



Exploring the Collaborative Co-Creation Process with AI: A Case Study in Novice Music Production

Yue Fu
University of Washington
Information School
Seattle, Washington, USA
chrisfu@uw.edu

Michele Newman
University of Washington
Information School
Seattle, Washington, USA
mmn13@uw.edu

Lewis Going
University of Washington
Information School
Seattle, Washington, USA
lgoing7@uw.edu

Qizui Feng
University of Washington
Information School
Seattle, Washington, USA
qfeng5@uw.edu

Jin Ha Lee
University of Washington
The Information School
Seattle, Washington, USA
jinhalee@uw.edu

Abstract

Artificial intelligence is reshaping creative domains, yet its co-creative processes, especially in group settings with novice users, remain under explored. To bridge this gap, we conducted a case study in a college-level course where nine undergraduate students were tasked with creating three original music tracks using AI tools over 10 weeks. The study spanned the entire creative journey from ideation to releasing these songs on Spotify. Participants leveraged AI for music and lyric production, cover art, and distribution. Our findings highlight how AI transforms creative workflows: accelerating ideation but compressing the traditional preparation stage, and requiring novices to navigate a challenging idea selection and validation phase. We also identified a new “collaging and refinement” stage, where participants creatively combined diverse AI-generated outputs into cohesive works. Furthermore, AI influenced group social dynamics and role division among human creators. Based on these insights, we propose the Human-AI Co-Creation Stage Model and the Human-AI Agency Model, offering new perspectives on collaborative co-creation with AI.

CCS Concepts

- Human-centered computing → Empirical studies in HCI.

Keywords

Human-AI Co-Creation, Creativity Model, Human-AI Interaction, Generative AI, Music, Novice, Agency

ACM Reference Format:

Yue Fu, Michele Newman, Lewis Going, Qizui Feng, and Jin Ha Lee. 2025. Exploring the Collaborative Co-Creation Process with AI: A Case Study in Novice Music Production. In *Designing Interactive Systems Conference (DIS '25)*, July 05–09, 2025, Funchal, Portugal. ACM, New York, NY, USA, 15 pages. <https://doi.org/10.1145/3715336.3735829>



This work is licensed under a Creative Commons Attribution 4.0 International License.
DIS '25, Funchal, Portugal
© 2025 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-1485-6/25/07
<https://doi.org/10.1145/3715336.3735829>

1 Introduction

Artificial intelligence has afforded a new era of creativity, where humans and machines collaborate to produce novel and innovative works. Human-AI co-creation has emerged as a growing field of interest, capturing the attention of researchers and practitioners alike. The goal is to leverage the strengths of both humans and AI to produce creative outcomes that go beyond what either could achieve alone [80]. Such synergy has found applications across various domains, including writing [59], fashion design [25], architecture [72], marketing and advertising [39], visual arts [84], and more.

In music production, AI-powered tools have become increasingly popular, supporting both expert musicians [19, 23] and novices [46, 51] in the creative process. A plethora of AI music generation and production tools powered by foundational models [50] have been developed, aiming to facilitate music creation by increasing efficiency, enhancing self-expression, and supporting idea generation. AI music tools can compose melodies [41], harmonize chords [74, 75], transcribe [63], improvise [53], suggest lyrics [78], and assist in mixing and mastering tracks [68]. For novice users, AI lowers barriers to entry, enabling individuals without formal training to engage in music production and express their creativity. Platforms like Suno [9], Google’s Magenta [3], and OpenAI’s MuseNet [6] advertised they can democratize music creation by providing users with accessible tools that augment their creative capabilities.

Despite the growing prevalence of AI in creative fields, there is a notable gap in understanding the detailed processes of how creators collaboratively co-create with AI over extended periods. Much of the existing research has been conducted in laboratory settings or through online studies that capture only a snapshot of the creative process—often within a single session lasting less than an hour, using toy creative problems, and does not lead to final usable creative output. However, real-world creative artifact creation is typically a prolonged endeavor that involves various creative working stages and unfolds over weeks or months. In fields like architecture, design projects can span years, and in music production, though varying by genre and scope, it often takes weeks for a professional production team to produce a track from the initial concept to the final release.

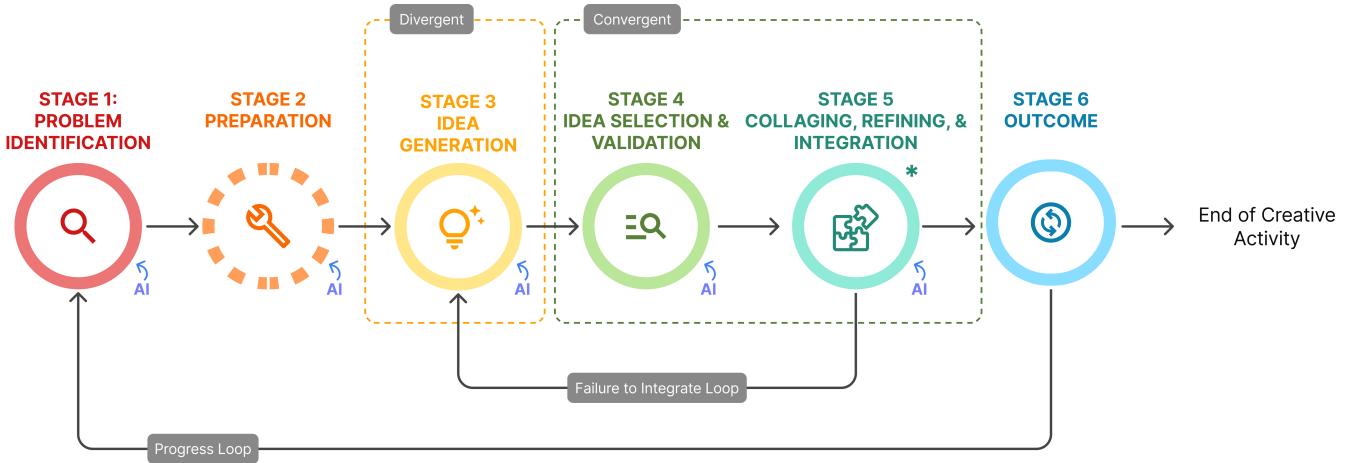


Figure 1: The Human-AI Co-Creation Stage Model. The diagram shows the creative process with AI support along a temporal dimension. As discussed in Section 4.1.2, introducing AI often shortens or bypasses Stage 2 (Preparation)—shown here with dashed boundaries. Although AI is praised for facilitating idea generation in Stage 3, it can also quickly lead to convergent stages in Stages 4 and 5, potentially linearize the creative process (see Section 5.3). Additionally, we identify a new Stage 5 (Collaging, Refining, and Integration)—marked by an asterisk—where creators must merge AI-generated and human-created elements. Should this piece-together process fail, users revert to Stage 3 to regenerate compatible outputs. Involving AI also facilitate achieving tangible results, fosters a “progress loop” [15], giving creators a sense of accomplishment and prompting them to revisit Stage 1 for subsequent creation cycles. Finally, AI can be viewed as an environmental factor that supports user autonomy, competence, or task involvement throughout the stages, with the exception of the end outcome (see Section 5.1.1).

Moreover, for much of creative processes, creation is a collaborative endeavor [64]. In disciplines such as design, tools like Figma [27] and Miro [5] have revolutionized how teams work together, providing ways for people to collaboratively brainstorm, vote, and comment. Similarly, in music production, collaboration is the backbone of the creative process. It usually involves a team of individuals fulfilling various functions, including lyricists, composers, mixing and mastering engineers, and more. Each role contributes specialized expertise, and the interplay among team members is crucial to the success of the final outcome. Using novice music production as a case study, our study aims to explore how AI tools influence collaborative dynamics and shape the stages of co-creation over an extended period. Music production represents an instructive example because it spans multiple creative roles (composition, arrangement, lyric writing, mixing/mastering) and often requires iterations to achieve a final outcome. Understanding both the collaborative dynamics and process of novice music production is essential for designing AI systems that effectively support collaborative co-creation activities between humans and AI. Specifically, we investigate three research questions:

- **RQ1:** What are the temporal stages of the co-creative process when novice users collaborate with AI in music production?
- **RQ2:** What relationships and roles do humans and AI assume in the collaborative music co-creation process?
- **RQ3:** How do users perceive AI in this process?

To address these questions, our case study explores how novice music creators work as a team to co-create music with AI over a period of 10 weeks. The setting of our study is an undergraduate capstone class at our institution, where students with no or limited

previous music production experience used AI tools individually and collaboratively to create music pieces. They explored various AI technologies for different stages of music creation, including composition, arrangement, lyric writing, mixing and mastering, etc.

Our research team reviewed each team’s artifacts, including course submission materials, process logs, published songs, group websites, and final presentations. We then conducted in-depth interviews with nine participants to understand their experiences, the social and technical challenges they faced, and how AI shaped their creative and collaborative processes.

