student's prior feedback to allow for new feedback personalization (e.g., marking progress made, referring to prior suggestions, akin to a "feed-forward" facility). The blending of rubric-alignment with personalization serves to provide rich, applicable feedback without creating one-size-fits-all comments students complained about. Education literature confirms personalized, criterion-referenced feedback to be most effective.
- **Instructor Oversight and Editing Interface:** Instructors are given a dashboard where they see AI feedback before it gets out to students (especially major summative feedback). The interface would display the AI-recommended feedback to each student in a skimmable, readable format. Instructors can then accept, revise, or supplement this feedback. This approach honors

instructors' expertise and current capabilities of AI – the AI can provide straightforward feedback while instructors can polish nuance points. One key factor is, if an instructor really doesn't have time, they can post AI feedback with minimal form tweaking for formative work, hoping it's better than nothing. On the other hand, for a major project, instructors can spend time crafting each note, drawing upon the AI response as a template. This makes the tool flexible for varying approaches to instruction, as well as time constraints. It also serves a secondary purpose: reviewing AI suggestions makes instructors think about both feedback habits and consistency, potentially benefiting feedback literacy in both ways (Carless & Winstone, 2020).

- **AI Feedback Style and Tone Customization:** There was a fear that the feedback would sound robotic or formulaic, or not in the instructor's voice. In response to this, the system allows instructors to define a preferred tone (e.g., warm/encouraging, critical/formal) and to contribute example feedback statements themselves. The AI model can use these to reflect the tone of the instructor in generated feedback. Essentially, the instructor "trains" AI using a few examples - e.g., how he/she typically gives praise vs. criticism. This function leverages transfer learning capabilities of modern NLP and makes the AI sound like a natural continuation of the instructor, not this foreign voice. This level of customization is required to get buy-in from instructors, and acceptance by students; students shouldn't be easily able to tell which feedback sentences are from the instructor, versus from the AI (except where transparency necessitates tagging, which we'll discuss later, under ethics).
- **Integration with LMS and Workflow:** Technical requirements underscore how critical it is for the AI engine to be implemented within the Learning Management System (LMS) or whatever submission portal a class utilizes (D2L Brightspace, Blackboard, Canvas, Moodle, etc.). The integration would happen through a plugin using LTI (Learning Tools Interoperability) or API. The reason for this is to not disrupt current workflows; instructors and students shouldn't have to log in to a second system, nor navigate a second submission. In this case, when a student submits an assignment in the LMS, that submission would automatically get funneled to the AI feedback engine. The engine generates feedback then auto fills a draft feedback box a teacher sees. The teacher can then finish and publish to the student within the LMS. Students would see feedback where they'd otherwise see feedback, with a note somewhere that part of it was machine-generated (for disclosure purposes). Through integration into current LMS spaces, adoption hurdles are fewer and so workflow is entirely smooth. This also aids scalability: integration means a tool can be shared across departmental levels and courses.
- **Multi-language Support and Clarity Checking:** Since we are operating with an international environment here (i.e., Kenyan and Saudi instructors giving feedback in a potential non-native language for them, and student readership alike), language assistance within the AI system is made available. The AI can serve to ensure feedback comments are written clearly enough in the chosen language (adequate English in our study environment) and even translated where appropriate. Less skilled writers among instructors can also rely upon the AI in suggesting grammar corrections or word clarifications. This provision was itself advocated by instructor feedback within the focus group concerning language impediment, and by literature referencing how useful a function AI can be in editing back to a clarity level. The AI can even review for constructive wording – wording feedback in a constructive, supportive form (which, according to a study, optimizes reception by students (Ferguson, 2011)). The AI can essentially be used as an "editor assistant" to produce feedback written clearly enough and worded politely enough, avoiding "cryptic or overly critical" feedback commonly plaguing overburdened instructors (Hounsell, 2007).
- **Immediate Formative Feedback to Students:** For low-stakes assignments or iterative drafts, the system can bypass instructor review and deliver instantaneous feedback to students. Consider a two-week sketchbook assignment or coding exercise in a studio course; once submitted, a student might get back AI feedback within minutes to work with before next class. This would be up to the instructor (enabled for those where timely response matters more than perfect accuracy). The benefit is having students continually always making progress, sort of like a virtual TA. Our student survey showed a preference for faster cycles; this addresses this directly. The philosophy fits under "feedback on demand" and use of automated means like quizzes or intelligent tutors to complement human review. Of course, for a final project or large grades, instructor moderation would remain.
- **Technical Architecture:** The specification (Appendix) stipulates a cloud infrastructure whereby the AI engine (maybe a hybrid rubric scoring algorithm and a generative model for text) would be remotely executed on secured hosts. Due to potentially intensive processing of a potentially large quantity of submission records, scalability would be a factor. The system would leverage available platforms wherever possible (e.g., API into a large language model for text, and another model for any image processing if included in the scheme). Data security procedures would be preserved through student submission records, along with returning feedback data, must be encrypted enroute, in keeping with FERPA and applicable data security legislation. No student data would ever be made available for further development of AI models without permission, to preserve privacy vulnerabilities. Essentially, each institution would have a choice to house the model behind its firewall or use a trusted third-party cloud with sound privacy agreements.
- **Ethical and Transparency Considerations:** The study acknowledges that use of AI in feedback raises ethical questions. Our design includes openness to users: students are to be notified when a feedback comment has been generated by AI (adding a small note "This comment was generated with the assistance of an AI tool and reviewed by your instructor"). This promotes credibility, while allowing critical thinking by students about where feedback comes from. In addition, the system necessitates a means by which students can rate or warn back against received feedback. If a student feels an AI comment is out-of-bound, not helpful, they (and the instructor) can provide feedback about it back to the system, to be fed back over time in order to improve the model (a form of supervised feedback loop for the AI). In addition, the AI will be trained upon a pedagogy-appropriate dataset (presumably by drawing upon prior

examples of quality feedback by instructors, supplemented by best-practice comments from literature) to minimize risk from bad, inappropriate, or biased feedback. The system can't make value judgments, nor personal remarks, but instead remains objective in its criticism. By incorporating such ethical limits, the study aims to protect this beneficial educational relationship where AI is a member of the teacher-student dialogue, not a black box authority.

- To summarize, then, the proposed AI Feedback System can be thought of as a very useful aid instrument that directly addresses perceived problems: it greatly decreases feedback turnaround time, makes more feedback accessible (every student gets more comment, since AI responds without fatigue), and does so in a structured form, that decreases instructors' work-load once setup. If implemented, it can be expected to lead to improved student learning outcomes and satisfaction. Students can get timely response for "How did I do?" questions as well as "How can I improve?" questions, while instructors can teach large studio courses with decreased burnout, channeling energies where best applied.

Development Partnership and Cost Considerations: Implementation of said foregoing system is a non-trivial piece of software work. Given our expertise in pedagogy and design, not in programming for AI, our inquiry investigated the potential co-venture with a commercial software development business in developing the system according to our specifications. It identified Codup.co, a software development business providing digital solutions, as a potential partner in this venture. Codup.co develops customized software and even performs AI-related work from its repertoire (Codup, n.d.). Of even more important, however, is how broad a spread there is in project budgets within it, showing they can scale up to a project this large; by industry reports, Codup project commitments have ranged from a low of $1,500 for small features to over $300,000 for large complex systems. This shows flexibility and capability to perform both development of an MVP (minimum viable product), and future expansion.

The study consulted Codup.co to get a projected figure for developing our AI feedback engine's MVP. The MVP would include the core features: LMS integration, AI-driven feedback generation engine (likely by repurposing a recognized NLP model through API), instructor portal, and minimal security/analytics. Based upon initial talks and benchmarks against comparable projects for comparable pieces of software with comparable specifications, the study estimates initial development costs at about $50,000–$75,000. This aligns with typical costs for custom AI software in 2025: industry reports have minimal AI education projects starting at $10k for minimal proof-of-concept prototypes to $100k plus, for full-featured ones. The study estimates $50-75k for code, design, testing, deployment, etc., spread out over development time of about 6 months. The cost would not be fully funded by our current $10k internal funding (which was primarily for exploratory purposes), so outside funding or institutional backing will be required to proceed to development. But by having the complete spec in our hand (prime deliverable for this project), this study is in a great position to approach further grants or investors, since it can state exactly what it needs to create and why.It is worth noting that using pre-trained AI models (via API) can be cost-effective for development but introduces ongoing usage costs (e.g., if using a third-party AI service that charges per request). The study's plan includes evaluating the trade-offs between developing an in-house model vs. using an API like OpenAI's. The latter might reduce up-front development time and cost, but would need a budget for operational costs. For a pilot scale (a few courses), these costs are modest (perhaps a few hundred dollars in cloud compute fees). If scaled to an entire university, it could be in the low thousands annually, still reasonable in comparison to human resource costs. In summary, Codup.co would introduce professional development practice in constructing the system in an effective and trustworthy manner. The study has considered estimated costs and finds them to be acceptable in relation to the potential impact. The detailed technical spec (Appendix) can be utilized as the blueprint for Codup engineers to close the gap between requirements of teaching and technical realization. This project is an exemple of university-industry collaboration for teaching and learning innovation.