We found that AI significantly accelerated ideation but shortened the preparation stage—potentially limiting the deeper domain knowledge acquisition that often develops early in creative processes. The stage of idea selection and validation became both more critical and challenging, as participants were frequently overwhelmed by abundant AI-generated options with bounded evaluation knowledge. We also identified a new stage—collaging, refining, and integrating diverse AI outputs with human-created elements to form a cohesive song. Additionally, AI mediated social dynamics and boosted self-efficacy among team members, yet it sometimes constrained emotional expression and artistic exploration. Based on our results, we propose two models, the Human-AI Co-creation Stage Model and the Human-AI Agency Model, which offer new perspectives on collaborative co-creation with AI and provide design implications for enhancing both the creative process and agency.

Our study contributes to the understanding of collaborative human-AI co-creation in the following ways:

- First, by focusing on a real-world, 10-week collaboration process, we capture the creative workflow across multiple stages and propose a **Human-AI Co-Creation Stage Model** that elucidates how creators and AI interact over time.
- Second, our analysis of collaborative teams, where members use AI both individually and collectively, reveals **AI tools mediate social dynamics and role assignment**.
- Third, by highlighting novice users' perceptions, we identify specific challenges and opportunities of AI-supported creativity in group settings. We then present a **Human-AI Agency Model** that argues for systems designed to nurture users' emotional expression, skill growth, contextual flexibility, and community building.

2 Related Work

2.1 Co-creation with AI in Music Production

Co-creation with AI refers to collaborative creative processes where humans and AI systems work together, leveraging respective strengths to achieve a shared creative goal. Human-AI co-creation represents a fundamental shift in how we interact with technology, moving from a tool-based relationship to a collaborative partnership [10].

Traditionally, music composition and production have relied heavily on human creativity and technical skill. Over the years, researchers in human-computer interaction and music education have explored this domain by developing digital musical tools—ranging from automatic mash-up systems and recording interfaces to gamified learning platforms and tutoring aids [38].

With the advent of AI-powered systems, new possibilities have emerged to automate and augment various stages of music creation, sparking discussions about where these tools fit into the creative process [48, 54] and how to ethically define roles and responsibilities for AI [22, 49, 56]. In particular, AI has been proposed as a means to support novice creatives. For instance, neural networks based system complemented by steering tools can restrict generative notes to particular voices and nudge output in high level directions, thereby enhancing a novice's sense of control and ownership [51]. Additionally, novices often lack foundational music knowledge and are deterred by the complexity of traditional production tools, making them prone to reduced motivation [47]. Researchers emphasize tailoring AI solutions to novices' specific needs, preserving enjoyment and intrinsic motivation throughout the creative process [47, 51].

However, previous studies have predominantly examined individual musicians' interactions with single-purpose AI tools in controlled settings [37, 52], leaving a gap in understanding how teams of creators, particularly novices, integrate AI tools across an entire production cycle. As AI evolves from discrete utilities to collaborative partnership [73], it is crucial to investigate how such technologies reshape collaborative processes. Our work explores how novice teams Co-create with AI tools over an extended music production period.

2.2 Introduction of AI Tools Used in the Case Study

Students in our study used a variety of AI tools. Here, we briefly introduce these tools and their functions to provide context for understanding our findings.

Traditionally, there are various production elements in music production, such as composition (creating music content, such as melodies and chord progression), arrangement (organizing musical elements), recording (capturing audio), mixing (balancing and adjusting individual audio tracks), and mastering (optimizing overall sound and tonal balance) [66]. Digital Audio Workstations (DAWs) like Ableton Live [4] serve as the modern central hub for music production, allowing creators to record, edit and manipulate both audio recordings and MIDI data (digital representations of musical notes) [21]. While DAWs have made professional-grade recording more accessible, they still require considerable technical skill to use effectively [21, 73]. Each part requires significant technical expertise and musical knowledge, creating substantial barriers to entry for novice creators [21, 66].

Recent AI tools have emerged to support different aspects of the production process. For composition and arrangement, text-to-audio generators like MusicGen [7] and Suno [9] can create musical elements from text descriptions. Tools like Samplab [8] can convert audio into editable MIDI data. For vocals, platforms like AIVA[2] can create synthetic voices or train AI models on recorded human samples. AI mastering services like LANDR [1] use machine learning to automatically optimize and balance tracks. In our case study, students experimented with a variety of these AI tools, each targeting different parts of the production elements. Appendix B provides a detailed summary of the AI tools students experimented and used, alongside their core functionalities.

2.3 Amabile's Stage Model of Creativity—Our Theoretical Lens

In the field of creativity research, scholars have proposed various models—such as the Four C Model [43], the Four P's model [32], the Five A's model [31], as well as stage-based frameworks like design thinking and Pearlman's model [60]—to capture the multifaceted nature of creative processes. Each framework offers unique insights into the origins, dynamics, and outcomes of creative work.

Amabile's Componential and Stage Model [11–13, 15] is among the most widely recognized and cited in this domain. It explains how creativity emerges from the interplay of (1) domain-relevant skills, (2) creativity-relevant processes, and (3) task motivation. Beyond these components, Amabile's model articulates five key stages of creative work: task identification, preparation, idea generation, idea validation, and outcome. These stages illustrate how individuals or small groups progress from recognizing a creative need to gathering information and skills, generating ideas, evaluating their viability, and deciding whether to accept a solution or loop back for further iteration. Additionally, Amabile emphasizes the role of the environment factors in fostering creativity, suggesting that supportive environment encourages risk-taking, autonomy, and open collaboration [13, 15].

We adopt Amabile's stage model as our primary theoretical lens (see Section 3.4) due to its broad applicability across creative domains and its longstanding influence in creativity research. Our findings illustrate AI's influences on each stage and need for a revision of the current model. Accordingly, we propose an updated version of Amabile's framework that illustrates how involving AI fundamentally transforms people's creative processes.

3 Methods

To explore how novice music creators co-create with AI tools over an extended period, we reviewed artifacts, process logs, and presentations by students who participated in a music production with AI course in the Spring of 2024. Additionally, we interviewed nine undergraduate students (out of 20 who attended the course). The interviews were conducted during the summer following the course completion, with each session lasting approximately 55 to 70 minutes. This study was reviewed by our institutional review board (IRB) and deemed exempt.

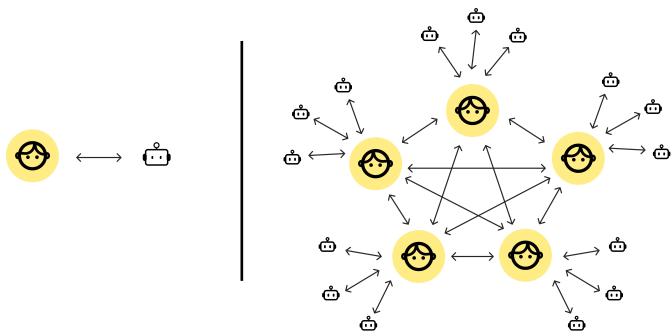


Figure 2: Conceptual Illustration of Individual Versus Collaborative AI Co-Creation. On the left, a single user interacts with an AI tool. On the right, it represents our case study context where students collaborated with each other and used various AI tools, reflecting a more complex and social situation.

3.1 Study Context

The participants were enrolled in a 10-week capstone course at our academic institution, designed to let students explore AI and music creation in the digital age. The course aimed to let students experience through the entire music production process—from initial ideation to final distribution—culminating in the release of their music on platforms such as Spotify and students' group websites.

Students were organized into teams of four to six members. Each team was tasked with collaboratively producing three tracks to be released as a single EP (see Figure 2 for a conceptual illustration of students' collaborative co-creation context). The instructors provided an introduction of various AI tools in the beginning of the course. The introduction included a demonstration of MusicGen via Google Colab and provided students curated links of other AI tools. After this initial introduction, students were encouraged to explore independently and choose tools best aligned with their creative vision and project requirements. Throughout the course,

students used a variety of AI tools to support different stages of music production, including lyrics writing, composition, melody creation, instrument, mixing and mastering, and for distribution purposes, EP cover and website design. They also incorporated their own singing and vocal recordings, sometimes renting professional recording studios to mix their voices with AI-generated outputs. The course structure included lectures, weekly presentations of each team's progress, and a final presentation showcasing their EP and reflections on the process.

3.2 Participants

Participants were recruited through the course instructor and teaching assistant after the course had concluded. We reached out to all 20 students who took the course, and nine agreed to participate in the study. Prior to the interviews, participants completed an initial survey to collect demographic information and their prior music production experience. The nine students we interviewed belonged to three teams: four students from the first team (P1, P2, P3, P4 - Team A), three from the second team (P5, P6, P9 -Team B), and two from the third team (P7, P8 - Team C). See Appendix A for participant demographics. Each participant received a US \$50 Amazon gift card as a token of appreciation for their time. See Appendix C for the major AI tools each team used during the music production process.