FRAMEWORK FOR PILOT TESTING, REFINEMENT, AND SCALING

With the needs assessment in hand and requirements defined, the way is clear to move forward on a planned implementation path. The study suggests a staged process beginning with small-scale pilots and adding additional phases incrementally, getting feedback at each step, i.e., an agile, iterative development process coupled with formal evaluation research. This strategy is guided by best practices in ed-tech innovation, including the Digital Promise Ed-Tech Pilot Framework, which recommends articulating needs, piloting, gathering data, and iterating prior to scaling.

Phase 1: Prototype Development and Internal Testing. In this stage, the development team (i.e., Codup.co working together with the research team) will create the MVP of the AI feedback system to specification. Once a working prototype is ready, we'll conduct some internal testing with a small number of test-users (perhaps some of the instructors from our focus groups and their teaching assistants). The goals are to identify any technical bugs, assess the quality of AI-generated feedback (does it make sense? is it on-topic?), and fine-tune parameters like the level of detail in comments. A simulation can also be conducted: input the system with some previous student work and compare the AI comments with the original instructor comments to determine how well it aligns or differs. This step guarantees that prior to including actual students, the system is secure and proficient enough. The system will be refined as needed (e.g., adjust the AI model if it's misinterpreting certain kinds of assignments). (e.g., tune the AI model if it's misunderstanding types of assignments).

Phase 2: Pilot Implementation in a Real Course. Next, we plan a pilot study in an actual instructional setting. Ideally, this would be a single course or a couple of course sections (around 20–40 students) taught by one of the instructors on our team. The choice of course will be strategic: a studio-based course where feedback is integral, and where the instructor is enthusiastic about trying the tool (like one of our focus group participants). Before the course begins, we will train the instructor on using the AI system and setting up assignments in it ("Train & Implement" step per Ed-Tech Pilot guidelines). During the course, the AI system will be used to generate feedback for some assignments. We will adopt an experimental mindset: for example, use the AI feedback on some projects while perhaps another project is done with traditional feedback, to compare outcomes. Data collection in the pilot will be comprehensive: student feedback perceptions via surveys (compared to baseline data from our current study), instructor workload logs (to see if time spent on feedback changed), and learning outcome measures (did the students who received AI-augmented feedback show equal or better project improvement or grades?). We'll also have qualitative check-ins: focus groups or interviews with the students and instructor after the pilot to gather their experiences. Since this is a critical proof-of-concept stage, careful analysis of pilot data will guide the next steps. If students report confusion or distrust of AI feedback, that's a red flag to address. If instructors find the interface cumbersome, we refine it.

Phase 3: Iteration and Refinement. Pilot results will guide us in adjusting both the AI system and the implementation plan. Perhaps the pilot indicates that the AI's language needs to be tweaked, or students could have gained from an orientation session to maximize the benefit of AI feedback. We will adjust accordingly. This phase might involve another development sprint with Codup.co to implement improvements. We foresee possibly having to tune the AI model with additional training data from the pilot (supervised learning off instructor corrections done during the pilot, so that the model can learn). The iterative nature of this aligns with agile development principles and forms a design-based research cycle in which interventions are iteratively refined in context. We might pilot a second time if we did make some major revisions – perhaps at a different university or with a different teacher to check for generalizability.

Phase 4: Scaling Up. Once the system proves its efficacy in pilot settings (e.g., improved student satisfaction without loss of academic quality, and confirmed reduction in instructor workload for feedback), the stage is set for scaling. Scaling can occur in multiple dimensions: more courses within the department, adoption by other departments (e.g., other disciplines that have project-based assignments), or even institutional-wide rollout if supported by administration. For scaling, it's not just the tool that matters but also policy and support. We would work with the university's Center for Teaching and Learning (CTL) to create documentation and training workshops for any instructor interested in using the AI system. We'd also ensure technical capacity for more users (infrastructure scale-up, possibly transitioning to a more robust server setup). During scaling, continuous monitoring is crucial – collecting usage data, performance data, and feedback from new users to ensure the system remains beneficial in diverse settings. The Ed-Tech Pilot Framework emphasizes summarizing and sharing results; in our case, we would publish our pilot findings in academic venues and share case studies with our institution to build buy-in. Throughout these phases, one must keep an eye on effectiveness and unintended consequences. For instance, does reliance on AI feedback change how students approach their work? (We would want it to make them more proactive, not complacent.) Does it affect the student-instructor relationship? (Ideally it frees time for more mentorship, but we should guard against any reduction in personal interaction.) These are research questions we'd explore in the pilot and beyond. We also need to ensure that using AI doesn't disadvantage any group of students. If any biases are detected (say the AI favors a particular writing style or misinterprets non-Western examples), we must address them through model adjustments or additional training data. Being vigilant during the scaling phase will help maintain equity and inclusion. This framework aligns with general recommendations for rolling out educational technology innovations in a controlled, learning-oriented way. By starting small, rigorously evaluating, and gradually expanding, we maximize the chances of success and minimize risks. Moreover, this phased approach allows stakeholders (students, instructors, administrators) to build trust in the system. Early successes can be champions for later adoption.

Limitations and Future Work: While this project was well conceived, there are a few limitations. First, data (especially focus group and survey responses) are self-reported attitudes. They tell us about feelings and experience, important ones to be sure, but indirectly about learning achievement itself. It is understood from educational theory timely, detailed feedback makes learning more effective, but this project did not measure change in student performance. Future studies seek to determine if use of the AI system truly produces more quality student work or grades (e.g., utilizing rubrics or external tests). Second, the samples, while geographically heterogeneous, remain small. Only 11 instructors and 150 students took part, and they may not be generalizable to all studio programs or to any other disciplines. Other disciplines may have unique requirements for feedback (e.g., a music performance studio vs. a design studio). This design thus may need to be adjusted outside the studio environment. The study conservatively attempted to focus on generally applicable qualities, yet care must be taken in making one-size-fits-all leaps. Pilot studies in alternative contexts (e.g., in STEM lab contexts, or writing studios) would be useful to test and refine the generality of the system.

Third, the study did not investigate potential negative consequences of AI feedback during needs assessment. Participants had optimistic views. Perhaps we have not yet been exposed to problems, e.g., student overreliance on AI or plagiarism problems (would feedback from AI be "assistant" in breaking rules? Most unlikely because it would equate to having a tutor, although we need to stress this issue clearly enough). Going forward, there will be need to have discussions on policy for appropriate use. Another limiting factor to each component of the project being totally realized is the current technical specification not yet being tested. There might be technological barriers in realizing all functionality. For example, flawless translation from visual design produced by AI remains a developing art. Implementation of reality might require scope adjustments. Finally, concerning funding: this study was funded by an in-house grant for this phase. Still more funds will be needed to develop and deploy the system, as described. There is ever-present uncertainty that even with a solid plan, securing additional monies or administrative approval might be challenging (budgets, changing priorities, etc.). This study has minimized this uncertainty by creating tangible deliverables (data, spec) making a case for additional investment more persuasive.

Future Work: The immediate next step is obviously to build and pilot the system, as described. From a scholarly perspective, this presents an opportunity to study human-AI collaboration in education in depth. The plan is to document the pilot implementation in a subsequent paper, evaluating not just efficacy but also the *process* of integrating AI into teaching practice. It will be enlightening to see how instructors' attitudes evolve as they use the system (does it reduce stress? Do they trust it over time?). Likewise, how do students adapt? There will be need to study if student feedback literacy improves when they get more frequent feedback; there's literature suggesting that more iterative feedback can help students learn to use feedback better (Winstone *et al*., 2017). Another area for future exploration is expanding the system's capabilities: for example, can we incorporate *peer feedback* alongside AI feedback? Studio classes often involve peer critique; perhaps the platform could host peer comments and use AI to summarize or highlight common points, creating a rich multi-source feedback environment. Also, this study envisions possibly applying this to other domains: a version for writing-intensive courses or lab report feedback, etc. Each domain might require training the AI on different data, but the principle remains: timely, specific feedback is universally beneficial. This study is also interested in the longer-term impacts on learning outcomes. Over multiple semesters, does the availability of AI feedback correlate with improved project quality or higher order thinking in student work? We might conduct quasi-experimental studies where some sections use the AI system and others don't and compare student outputs via blind assessment. This would provide evidence for or against the value-added of the system.

In terms of technical future work, an intriguing direction is to incorporate learning analytics e.g., the system could track common issues across many students and provide that analysis to the instructor or the class ("Here are three areas where the class as a whole struggled, according to the AI's analysis"). This could inform targeted mini-lectures or resources. It essentially closes the loop at the class level, not just individual level. Doing so responsibly (ensuring the analysis is correct) would be an innovation in how AI might not only give feedback to students but also feedback to instructors about their teaching or assignment design.

CONCLUSION

In conclusion, this project revealed a key shortcoming in learning in a studio environment and charted how this can be addressed through technology. The study found that both tutors and students are disillusioned by lateness and brevity in feedback within project-based artistic learning environment, with potentially hindered iterative learning and frustration. In response, the study proposed a feedback system driven by artificial intelligence and crucially formulated a full technical requirement rather than leaping into development. This prudence ensured that the solution imagined rests firmly upon actual user requirements and pedagogic goals.