3.3 Procedures and Materials

We collected all the course homework materials from the instructor and teaching assistant, and reviewed the materials with participants' consent. This course provided a unique opportunity to examine the co-creation process systematically, as it explicitly required students to document their progress each week. These process materials encompassed iterations of group music concepts, pre-production plans, role assignments, ideation logs, reflections on experimenting with new AI tools, iterations of songs and lyrics, feedback from listening sessions, and participants' reflections on the feedback. The research team internally reviewed all these materials, including class presentations, collaborative online canvases (such as Figma), final lyrics, produced EPs available on Spotify, and each team's website for publicity and distribution. One research team member served as the teaching assistant for the course. This member provided a detailed, reflective document outlining students' experiences, challenges, processes, and specific AI tools explored during the class. This documentation was compiled from personal recollections and process logs submitted by students throughout the course.

The research team met weekly to discuss and analyze the collected materials, identifying emergent themes and formulating potential research questions. Then three team members independently drafted semi-structured interview questions based on our agreed research questions. Over the next three weeks, the team discussed and refined interview questions, resulting in a final interview protocol comprising four main parts. The first part asked participants about collaboration dynamics. We explored the students' collaborative roles within their teams and their perceptions of AI's role in production. We also asked about the benefits and challenges faced when incorporating AI into a team setting. The second part asked

about students' experiences with the music production process, focusing on how the involvement of AI technologies affected various music production stages and their individual workflow. The third part of the interview protocol explored participants' reflections on the 10-week music-making with AI journey. We asked their perceptions of how AI influenced their self-efficacy, creative expression, sense of ownership, and control and pride over the music pieces. During each interview session, the interviewer responsible for reviewing the team's material. Interviewers frequently incorporated excerpts from these materials, including listening recorded EPs together with interviewees, to prompt their specific memories and reflections.

Most interviews were conducted in the summer to optimize participants' recollection of their experiences (the course was held during the spring quarter). Only one interview was conducted later after the summer break due to participant availability. Interviews were designed to last 60 minutes. All interviews were audio-recorded and transcribed. We anonymized and securely stored all audio recordings.

3.4 Data analysis

We employed an abductive approach to analyze and understand the case by both existing related theories and empirical data collected. The abductive approach is often used in case studies [62] and the process systematically combines theoretical framework, empirical data, and case analysis [26]. The aim of our abductive approach is to provide ideas and a subsequent tentative theory [62].

We began analyzing the data and reviewing theories concurrently with the interviews. During the first round of coding, we used Reflexive Thematic Analysis [26]. One team member read the transcripts and revisited the audio recordings multiple times. The teammate extracted quotes and organized them into preliminary themes. The research team met weekly to discuss initial understandings, share insights, and compare interpretations of these quotes and themes. Through an iterative process, we agreed upon themes related to participants' usage, perceptions, and experiences of AI in the music production process. These themes encompassed various aspects such as collaboration dynamics, challenges and benefits of using AI, the influence of AI on creativity and workflow, and participants' perceptions of AI in the music production process.

As our analysis progressed, we identified an alignment between the emerging themes and Amabile's Model of Creativity [13], which delineates creative stages. Recognizing this congruence, we integrated Amabile's model as a guiding framework to further structure and interpret our data. During subsequent meetings, the research team discussed how our data both supported and extended Amabile's model. We reached a consensus to adopt the theory for the next phase of analysis. This involved re-examining part of the data with Amabile's model in mind, and coding transcripts according to the creative stages of the model. We remained open to the possibility of modifying the model based on our findings.

The research team collaboratively refined the themes to reflect the expanded framework of creativity. Using an online collaborative tool (Figma), one researcher organized related transcript quotes into stages of creativity and human-AI collaboration themes.

During a half-day data analysis workshop, four researchers collaboratively mapped participants' quotes, identified opportunities and challenges at each creative stage, and analyzed the workflow and iterations between stages in Figma. This workshop also focused on understanding human-AI collaboration within the creative process. Themes related to human-AI collaboration and participants' perceptions of AI were documented and refined alongside the expanded creative stages. One researcher then synthesized the workshop findings in Figma, illustrating the expanded stages of creativity, directional links between stages, and overarching themes of human-AI collaboration. Subsequently, the lead author consolidated all the coded data, themes, and interpretations from the Figma board, as presented in the Results and Discussion sections.

4 Results

4.1 Collaborative Co-creation Stages with AI

4.1.1 Stage 1: Problem and Task Presentation.

Defining Creative Vision and Objectives. In the initial stage, AI played little role. Participants focused on defining what they wanted to create, identifying the central theme, vision, and clarifying the message or "story" they aimed to convey. They discussed broad questions such as "*what kind of story [we] want to tell and the feeling [we] want the sound to resonate*" (P9). One participant explained, "*we need to come up with the problem statement that we can explore different AI usage in the music production process*" (P3), while another noted the importance of aligning their vision by understanding "*what we want to express [and] what we lack*" (P4). These discussions not only guided their creative goal but also helped the teams form a collective motivation for the project.

Establishing shared goals and building consensus required active collaboration. Teams often held brainstorm sessions during which members presented ideas, preferences, and musical inspirations. These early meetings were critical for clarifying a *big idea* or thesis for their music projects.

Assigning Roles and Responsibilities. Once the project vision began to crystallize, teams allocated tasks to specific members based on interests and skill sets. As one participant explained, "*We divide up the work at the beginning during the brainstorming stage... once we settled on a theme and genre, we [assign] the roles*" (P7). Some teams assigned imaginary roles to AI, referring to AI as a "teammate" (P4) and one group even gave it a playful name, "Strawbae" (P8), which aligned with their band's identity, signaling their intent to integrate AI as a meaningful player in this process.

We found this stage to be relatively brief, typically lasting only a single brainstorming or workshop session. During this time, students were assigned specific tasks that they would later support individually by using AI tools in the subsequent stages of the music production process.

4.1.2 Stage 2: Preparation: Building Up Relevant Information and Skills.

In a traditional creative process, the preparation stage involves gathering knowledge, researching relevant information, and developing domain skills, which involves a great deal of learning and time-consuming [14, 17, 61]. However, in our case, students advanced quickly from Stage 1 (problem presentation) to Stage 3

(ideation). AI enabled fast idea generation without much preparation, reducing the need for knowledge-building at the outset.

When asked about their production process, none explicitly mentioned a distinct preparation or research phase. Their “*production plan*” documents submitted as class assignments showed an immediate pivot from initial brainstorming to active ideation. While the 10-week capstone structure likely created external pressure to progress quickly, participants repeatedly noted that AI helped them dive into music creation with minimal prior expertise. Interestingly, several participants described using AI generation as a way to build domain knowledge. One student explained:

“We would create a sample using software [AI]. Then downloaded it, fed it into [an audio-to-text] AI tool. It [the tool] would read that music clip and generate text based on how it sounded... From there, we put that text into another tool to see what it would generate” (P8).

Through feeding musical snippets into AI and receiving feedback or textual descriptions, students began to understand specific terminology and refine their vocabulary to prompt other AI systems. Another participant mentioned that an AI tool provides sample clips and patterns for the creator to generate and combine ideas, saying it “*helped [us] think creatively about what electronic or techno music we wanted to create*” (P7), even though the final tracks did not incorporate that tool’s outputs.

Overall, we observed a blurring of the preparation and ideation stages. Rather than devoting significant time to mastering musical fundamentals, novice users relied on trial-and-error with AI to gain the knowledge needed for creation. This stands in contrast to the long “warm-up” periods often noted in creativity literature [12, 71, 77].

4.1.3 Stage 3: Idea Generation.

The third stage of creativity involves generating ideas and exploring solutions. However, with AI, participants described a fundamentally altered process. Traditionally, individuals draw from their cognitive pathways and environmental resources [14], but AI enabled them to produce music with less reliance on human cognitive and environmental resources. Our data suggest that AI served as a powerful springboard for inspiration, speeding up the ideation process, and a potential source of idea fixation on initial AI’s output.

AI as a Springboard for Inspiration. Many participants praised AI for jumpstarting their music making, even when they ultimately did not incorporate some of the AI-generated content in the final work. One said, “*you can get a lot of cool ideas, and especially for people who don’t have any background in music. It does, I think, increase accessibility around how to create music or getting ideas*” (P8). Another echoed, “*We ask ChatGPT to help us create lyrics on a certain theme... we might not use them directly, but they inspire us*” (P7). Even when the team chose not to use the AI outputs, they perceived AI as useful. For example, one team generated musical clips via tools like Suno only to discard them while inspired by the ideas from the AI suggestions (P8).

Direction Setting and the Risk of Idea Fixation. Beyond providing inspiration, AI often guided participants’ musical direction. Several noted that “*it can quickly generate different types of songs with*

different styles... which sets the basis [for discussion]” (P4). Others found text-to-audio tools useful for “*getting a general vibe... [that] gave us some direction when we were lost*” (P8). In many cases, the AI’s initial output shaped the foundational chords, beats, or melody line, and for some groups, effectively “*70 or 80 % of the whole song*” (P1).

However, this convenience came with a risk of over-reliance or idea fixation [40]. As one participant put it, AI may “*make you throw away ideas that could have been good if you had just developed them*” (P6). Some team members acknowledged their limited knowledge, which prompted them to “*just go with what [AI] gave us*” (P6). Novice status and AI’s “impressive” capabilities (P2, P4, P5) frequently led participants to adopt AI outputs without extensive exploration of alternatives.

Accelerating the Ideation Workflow. Participants repeatedly cited AI’s speed and efficiency in generating content. One student called AI “*really efficient... [it] can give us a whole song in maybe 30 seconds*” (P3). Others described it as a “*shortcut*” (P9) for tasks like chord generation and beat making. One student explained, “*It would have taken more time if we didn’t have AI... we were able to find chord progressions that fit well together*” (P8). Specifically, participants repeatedly appraised AI’s ability to quickly generate lyrics (P2, P3, P5, P6, P8), beats and chord progressions (P1, P7), and simulate people’s singing voices (P1, P4, P7, P8).

4.1.4 Stage 4: Idea Selection and Validation.

After generating a plethora of ideas generated by AI in Stage 3, participants moved on to selecting these ideas. As one student succinctly put it, “*The output that AI generates needs to be selected by humans... The final choices are made by humans*” (P1). Traditionally, individuals use domain-relevant skills acquired during Stage 2 to evaluate ideas for correctness or appropriateness [14]. However, because participants had little knowledge-building preparation, and AI could quickly produce numerous options, selecting and validating ideas became challenging.

The Challenge of Validating Abundant AI-Generated Ideas. Participants frequently felt overwhelmed by the sheer number of AI-generated options, finding it time-consuming to sift through them and decide which ideas merited further development. One participant recalled, “*We spent a large chunk of time trying to find something that we felt fit... And then it did become a little frustrating because we want to make progress on the song*” (P8). Another participant mentioned generating a wide range of lyrics then need to filter out subpar content: “*That’s not to say I was blindly accepting AI stuff... a lot of the lyrics were just really corny*” (P6). Lacking formal rubrics or deep domain expertise, many teams relied on gut feelings and intuition. One participant described hearing a promising AI-generated beat at the fifth iteration and disregarding the rest: “*It just really resonate[d]... We couldn’t listen to the sixth or seventh iteration. We just stuck with that fifth one*” (P1). Another noted that, “*We tried a lot of stuff with AI but only a small fraction made it into the final song*” (P9). For these novice producers, the subjective sense of what “sounds good” often served as the default standard.

While the abundance of outputs initially seemed advantageous, it often led to decision paralysis when students were faced with numerous choices [65], especially under the tight deadlines of a 10-week capstone course. One novice student noted their team was

unable to appraise the various ideas generated by AI, “*At first, every idea was like a good idea. So we just tried to throw things together, but eventually, we realized that there were too many tracks and it was too cluttered.*” (P6). Another participant noted that the selection phase was time-consuming, “*I think there was definitely a lot of inspiration we drew from what we generated using AI..But it was frustrating when we spent so much time.*” (P8).

Time constraints yet prevented endless iteration. By limiting the available time for ideation, the course structure forced teams to converge on a final choice. As one participant mentioned, “*If there is no time limit, we will have endless discussions*” (P4). Several participants recognized the need for a decisive leader to streamline the process, “*We have to arrive at a conclusion very quickly...[A] decisive person...makes the decision. We just used the idea with some adjustment*” (P2). Thus external pressures and clear decision-making role can help manage the potentially overwhelming influx of AI-driven ideas.

Balancing Individual vs. Collaborative Selection. Teams differed in how they worked on idea selection and validation, with some emphasizing individual work before collectively deciding, while others favored a collaborative process from the outset. Certain participants iterated on AI outputs independently, “*I iterated many times...then chose a good one and showed it to the group*” (P5), picking out viable lyrics or melodies from various generations on their own. In contrast, other teams tackled selection more communally. One group “*had a session where we each pulled up ten to twenty iterations and then shared them*” (P7). Another group “*used Suno maybe 50 times and chose the best melody together*” (P3). Whether individually or collectively driven, both approaches tried to reach a consensus efficiently and respected each teammate’s contribution.

4.1.5 Stage 5: Collaging, Refining, and Integration with Human-Created Elements (The New Stage Caused by Involving AI).

As a result of incorporating multiple AI tools and outputs into the creative process, participants introduced a new collaging and refinement stage in which they assembled materials from various AI sources into a cohesive musical work. In our case study, students typically found themselves working with scattered AI outputs often generated from different AI tools—melodies, chord progressions, vocal snippets, and sound effects—none of which fit together seamlessly. Additionally, they often needed to refine and adjust these fragments, then combine them with human-created elements such as raps, vocals, and instrumentation. In this stage, creators had to integrate an assortment of discrete pieces into a cohesive musical composition, much like stitching together diverse elements in a visual collage.

One student mentioned, “*everyone may use AI to do some different parts*” (P4), and another explained, “*we used a lot of [AI tools], then we combined ideas... I actually spent [a great deal of] time pulling them together*” (P2). The same student described in charge of putting teammates’ AI-generated chords, beats, or lyrics and then using a digital audio workstation to merge and modify them: “*It starts with a melody...chords from another person...and lyrics from another teammate*” (P2). P8 summarized succinctly, “*We blended the outputs*” (P8).

Necessity of Refinement and Adjustment. Integrating AI outputs also demanded refinement of each AI output. One participant

said, “*AI just produced the idea, and most importantly, it is our work in editing that*” (P2). Another explained how they had to alter lyrics for better rhyme or tweak generated chords to suit the group’s performance, “*If you want a comprehensive, emotional song, you have to do some refinements by yourself*” (P3). Some misalignments, such as slight key differences or awkward transitions, forced participants to manually fix them to meet their musical standards (P2).

Inflexible AI Audio Output. Many teams found audio outputs from AI to be particularly challenging to work on. Participants often describe AI audio as “*fixed*” (P5) and “*hard to manipulate*” (P6). They pointed out that most AI systems only provide WAV or MP3 files, preventing users from easily adjusting individual notes or tracks. One student said, “*tools like that are really cool, but it’s really hard to like deconstruct that and try to make something on your own with it because it’s already like a finished product*” (P6). Across participants, students emphasized their desire for more control and flexibility over AI’s output.

Difficulty Mixing Human and AI Components. In addition, when blending AI-generated segments with live instruments or vocals, participants ran into further complications: “*We tried to get a melody by integrating another instrument...that wouldn’t work because if you combine these two elements, it was just weird*” (P3). Teams reported that once they had recorded human vocals, synchronizing or aligning with personal performance nuances was frustratingly complex. As one participant summed up, “*It’s really hard to incorporate [AI output] without it taking over the song...like 95% AI and 5% us*” (P6).

Time-Consuming Editing Processes. Ultimately, assembling and refining these disparate pieces required significant time and skill. One participant described manually entering notes into Ableton, spending “*four to five hours just editing the chords*” (P2). The same student observed that mismatches in AI’s default musical keys forced students to retune or rewrite entire parts: “*It’s all black keys, but we sing in white keys, so it’s a serious problem*” (P2). During this stage, considerable human effort is required to integrate into a polished final product.

Collectively, students often felt AI was falling short at this point in the process. While it could supply large quantities of raw material, integrating those diverse outputs into a polished track demanded significant human effort and expertise.

4.1.6 Stage 6: Outcome.

Amabile’s original model posits three possible outcomes for creative efforts: success, failure, or partial progress toward a goal [12]. In our study, AI mitigated the risk of “no reasonable response possibility generated,” [14] providing participants with “psychological safety” [14] for further exploration. As one student affirmed, “*I feel like we succeeded. I don’t think I ever would have made a song if I didn’t have AI help me*” (P6). Even if the first song did not meet everyone’s exact expectations, AI frequently offered enough momentum to ensure progress. By the end of the course, each team produced and published three complete EP tracks—an accomplishment that students unanimously viewed as a success. The shared artifacts students created, including published EPs, helped them solicit feedback from peers, friends, and family. This feedback cycle reinforced what Amabile describes as a “progress loop,” [15]

whereby tangible advancement in meaningful work boosts intrinsic motivation and improves creative performance in subsequent iterations.

During this final stage, participants also developed more objective views of AI's limitations, particularly regarding the challenges of integrating and refining its outputs (see Stage 5). These insights prompted them to adapt how they would like to use AI in the next round. In addition, knowledge and experience gained from the initial trial is added to the participant's domain skills. Over this process, students developed confidence and agency, as one remarked, *"Making an album with other people and AI sounded really hard at the beginning, but at the end, we found it wasn't hard with teammates and AI"* (P7). Notably, students often decide to rely more on their own rather than AI for future iterations. Collectively, our findings indicate that during this stage, students consciously reassessed the appropriate level of AI involvement, reevaluated their expressive needs, and integrated the domain expertise they had acquired before embarking on future iterations.