Our key contributions are:

- A detailed description of SBL feedback requirements from multiple perspectives, and
- A full design specification for a useable AI feedback system as a roadmap for implementation.

The technical contribution here lies in turning technical specifications into a scholarly artefact informed by research, something others can develop from. This study did not deploy nor evaluate the AI system in this work; its arguments are thus for envisaged benefits and deductive implications. The actual learning and teaching impact would need to be evaluated in future deployments. Nonetheless, the vision for the system remains exciting. If implemented, a feedback system described above would have the capability to revolutionize studio learning by making timely formative feedback for each student feasible even where large classes are taught. It would allow teachers to recover more control over their load to spend more time directly with students and in creative feedback, with a guarantee that automatic routine feedback would follow. Broadly, this fits into the trend for capitalizing on more scalability and consistency in pedagogy by enhancing people with AI (Luckin *et al*., 2016). It shows how one can use AI to enable a pedagogy (studio learning) known to be high impact but labor-intensive and thus provide high-quality feedback to more students without corresponding teacher effort. This work is also made more robust by being cross-cultural. Having had both teachers and pupils from a variety of countries makes it more certain that both the elicited requirements and solutions are applicable everywhere. The underlying problems with timely and effective feedback are not unique to any educational institution; thus, a solution by a machine can largely be applicable. Of course, localization (e.g., language, compatibility with local LMS) would be necessary, but the core notion would remain the same.

Finally, this study acknowledges here the mechanisms for support allowing for this work. The Stephen F. Austin State University JACK STARS internal funding provided critical funding to proceed with the needs assessment. This sort of seed funding is invaluable in getting novel ideas to a point where more substantial major funding can be applied. The foundation is here now, and with ongoing funding, this study proposes bringing the AI feedback system from ideas to reality. The journey from a realized pedagogical pain point to actual practice is long and requires cooperation from across disciplines. This project reflects this cooperation, bringing together expertise in education, computer science, and user-centric design. Without any reservation, this work engenders like-minded interdisciplinary projects toward improving learning and teaching with cutting-edge technology, ever in answer to real student and teacher needs.

ACKNOWLEDGMENTS

The author wishes to thank the instructors and students who participated in this study for their time and insights. This project was supported by the Stephen F. Austin State University JACK STARS internal grant (Award of $10,000), which provided funding this research. The proposed AI system concept has benefited from discussions with the development team at Codup.co, whose input on technical feasibility and cost estimations are gratefully acknowledged. Any opinions, findings, and conclusions expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agency or collaborators.

REFERENCES

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101.

Carless, D., & Winstone, N. (2020). Teacher feedback literacy and its interplay with student feedback literacy. *Teaching in Higher Education, 28*(1), 150–163.

Clutch. (2023). Codup (company profile) – *Reviews and Ratings*. Retrieved from Clutch.co: *"Codup's pricing varies significantly, with projects ranging from $1,500 to over $300,000…"*clutch.co

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and Conducting Mixed Methods Research* (3rd ed.). Thousand Oaks, CA: SAGE.

Digital Promise. (2016). *Ed-Tech Pilot Framework* (Eight-Step Guide to Edtech Pilots). DigitalPromise.org. digitalpromise.orgedscoop.com

Ferguson, P. (2011). Student perceptions of quality feedback in teacher education. *Assessment & Evaluation in Higher Education, 36*(1), 51–62.

Hanmore, T., Moon, C. C., Curtis, R., Hopman, W., & Baxter, S. (2023). Is time really of the essence? Timeliness of narrative feedback in ophthalmology CBME assessments. *Medical Teacher, 45*(5), 480–486. https://doi.org/10.1080/0142159X.2023.2274286

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112.

Hounsell, D. (2007). Towards more sustainable feedback to students. In D. Boud & N. Falchikov (Eds.), *Rethinking Assessment in Higher Education* (pp. 101–113). London: Routledge.

Lynam, A., & Cachia, M. (2018). Students' perceptions of feedback in higher education. *Assessment & Evaluation in Higher Education, 43*(8), 1302–1313.

used to assist in drafting and organizing the technical specifications of the proposed AI Feedback System, based on researcher-provided inputs and contextual data.)