4.2 Human-AI Collaboration

Incorporating AI into the collaborative co-creation process reshaped social dynamics and influenced how students perceived their roles in the creative activity. While some teams treated AI as an additional "teammate," participants acknowledged that AI differs fundamentally from human collaborators. This section details how AI mediated social interactions, how participants conceptualized AI's roles, why they valued human emotional input, and how they retained control over AI's contributions.

4.2.1 AI-Mediated Social Dynamics.

Sensitivity in Human-Human Collaboration. Creating music as a team involves emotional risk and personal investment, making feedback exchanges especially delicate. One participant described learning "*when to step in and give feedback and when to let people do their own thing*" (P6), highlighting the tension between honest critique and respecting others' creative autonomy. Many admitted they were too polite to challenge ideas they did not like. *"People are a little hypocritical... nobody wanted to say bad things"* (P5). Another participant observed how each member's different tastes demanded ongoing compromise: *"We each have different interests... there's got to be some compromise there"* (P6). Because most were still getting to know each other, they often defaulted to being "*receptive to everyone's ideas*" (P6), even at the cost of constructive critique.

Emotional Ease and Honest Critique with AI. In contrast, participants felt more comfortable offering blunt critiques of AI outputs. One student noted, *"It's easy... you don't have to worry about hurting AI's feelings"* (P6). They routinely ignored or modified AI outputs generated by themselves or others without fearing interpersonal conflict. P6 explained, *"When your teammate shares something [they made]... you have to treat them gently, whereas with AI, it's easy to ignore 90% of what it says"* (P6). A different participant even credited AI for reducing social barriers around sensitive song topics, remarking that *"a judgment-free AI allowed us to bring walls down and execute on our vision"* (P9). While AI was treated as a creative partner, its emotional neutrality made it easier to critique and reject its suggestions, a stark contrast to human-to-human interactions.

4.2.2 Perceived Roles of AI in Music Production.

AI as a "Team Member" and Co-Creator. Some participants described AI as a creative partner that actively contributed to idea inspiration, lyric composition, and other roles. As one noted, *"It acted as both a songwriter and a lyric writer... like two group members at once"* (P2). Another participant referred to AI as *"a fifth person in the room... someone to bounce ideas off of"* (P6). Some teams even anthropomorphized the AI, giving it a name such as "Strawbae" to signal its perceived membership in the band. However, expectations of AI's ability to maintain a team member role over time in creative collaboration often fell short. According to one student, *"We wanted Strawbae to be a band member, but... we found it was really difficult to give AI that kind of character"*, and by the end of the term, they had *"less confidence"* in it than they initially expected (P8).

AI as Technical Assistant or Domain Expert. Others viewed AI primarily as a powerful assistive tool, filling technical gaps where team members lacked expertise. One student believed AI should *"provide technical support and creative inspiration"* (P3), while another admitted they *"regard AI as having more expertise than us"* (P6), especially in mixing, mastering, or specific stylistic domains. Additionally, P2 admitted AI takes on tasks outside the participants' expertise, such as creating visuals for album art.

4.2.3 Valuing Human's Emotional and Creative Expression.

AI's Perceived Mechanical Output and Lack of Emotional Depth. Participants frequently criticized AI-generated music as lacking the emotional depth that human-made compositions convey. One student commented, *"it lacks a certain amount of rawness you get from music made by humans"* (P8). Another used a metaphor to describe AI outputs as *"too simple... like orange juice without added ingredients"* (P2), highlighting their view that AI compositions do not capture the emotional nuance that resonates with listeners. A third participant recognized AI's *"excellent technical processing,"* yet lamented its inability to supply *"emotional expression, the space where human creativity remains irreplaceable"* (P3). Others echoed similar sentiments, describing AI music as *"mechanical and stiff"* or *"formulaic and monotonous"* (P5).

Importance of Emotional Connection. Students underscored the emotional connection between listeners and creators. One participant explained that listeners *"want to be attached to the real people behind it"* (P6), arguing that a composer's lived experience or personal identity shapes how an audience emotionally engages with the work. Another student asked how one could meaningfully relate to an AI-generated piece: *"We relate symphony number five with Beethoven, his own experience and his own idea, his own expressions... But if this song is also written by a certain kind of AI model, maybe it reminds me of Beethoven, but it's not really written by Beethoven. How can we relate this piece with the AI model itself?"* (P1).

Several participants stressed the need to retain human emotional involvement, even as AI becomes more capable. One argued that *"final artistic expressions still need deep human involvement and emotional investment"* (P3). One participant summarized succinctly, *"the optimal blend of AI and human creativity lies in finding a balance where AI's efficiency complimenting human emotional expressiveness"* (P3), a stance reflecting the overall consensus.

4.2.4 Value Ownership.

Participants expressed concerns around maintaining artistic ownership when collaborating with AI tools. The possibility of AI overpowering human contribution created a sense of unease. One team explicitly preferred to use multiple, less powerful tools instead of relying heavily on a single, advanced “*Music AI*” solution due to concerns that it might “verge on making its own song” (P6). Participants emphasized maintaining artistic ownership when collaborating with AI tools, as one participant articulated, “*I think the most important lesson I learned is to not let AI lead the way because music production in the end is like expressing your own identity and idea*” (P7). Another participant further echoed this perspective by asserting: “*I am expressing myself. It's my own manner. It might be different for other musicians, but my own manner is like, if AI is going to do it for me, it's AI's. It's not mine*” (P1). This highlights a collective determination to preserve personal creative identities and avoid losing themselves to AI-driven processes.

4.2.5 A Conscious Choice to Retain Control.

Participants also demonstrated a strong and deliberate effort to retain control over the music production process. They consistently described actively engaging with, questioning, and editing AI-generated content rather than accepting it without critical review. One student explained how they would “*question each line*” in AI-created lyrics, ultimately scrapping many of them (P6). Another participant noted that “*AI generates better results, but at the end it's mainly us who wrote the prompts and chose... whether to use the end results*” (P7), highlighting their conscious decision to maintain authority over the creative process. Over time, as participants became more experienced, they gravitated towards AI tools offering more adjustable parameters, thus providing greater flexibility and control. They repeatedly emphasized that human goals and critical judgment should remain central, saying clearly that “*we can use AI to generate our idea, but never let it replace our thinking—we have to have our own thinking*” (P2). They concluded that humans must ultimately “*select*” from AI outputs (P1), ensuring their control of the process and outcome.

4.3 AI's Impact on Creative Expression and Self-efficacy

Participants expressed both positive and negative views regarding AI's influence on their creative expression, praising its ability to boost confidence while also critiquing its constraints on artistry and collaboration.

4.3.1 Enhancing Creative Confidence and Self-Efficacy.

Many students credited AI with bolstering their creative confidence and enabling them to explore musical ideas they might not have attempted otherwise. One participant acknowledged they “*never would have made a song*” without the reassurance provided by AI (P6), explaining that simply knowing there was a tool to fill gaps in their knowledge empowered them to experiment their ideas. Another described the experience of using AI as “*pushing me outside my comfort zone*,” which eventually led to personal growth and a willingness to try to produce music in the future, “*if that's not like what I'm gonna do for the rest of my life, it could still be like a fun side hobby*” (P8). For newcomers to music production, AI

mitigated the “*cold start*” problem and made the creative process more inviting (P7). Consequently, many felt a boost in self-efficacy, taking pride in what they had achieved and viewing the experience as a stepping stone for future creative endeavors.

4.3.2 Potential Constraints on Artistic Expression.

However, participants with stronger musical backgrounds or a desire for deeper artistic control sometimes viewed AI as limiting their creative expression. One student lamented that relying on Suno to generate entire songs reduced opportunities for collaboration and hands-on experimentation, leading them to “*spend less time with each other...or with real physical instruments*” (P7). Others found that constantly regenerating AI outputs became “*overwhelming and disruptive*” (P1), sapping their motivation to pursue personal melodies or ideas. In some cases, frustration with subpar or repetitive AI content generation caused teams to “*spend so much time*” trying to salvage material that ultimately failed to meet their artistic vision (P8). While AI undoubtedly boosted confidence for novices, it could also stifle the creative instincts of those who wished for a more human-driven and embodied approach to music-making.

5 Discussion

5.1 Design Implications for the Human-AI Co-Creation Stage Model

In this subsection, we connect the results in Section 4.1 and propose a Human-AI Co-creation Stage Model (see Figure 1). The model builds on and extends Amabile's original model of creativity [12]. We also discuss the challenges and opportunities AI offer and propose a list of design implications for each stage of the co-creation process.