Pang, T. Y., Kootsookos, A., & Cheng, C.-T. (2024). Artificial Intelligence Use in Feedback: A Qualitative Analysis. *Journal of University Teaching and Learning Practice, 21*(6), 1–20. *(Published April 30, 2024).*open-publishing.orgopen-publishing.org

Quality Assurance Agency for Higher Education (QAA). (2018). *UK Quality Code for Higher Education: Advice and Guidance – Assessment*. (Guideline quote: "Feedback comments are provided in sufficient time to enable students to enhance their performance in subsequent assessment tasks.")files.eric.ed.gov

Williams, J., & Kane, D. (2012). Engineering programmes in the UK: the student feedback experience. In *Enhancing Learning and Teaching in Engineering* (HEA Conference Proceedings). *(Quote on student comments about late feedback)*files.eric.ed.gov

APPENDIX:

Proposed AI Feedback System Technical Specifications

This appendix presents the detailed technical specifications and requirements for the AI Feedback System as derived from the needs assessment. It is intended to guide software development and serve as a reference for stakeholders involved in implementation.

SYSTEM OVERVIEW AND OBJECTIVES

- **System Name:** (Tentative) *Studio Feedback Assistant (SFA)* – An AI-driven feedback support tool for studio-based learning.
- **Primary Goal:** To provide timely, detailed, and personalized feedback to students on creative assignments, while reducing instructor workload associated with grading and feedback.
- **Scope:** Initially designed for text-based and image-based assignments in design studios (e.g., architecture, art, engineering design). Future extensions could apply to other domains.

Users

- ***Instructors –*** set up assignments, review/edit AI feedback, oversee the process.
- ***Students –*** receive AI-generated feedback through existing course platforms and respond to it.
- *Administrators/Researchers* (secondary) – monitor usage, aggregate data for program improvement.

FUNCTIONAL REQUIREMENTS

Automated Feedback Generation: For each student submission, generate formative feedback addressing key aspects of the work. The feedback should include:

- At least 2–3 positive comments (what was done well).
- At least 2–3 constructive suggestions for improvement.
- Reference to assignment criteria or rubric categories.
- Approximate length: 100–300 words total (adjustable by instructor preference).
- Accuracy: Feedback content must be relevant to the submission (e.g., no generic text unrelated to the student's work).

Op

Rubric Alignment: Allow instructors to input or select criteria/rubric for each assignment. The AI feedback should explicitly tie comments to these criteria (e.g., "On *Concept*: …", "Regarding *Technical Execution*: …").

Personalization: Incorporate student-specific context where available:

- Recall previous feedback given to the same student (if any) to avoid repetition and acknowledge progress ("This is an improvement from your last draft in terms of composition…").
- Use the student's name in feedback for a personal touch (optional, based on instructor setting).

Language and Tone Customization

- Provide settings for feedback tone (e.g., formal, conversational, encouraging).
- Allow instructor to provide example feedback sentences or style guidelines. The system adapts the output to match this style.
- Ensure language is clear and at an appropriate reading level (avoid overly technical jargon unless in discipline context; avoid slang unless instructor uses it).

Immediate Low-Stakes Feedback Mode: For assignments flagged as "practice" or drafts, automatically release AI feedback to students upon generation (no instructor approval needed), to maximize immediacy.

Instructor Review Mode: For higher-stakes or final assignments, route AI-generated feedback to an instructor dashboard for review/editing *before* student release.

- Instructors can edit the text inline or approve as-is.
- Instructors can add additional comments manually.
- The interface should make it easy to see AI suggestions and modify them quickly (e.g., suggestions highlighted).

Feedback Release Controls

- Instructors can publish feedback to all or individual students once they are ready.
- Option to "publish with delay" (e.g., generate immediately but release to students at a specified time, if instructor prefers to coordinate feedback release).
- Option to revert to manual-only feedback for certain students or assignments (in case AI output is unsatisfactory, instructor can choose to write from scratch).

Peer Feedback Integration (future feature): (Not in MVP but planned) Enable input of peer comments so AI can integrate or summarize peer and AI feedback in one coherent message to the student.

TECHNICAL SPECIFICATIONS

Architecture

Client Side: Web-based interfaces:

- *Instructor Dashboard*: Accessible via web browser (and possibly within LMS as an embedded frame). Shows class roster and submissions, AI draft feedback, editing tools.
- *Student View*: Students see feedback in LMS assignment feedback area or a web portal, with AI comments labeled as needed.

Server Side: Cloud-hosted application with modular components:

- ***AI Engine*:** NLP Model (e.g., GPT-based or similar) hosted on a server or via API. This component handles text generation for feedback. Possibly separate sub-modules for image analysis if needed.
- ***Application Logic*:** Orchestrates data flow (receives submission, calls AI engine with proper prompts, returns result, saves database).
- ***Database*:** Stores submissions metadata, feedback generated, user profiles (could be integrated with LMS DB via API or a separate DB linked by student IDs).
- ***Integration Layer*:** APIs or LTI for communication with LMS (for receiving submission events, sending back grades/feedback).