5.1.1 Protean Roles of AI and Supporting Initial Problem Space Exploration.

Many students in our study were initially uncertain about AI's potential, often personifying it as a “fifth teammate”. However, AI is neither human nor strictly confined to any single role [76]. Instead, it can perform multiple functions—from generating chord progressions to synthesizing vocals in our case. Viewing AI as an environmental factor [15, 57] or resource, rather than a discrete collaborator, may offer a more adaptable and efficient framework for creativity. **Environmental factors** refer to external influences that either support or hinder the creative process. Using this perspective, the presence of AI can impact an individual's intrinsic motivation, access to resources, and ability to engage in creative thinking.

AI systems, particularly those designed for novice creators, could significantly enhance initial exploration of problem and solution spaces. Novices, limited by their domain knowledge, often struggle with effectively navigating creative problem spaces and identifying directions. AI system designers should consider mechanisms that **externalize and illuminate the creative problem spaces** for users, providing structured guidance or exploration cues. For example, systems could incorporate high-level overviews of possible musical genres, thematic prompts, or even visual inspiration in the form of preliminary mood boards, thereby helping novice creators systematically scope and understand the creative possibilities at their disposal.

Moreover, incorporating **multi-agent AI systems that simulate various persona roles** can further enrich the initial problem-space exploration by introducing multiple perspectives. Such multi-agent systems, as demonstrated in recent collaborative AI frameworks [30], can foster diverse dialogues that human collaborators may not naturally consider due to limitations in their experience or domain-specific knowledge. Encouraging users or creative teams to interact with multiple AI personas rather than assigning a singular, narrowly defined function to AI may thus facilitate richer, more multidimensional problem space exploration.

5.1.2 Supporting Preparation and Domain Knowledge Building.

In our study, the Preparation stage (Stage 2) was often compressed or entirely bypassed, with participants swiftly transitioning from Task Identification (Stage 1) to Idea Generation (Stage 3). We attribute this rapid progression to the availability of powerful AI tools that facilitate quick ideation, thereby reducing the perceived need for an extensive “warm-up” or preparatory period. However, the preparation phase and the development of domain knowledge are critical not only for understanding the creative activity initially but also for effectively evaluating ideas during Idea Selection and Validation (Stage 4). The absence of a thorough preparation stage may undermine creative outcomes, as creators under external pressure might favor more predictable, AI-generated solutions without sufficient domain expertise [14].

Despite these challenges, there is significant potential for AI to support more systematic knowledge and skill exploration. Foundational AI models can provide domain-specific knowledge, sense-making support, and contextual explanations tailored to novices’ needs. Designers should therefore **integrate learning objectives** within AI systems, especially when these systems are expected to serve large novice user populations. For example, in the domain of image generation, tools like CreativeConnect [24] enable users to analyze reference images, break down suggested keywords into categories, and recombine elements to create new images with AI assistance. By delivering contextual explanations and tailored domain knowledge, AI can provide a structured preparation phase before engaging in ideation.

Moreover, **intentionally delaying the introduction** of the AI-driven ideation phase to incorporate a “training the brain” period may foster deeper cognitive engagement and mitigate the risk of early cognitive fixation on suboptimal ideas [33]. Aligning this design approach with educational theories of “productive struggle” [79], it is beneficial for users to grapple with challenging problems during this stage before accessing AI-generated ideas. By deliberately delaying immediate AI idea generation, designers can encourage deeper cognitive engagement, mitigate the risks associated with overreliance, and reduce anchoring biases.

5.1.3 Supporting Divergent Inspiration and Preventing Convergent Fixation.

Students in our study consistently praised AI for its ability to rapidly generate numerous creative ideas, serving as a powerful tool for inspiration and exploration. As multiple studies have noted, AI systems can function effectively as brainstorming aids [45, 67, 82], offering a seemingly endless array of possibilities. However, this capability also tends to drive an early convergence in the ideation

process. Traditionally, creative ideas emerge gradually, often following a prolonged incubation phase during which vague ideas evolve into clear, innovative concepts [77]. In contrast, AI frequently delivers near-final or highly refined outputs. This immediate availability of polished ideas can narrow a creator’s exploratory range, leading them to focus on *perfecting specific prompts* rather than venturing into broader, alternative creative pathways. To avoid this pitfall, it becomes crucial for AI system designers to **actively encourage idea divergence** rather than merely providing the most straightforward or obvious ideas. Potential design features such as **“big-picture” overviews** of stylistic or thematic **variations** can encourage ongoing divergence and prevent users from locking onto the first promising AI-generated concept. Also, AI systems that prompt reflection on initial goals (Stage 1) can support users’ metacognition and help them stay open to broader creative directions.

Additionally, inspired by human practices in art and design, introducing **intentional ambiguity** into AI outputs can serve as a resource for imagination and interpretation [29]. Rather than presenting polished and fully realized ideas, AI systems could offer incomplete or ambiguous ones, such as rough narratives or metaphorical prompts, to intentionally invite human creators to fill in the gaps, select resonant ideas, and infuse personal meaning. In practice, AI systems might feature multiple rough drafts or even a “wild ideas” mode, explicitly signaling to users that these suggestions are intended as creative seeds.

AI-enabled creativity also shifts the burden from generating ideas, which traditionally a significant obstacle, to selecting and evaluating them. Participants in our study expressed decision fatigue [35] when confronted with numerous AI-generated options. Without extensive domain knowledge or preparation, they often defaulted to intuitive judgments, such as “it sounds good”. Creators may gravitate toward a *satisficing* choice, selecting an adequate option rather than striving for the optimal one, an example of bounded rationality [70]. Being highly subjective, music domain may validate such a gut-level approach. However, over-reliance on intuition risks cognitive biases, akin to System 1 thinking overshadowing more deliberate System 2 reasoning [42] and intuition judgment may not apply to other co-creation domains such as marketing and architecture design. To address this, AI tools could integrate **evaluation guidelines and metrics**, or offer scaffolding that helps creators formulate their own. By clarifying the criteria for “success,” such features help users more accurately gauge the quality and relevance of AI outputs. Also, **automated curation features might filter out obviously low-quality outputs** and provide explanations for each dismissal, reducing cognitive overload. Clustering or categorizing similar ideas further streamlines comparison and selection, allowing users to focus on the most promising concepts and expand their domain knowledge.

5.1.4 Challenges in Collaging, Refining, and Integration.

Our findings reveal a new stage that extends beyond Amabile’s original framework: Stage 5, focused on collaging, refining, and integrating AI outputs with human-produced material. Many described the process as time-consuming and technically challenging,

with AI outputs often arriving in uneditable formats that impede integration. Consequently, promising ideas are sometimes abandoned simply because they cannot be made to “fit” with other elements.

To address these integration challenges, AI systems should aim to produce more **flexible and easily editable outputs**. Specifically, designers could design AI-generated content to be available in formats that facilitate detailed adjustments—such as separate MIDI tracks, discrete vocal stems, or clearly defined audio layers. Incorporating **real-time or in-situ editing capabilities** directly within the AI generation interface could enhance workflow efficiency and creative control. Moreover, it is critical to streamline and optimize the workflow pipeline between different AI tools and digital audio workstations. Seamless export-import functionalities should be prioritized, allowing AI outputs to transition smoothly into DAWs or other software environments to reduce the manual overhead.

5.1.5 Sustaining Progress and Motivation.

Our results (Section 4.1.6) show all groups completed an EP consisting of three songs, demonstrating involving AI leads to success or at least help human creators gain some progress towards creative goal established in Stage 1 (Task Presentation). By achieving tangible results, creators experience a “progress loop” [15], where a sense of accomplishment boost motivation and confidence, prompting them to loop back to Stage 1 for another cycle of ideation. This positive feedback cycle can foster high levels of creativity over multiple iterations.

AI system designers should, therefore, consider mechanisms that extend beyond facilitating content production to **actively supporting sharing, feedback, and broader social interactions**. As indicated by recent developments in the generative AI field—such as Suno CEO’s vision of seeing Suno as a form of social media [69], and OpenAI’s plans to develop social media capabilities potentially influenced by GPT-4o’s successful image generation feature [58]—the social dimensions of creative AI platforms are increasingly recognized as central to user motivation and engagement. Sharing creative products with others can be intrinsically rewarding, heightening creators’ sense of accomplishment and promoting further exploration and deeper involvement with creative tasks.

Moreover, providing **reflective evaluation** features at the final outcome stage could further sustain creators’ motivation by helping them critically assess both human and AI contributions to their completed works. Such reflective tools empower creators to make informed decisions about their subsequent use of AI, ensuring future engagements align closely with their evolving creative goals and personal preferences.

5.2 AI-mediated Social Dynamics and Roles

Creativity is often framed as a deeply social process [44], especially in domains like music where improvisation, composition, and performance traditionally hinge on collaboration and interpersonal exchange. The creative collaboration has been proposed to be “distributed” among audiences, materials, embodied actions, and the socio-cultural environment, which expand the collaboration concept to beyond direct human collaboration [20]. Our data shows AI can assume dual functions within a creative team—it can appear as a “teammate” to whom creators assign a name (e.g., “Strawbae”), or

it can serve as a environmental factor that shapes the creative environment without behaving as a conventional collaborator, serving as technical tools and socio-cultural contexts.