Scaling & Performance: The system should handle simultaneous submissions (e.g., if a deadline causes dozens of submissions at once). Aim for generation time per submission < 2 minutes under moderate load (50 concurrent). Use asynchronous job queues for AI processing to manage high load. The architecture should be horizontally scalable (e.g., containerized services that can be replicated).
Technologies: TBD by development team, but likely: a web framework (e.g., React for front-end, Node/Python for backend), an AI platform (OpenAI API, HuggingFace models, or custom PyTorch/TensorFlow model), and a SQL/NoSQL database. Ensure chosen tech stack supports required integrations.

AI Model Details

- Initially leverage a pre-trained Large Language Model (LLM) for text generation (for example, GPT-4 or an open-source equivalent). Use prompt engineering to guide the model: prompt will include the assignment criteria, possibly an example of ideal feedback format, and the student's submission (or key details extracted from it).
- For image-based assignments, use an approach where the student provides a short description or the instructor provides expected features to check. Alternatively, integrate with any existing image-to-text critique models (still experimental; likely not MVP).
- Implement a content filter on AI output: ensure no inappropriate or offensive content in generated feedback (though unlikely, since training data is academic context, but a safety net is prudent).
- The model should be *fine-tunable*: we will collect a dataset of (submission -> instructor feedback) pairs during pilot to fine-tune the AI for better accuracy in our domain.
- **Accuracy and Testing:** Before deployment, test the AI on a sample of known assignments with known feedback to evaluate correctness. Adjust prompts or model choice as needed to improve alignment.
- Provide a mechanism for continuous learning: flagged corrections by instructors can be fed back to improve the model (this might be done offline by retraining periodically).

Integration with LMS

- Support common LMS systems (Brightspace, Canvas, Blackboard, etc.) through their APIs:
- When a student submits an assignment, LMS triggers our system (via webhook or periodic polling).
- Our system pulls the submission file/text. (Formats: PDF, DOCX, images, etc. – for text extraction, use PDF/DOC parsers; for images, possibly allow students to upload a brief self-description alongside).
- After feedback generation, push the feedback text back to LMS using the grading API (as an instructor comment draft or directly as feedback, depending on mode).
- Alternatively, deploy via LTI: the instructor adds our tool as an LTI assignment. Students submit within our tool interface which then passes grades back. (This route may be chosen if API integration is too complex initially).
- Data syncing: ensure student identifiers are consistent (use LMS user ID to link data).
- Security: our system should not expose LMS credentials; use OAuth tokens provided by LMS for API.

User Interface (UI/UX)

Instructor Dashboard UI

- Show a list of assignments; selecting an assignment shows a table of students, submission status, and feedback status (e.g., “Generated”, “Edited”, “Sent”).
- Clicking a student opens the feedback editor: a text box pre-filled with AI feedback. Changes are auto saved.
- Provide quick edit tools: e.g., a button to regenerate a section of feedback (if instructor isn't happy with one part), or insert a common comment from a library.
- Allow instructors to see original submission side-by-side (if text, display it; if image, show thumbnail).
- Indicate where AI mentions each rubric criterion (maybe highlight terms).
- Overall, UI should be clean, not cluttered, as instructors might use it while grading dozens of students rapidly.

Student UI

- If within LMS, it appears as normal feedback comment text. Optionally, we might format AI feedback distinctly (e.g., sections titled “AI Feedback” vs “Instructor Notes”).
- If through our portal, a simple interface showing their submission and the feedback. Possibly allow a reply or acknowledgment feature (students could ask for clarification, which goes back to instructor).

Mobile access: Ensure the student view (and possibly instructor view) are mobile-friendly, as students might check feedback on phones.

Accessibility: All UI should meet accessibility standards (WCAG 2.1 AA): screen-reader compatible, proper contrast, etc., so that all users including those with disabilities can use the system.

Data Management and Privacy

- **User Data:** Collect minimal personal data. Essentially, we use student name/ID from LMS and their assignment content. We do not need demographic or other personal info.
- **Storage:** Store assignment content and feedback securely. If possible, avoid storing full student submissions long-term (could process and discard, keeping only feedback and maybe a link to original in LMS). Feedback messages and any model outputs will be stored (for audit and improvement).

- **Privacy Compliance:** Adhere to FERPA (for educational records privacy in the U.S.) – the feedback and student performance data are protected. Thus, any data used to further train the AI must be de-identified or aggregated. We will obtain necessary consents if we use data for research beyond internal improvement.
- **Security:** Use HTTPS for all data transfer. Encrypt sensitive data at rest. Implement authentication and role-based access (instructors can only see their classes, students only their feedback). Regular security testing and compliance with university IT standards.
- **Data Retention:** By default, keep data for at least one semester after course ends (to allow review and grade appeals). Possibly purge or archive after a year unless needed for research (with approvals).

Logging and Analytics

- Log all AI interactions: when it generated feedback, what prompt was used, the output given, time taken, etc. This helps debug issues (e.g., if nonsense feedback was generated, we can trace why).
- Provide instructors with some analytics: e.g., a summary of common themes the AI saw (if we implement that), or even how many students had similar feedback points. This can guide class-wide discussions ("I noticed many of you got feedback about X, let's address that together").
- Analytics for admin/research: overall usage, average feedback length, turnaround times, etc. that can demonstrate impact (e.g., "feedback was provided within 1 day on average in courses using the system, compared to 10 days without").

DEVELOPMENT PLAN AND MILESTONES

- **Phase 0 – Planning and Design:** (Current stage) Completed needs assessment and specifications. *Milestone:* Detailed spec document approved by stakeholders.
- **Phase 1 – MVP Development:** Core features (automated feedback generation, LMS integration, instructor/student interfaces) developed and unit-tested. *Milestone:* Working prototype on staging server; basic end-to-end functionality with test data.
- **Phase 2 – Internal Testing and Quality Assurance:** Use sample assignments to test AI output quality; refine prompts/model. Fix UI/UX issues from test users' feedback. Security and integration testing with LMS sandbox. *Milestone:* MVP ready for pilot (all critical bugs resolved; security passed).
- **Phase 3 – Pilot Deployment:** Deploy in live course (small scale) as described in Discussion. Provide on-site/on-call support during pilot. Collect feedback and identify needed refinements. *Milestone:* Pilot evaluation report (including user feedback and any new requirements).
- **Phase 4 – Iteration:** Implement improvements from pilot feedback (could be UI tweaks, model tuning, new features like easier editing tools or additional prompts for specific scenarios). *Milestone:* Version 1.1 of system.
- **Phase 5 – Extended Pilot/Beta:** Deploy in more courses or a second institution to test scalability and different contexts. Monitor system performance under heavier load. *Milestone:* Approval for campus-wide (or broader) adoption.
- **Phase 6 – Scale Up and Maintenance:** Harden the system for production (load balancing, fail-safes), implement any remaining nice-to-have features (peer feedback integration, advanced analytics). Establish maintenance routines: model updates (e.g., incorporate new LLM versions when available), regular data audits for accuracy, user support channels (helpdesk). *Milestone:* Full release (v2.0) and transition to regular operation with support documentation.

Projected Timeline: Approximately 6 months for Phases 1–3 (development through initial pilot), another 3–6 months for Phase 4–5 (iteration and second pilot), so about 1 year from now to have a thoroughly tested product ready for wider use. Scaling and maintenance would be ongoing.

Projected Development Costs: (These are preliminary and subject to revision)

- *MVP Development:* $50,000 (covers developer team effort for ~4–5 months, including AI model integration which may involve API costs).
- *Pilot Support & Iteration:* $10,000 (covering tweaks after pilot and part-time support).
- *Scaling Infrastructure:* $5,000 (for enhanced hosting, additional servers, etc., if needed).
- *Total ~$65,000* for first year development. This aligns with industry cost estimates for similar AI EdTech projects.
- Ongoing costs: AI API usage (estimated a few cents per feedback; for 1000 feedback instances, perhaps $50 – negligible at pilot scale, but monitor if scaling to tens of thousands), maintenance developer time (could be $10-20k/year for updates), hosting ($1-2k/year).
- These figures will be refined in a detailed budget proposal for external funding. We anticipate seeking grants (e.g., NSF educational innovation) or institutional tech funding to cover this.

FUTURE FEATURES (POST-IMPLEMENTATION ENHANCEMENTS)

- **Adaptive Learning Integration:** Over time, use accumulated data to suggest course improvements. E.g., if 80% of students receive the same critique, suggest curriculum adjustment or provide additional resources to students in that area.
- **Student Feedback Literacy Tool:** Include a student-facing module that guides them on *how to interpret and use feedback* (whether AI or human). For instance, interactive reflection prompts after reading the feedback.

- **Multi-Modal Feedback:** Explore AI generating audio or video feedback (text-to-speech summaries, or image annotations on student work to point out specific areas on a design).
- **Gamification Elements:** To encourage engagement, students could earn badges for acting on feedback or for seeking clarification – fostering a culture of iterative improvement.

This technical specification is intended to serve as a living document. As we progress through development and pilot testing, we will update the specifications to reflect new learnings and changing requirements. The collaborative input of educators, developers, and students will continue to shape the evolution of the AI Feedback System, ensuring it remains aligned with the goal: enhancing learning in studio-based education by delivering the right feedback at the right time.

****