These AI-mediated social dynamics influenced how participants critiqued work and navigated interpersonal relationships. While students often hesitated to offer candid, potentially harsh feedback on their peers’ contributions, they felt no such inhibitions about criticizing AI output—even when it was co-created by a human teammate. This tendency resonates with findings from AI-mediated communication research [28, 34], where research shows individuals partly blame AI rather than attributing all responsibility to their conversation partners [36]. In group settings, AI can thus function as a buffer or scapegoat, easing interpersonal tensions by absorbing some share of blame or critique. To further facilitate productive critique while managing interpersonal tensions, platform designers could consider implementing collaborative systems that **enable anonymous and equitable feedback**. Such anonymity can mitigate social pressures and biases, fostering honest evaluation of both human and AI-generated content, and contributing to healthier collaborative dynamics. Another approach to facilitate collaboration involves explicitly positioning AI as a **mediating agent** within social interactions. Such mediating roles for AI may enhance team cohesion, providing a balanced distribution of social accountability.

Our study also highlights how AI-driven successes may **foster inflated self-assessment** among novice creators. Many participants lauded their final works as their own achievements, even if AI played a decisive role in shaping the music or lyrics. This self-attributed success can cultivate optimism and self-efficacy, motivating further experimentation with AI. However, this self-attribution may be somewhat illusory, given that the compressed preparation phase in our study might have left creators with limited domain knowledge, thereby constraining their ability to critically appraise the quality of their outputs.

5.3 Human-AI Agency Model in Co-Creation

We propose a Human-AI Agency Model (see Figure 3) that positions users in the driver’s seat, promoting not only better creative outcomes but also a deeper sense of intrinsic value and enjoyment throughout the creative process. Historically, creative activities (e.g., art-making, music-making, research) have intrinsic values that are revealed from repeated engagement, not solely from the quality of the final artifact [83]. While AI can greatly augment the creative process, particularly for novices lacking the skills to begin, it can also detract from the enjoyment and sense of agency inherent in creativity.

One-click AI generation tools risk diminishing human agency, stripping away the exploratory “aha” moments. Moreover, while the positive affect associated with early AI successes encouraged broader cognitive exploration, one-click generation tools risk reducing creators to mere prompt engineers—limiting deeper engagement, collaboration, emergence, and the nuanced evolution of their creative visions. Consequently, using AI to support creative activity may also lead to a more *linearized* workflow: following specific process of prompting, validating, and refining. This tension underscores the importance of contextual flexibility—helping creators

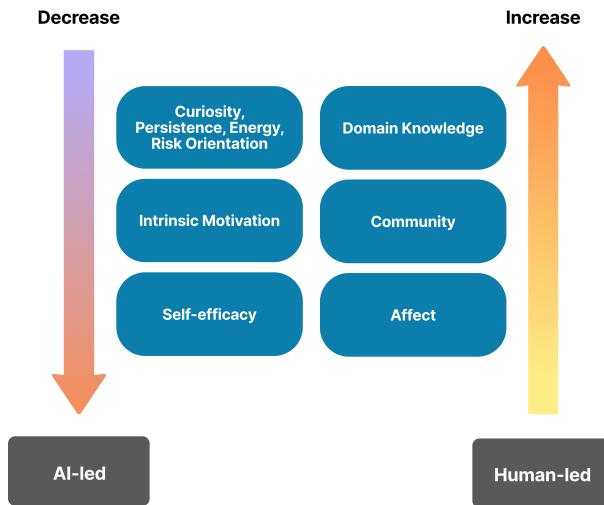


Figure 3: The Human-AI Agency Model. AI-led creation tends to lower motivation, self-efficacy, and domain growth, etc., whereas human-led processes can increase these factors, supporting creators' long-term growth and intrinsic enjoyment of creativity (see Section 5.3) for discussion.

harness AI's efficiency while preserving the open-ended, playful spirit that fosters exploration and serendipitous discoveries.

Additionally, participants consistently emphasized that certain aspects of music-making, particularly emotional expression, should not be replaced by AI. This resonates with broader concerns about AI-generated media and the risk of fabricating emotion or persona [18, 81]. While an audience might not differentiate between human and AI-driven musical output, the intrinsic reward of conveying personal feelings and building a human-to-human connection is central to the creative experience. AI system designers should focus on supporting people's emotional expression but not using AI to automate and replace human input.

We also argue that it is important to think beyond the transactional use and immediate outputs [55] of the AI tools and think about how these tools can support humans to grow as a creator over time. This includes nurturing curiosity, honing heuristics, and building up intrinsic motivation. An ideal system might adopt an "easy to use, hard to master" model, offering novices a gentle learning curve while allowing experienced users to refine more sophisticated skills. By reinforcing self-efficacy and providing structured pathways for skill acquisition, AI can act not just as an accelerator for creativity but as a coach-like guide.

Finally, much of today's AI co-creation revolves around individual interactions. Yet music production, and many other creative fields, often thrive on collaboration and social sharing. Future AI designs could integrate multi-user environments, enabling teams to compose and refine content together, share updates, and build social communities around their work. These environments can facilitate *publication and feedback loops*, allowing creators to share outputs with peers, family, or broader audiences for continuous dialogue and engagement.

Our findings suggest several key implications for the design of human-AI co-creation systems. First, designers should **facilitate controlled serendipity** [16] and **contextual flexibility**, encouraging non-linear workflows that allow broad experimentation and support spontaneous, cognitively emergent discoveries. Second, platforms should explicitly **foster growth-oriented skill development** by integrating tools to track user progress, providing tailored tutorials, contextual tips, and adjustable challenges to sustain long-term mastery and intrinsic motivation. An ideal co-creative AI system might employ an "**easy to use, hard to master**" approach, inspired by game design principles, offering accessible interfaces for novices while gradually revealing advanced, nuanced controls as users develop proficiency. Third, AI systems should actively **support emotional expression without replacing it**, avoiding simplistic automation of expressive tasks, and instead enabling creators to incorporate their personal stories, imagery, or poetic expressions beyond basic emotional cues. Finally, considering the inherently social dimension of creativity, AI design should **facilitate collaborative co-creation and social sharing**, expanding beyond single-user scenarios to multi-user environments. Such platforms would help enable real-time collaboration, publishing, and iterative feedback, fostering vibrant creative communities and sustained user engagement.

6 Limitations and Future Work

Our study has several limitations. First, we have a limited number of participants involved in the interview process. Second, the structured course—with its assignment deadlines and the requirement to produce three songs—introduced external pressures that may have constrained participants' creative processes relative to real-world music production. Third, the rapid formation of student teams in the beginning of the course likely inhibited the development of the deep interpersonal rapport typical in long-term professional collaborations. Future work should extend this research to larger and more diverse populations, examine longitudinal creative collaborations in professional or community settings, and explore how advanced, domain-specific AI systems integrate into extended, real-world co-creation processes.

7 Conclusion

Our study explored how novice music creators collaboratively co-create music with AI over a 10-week period. We found that while AI accelerates idea generation, lowers barriers to entry, and boosts novices' confidence, it also introduces unique challenges. Notably, AI involvement compresses the traditional preparation stage and gives rise to a new collaging and refinement stage. Furthermore, AI-mediated social dynamics alter how students critique one another's work and negotiate roles. By situating our findings within an updated Human-AI Co-Creation Stage Model, we demonstrate how AI influences every stage of the creative process—from initial task identification to final outcome assessment. In addition, we propose a Human-AI Agency Model that emphasizes the need to preserve human emotional input and argue for intrinsic value of creative activities. We suggest future human-AI co-creation systems should foster creativity, collaboration, and skill development, enhancing

rather than eclipsing the essential human experiences of artistic expression.

References

- [1] 2025. Ableton Live. <https://www.ableton.com/en/live>
- [2] 2025. AIVA: Your personal AI music generation assistant. <https://www.aiva.ai>
- [3] 2025. Google Magenta. <https://magenta.tensorflow.org/>
- [4] 2025. LANDR: Your one-stop music production powerhouse. <https://www.landr.com>
- [5] 2025. Miro. <https://www.miro.com>
- [6] 2025. MuseNet. <https://openai.com/index/musenet/>
- [7] 2025. MusicGen | Advanced AI Music Generation. <https://musicgen.com/>
- [8] 2025. Samplab: The next evolution of sampling. <https://www.samplab.com>
- [9] 2025. Suno: Make a song about anything. <https://www.suno.ai>
- [10] Just Think AI. 2025. Human-AI Co-Creation: Collaborative creation processes between humans and AI. <https://www.justthink.ai/blog/collaborative-creation-processes-between-humans-and-ai>
- [11] Teresa Amabile. 2011. *Componential theory of creativity*. Harvard Business School Boston, MA.
- [12] Teresa M Amabile. 1983. The social psychology of creativity: A componential conceptualization. *Journal of personality and social psychology* 45, 2 (1983), 357.
- [13] Teresa M Amabile. 1988. A model of creativity and innovation in organizations. *Research in organizational behavior* 10 (1988).
- [14] Teresa M Amabile. 2018. *Creativity in context: Update to the social psychology of creativity*. Routledge.
- [15] Teresa M Amabile and Michael G Pratt. 2016. The dynamic componential model of creativity and innovation in organizations: Making progress, making meaning. *Research in organizational behavior* 36 (2016), 157–183.
- [16] HKG Andersen and Peter Knees. 2016. Conversations with expert users in music retrieval and research challenges for creative MIR. In *17th International Society for Music Information Retrieval Conference (ISMIR 2016)*. International Society for Music Information Retrieval, 122–128.
- [17] A Bain. 1874. *The Senses and the Intellect*. D. Appleton & Company.
- [18] Vian Bakir, Alexander Laffer, Andrew McStay, Diana Miranda, and Lachlan Urquhart. 2024. On manipulation by emotional AI: UK adults' views and governance implications. *Frontiers in Sociology* 9 (2024), 1339834.
- [19] Berker Banar, Nick Bryan-Kinns, and Simon Colton. 2023. A tool for generating controllable variations of musical themes using variational autoencoders with latent space regularisation. In *Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence and Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence and Thirteenth Symposium on Educational Advances in Artificial Intelligence (AAAI'23/IAAI'23/EAAI'23)*. AAAI Press, Article 1990, 3 pages. <https://doi.org/10.1609/aaai.v37i13.27059>
- [20] Margaret S Barrett, Andrea Creech, and Katie Zhukov. 2021. Creative collaboration and collaborative creativity: A systematic literature review. *Frontiers in Psychology* 12 (2021), 713445.
- [21] Adam Patrick Bell. 2015. Can we afford these affordances? GarageBand and the double-edged sword of the digital audio workstation. *Action, Criticism & Theory for Music Education* 14, 1 (2015), 43–65.
- [22] Oliver Bown. 2021. Sociocultural and design perspectives on AI-based music production: why do we make music and what changes if ai makes it for us? *Handbook of Artificial Intelligence for Music: Foundations, Advanced Approaches, and Developments for Creativity* (2021), 1–20.
- [23] Nick Bryan-Kinns, Ashley Noel-Hirst, and Corey Ford. 2024. Using Incongruous Genres to Explore Music Making with AI Generated Content. In *Proceedings of the 16th Conference on Creativity & Cognition* (Chicago, IL, USA) (C&C '24). Association for Computing Machinery, New York, NY, USA, 229–240. <https://doi.org/10.1145/3635636.3656198>
- [24] DaEun Choi, Sumin Hong, Jeongeon Park, John Joon Young Chung, and Juho Kim. 2024. CreativeConnect: Supporting Reference Recombination for Graphic Design Ideation with Generative AI. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 1055, 25 pages. <https://doi.org/10.1145/3613904.3642794>
- [25] Hanhui Deng, Jianan Jiang, Zhiwang Yu, Jinhui Ouyang, and Di Wu. 2024. Cross-GAI: A Cross-Device Generative AI Framework for Collaborative Fashion Design. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 1 (2024), 1–27.
- [26] Anna Dubois and Lars-Erik Gadde. 2002. Systematic combining: an abductive approach to case research. *Journal of business research* 55, 7 (2002), 553–560.
- [27] Figma. 2025. Figma | Think bigger. Build faster. <https://www.figma.com>
- [28] Yue Fu, Sami Foell, Xuhai Xu, and Alexis Hiniker. 2024. From Text to Self: Users' Perception of AIMC Tools on Interpersonal Communication and Self. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–17.
- [29] William W Gaver, Jacob Beaver, and Steve Benford. 2003. Ambiguity as a resource for design. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 233–240.
- [30] Pratik Ghosh and Sean Rintel. 2025. YES AND: A generative AI multi-agent framework for enhancing diversity of thought in individual ideation for problem-solving through confidence-based agent turn-taking. *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25), April 26-May 1, 2025, Yokohama, Japan* 1, 1 (2025).
- [31] Vlad Petre Glăveanu. 2013. Rewriting the language of creativity: The Five A's framework. *Review of general psychology* 17, 1 (2013), 69–81.
- [32] Aleksandra Gruszka and Min Tang. 2017. The 4P's creativity model and its application in different fields. In *Handbook of the management of creativity and innovation: Theory and practice*. 51–71.
- [33] Xinyu Guo, Yi Xiao, Jiaqi Wang, and Tie Ji. 2023. Rethinking designer agency: A case study of co-creation between designers and AI. *IASDR 2023: Life-Changing Design, 9-13 October, Milan, Italy* (2023).
- [34] Jeffrey T Hancock, Mor Naaman, and Karen Levy. 2020. AI-mediated communication: Definition, research agenda, and ethical considerations. *Journal of Computer-Mediated Communication* 25, 1 (2020), 89–100.
- [35] David Hirshleifer, Yaron Levi, Ben Lourie, and Siew Hong Teoh. 2019. Decision fatigue and heuristic analyst forecasts. *Journal of Financial Economics* 133, 1 (2019), 83–98.
- [36] Jess Hohenstein and Malte Jung. 2020. AI as a moral crumple zone: The effects of AI-mediated communication on attribution and trust. *Computers in Human Behavior* 106 (2020), 106190.
- [37] Cheng-Zhi Anna Huang, Tim Cooijmans, Adam Roberts, Aaron Courville, and Douglas Eck. 2019. Counterpoint by Convolution. [arXiv:1903.07227 \[cs.LG\]](https://arxiv.org/abs/1903.07227) <https://arxiv.org/abs/1903.07227>
- [38] Eric J Humphrey, Douglas Turnbull, and Tom Collins. 2013. A brief review of creative MIR. *ISMIR Late-Breaking News and Demos* (2013).
- [39] Tasin Islam, Alina Miron, Monomita Nandy, Jyoti Choudrie, Xiaohui Liu, and Yongmin Li. 2024. Transforming digital marketing with generative AI. *Computers* 13, 7 (2024), 168.
- [40] David G Jansson and Steven M Smith. 1991. Design fixation. *Design studies* 12, 1 (1991), 3–11.
- [41] Zeqian Ju, Peiling Lu, Xu Tan, Rui Wang, Chen Zhang, Songruoyao Wu, Kejun Zhang, Xiangyang Li, Tao Qin, and Tie-Yan Liu. 2021. Telemelody: Lyric-to-melody generation with a template-based two-stage method. *arXiv preprint arXiv:2109.09617* (2021).
- [42] Daniel Kahneman. 2011. Thinking, fast and slow. *Farrar, Straus and Giroux* (2011).
- [43] James C Kaufman and Ronald A Beghetto. 2009. Beyond big and little: The four c model of creativity. *Review of general psychology* 13, 1 (2009), 1–12.
- [44] James C Kaufman and Robert J Sternberg. 2010. *The Cambridge handbook of creativity*. Cambridge University Press.
- [45] Jinggoog Kim and Mary Lou Maher. 2023. The effect of AI-based inspiration on human design ideation. *International Journal of Design Creativity and Innovation* 11, 2 (2023), 81–98.
- [46] Atsuya Kobayashi, Tetsuro Sato, and Kei Tateno. 2024. Novice-Centered Application Design for Music Creation. In *Companion Publication of the 2024 ACM Designing Interactive Systems Conference* (IT University of Copenhagen, Denmark) (DIS '24 Companion). Association for Computing Machinery, New York, NY, USA, 125–129. <https://doi.org/10.1145/3656156.3663695>
- [47] Atsuya Kobayashi, Tetsuro Sato, and Kei Tateno. 2024. Novice-Centered Application Design for Music Creation. In *Companion Publication of the 2024 ACM Designing Interactive Systems Conference* (IT University of Copenhagen, Denmark) (DIS '24 Companion). Association for Computing Machinery, New York, NY, USA, 125–129. <https://doi.org/10.1145/3656156.3663695>
- [48] Adam Hollmén Larsen and Jichen Zhu. 2024. Ideary: Facilitating Electronic Music Creation with Generative AI. In *Companion Publication of the 2024 ACM Designing Interactive Systems Conference* (IT University of Copenhagen, Denmark) (DIS '24 Companion). Association for Computing Machinery, New York, NY, USA, 275–278. <https://doi.org/10.1145/3656156.3663731>
- [49] Kyungyun Lee, Gladys Hitt, Emily Terada, and Jin Ha Lee. 2022. Ethics of Singing Voice Synthesis: Perceptions of Users and Developers.. In *ISMIR*. 733–740.
- [50] Shansong Liu, Atin Sakkeer Hussain, Chenshuo Sun, and Ying Shan. 2023. M²UGen: Multi-modal Music Understanding and Generation with the Power of Large Language Models. *arXiv preprint arXiv:2311.11255* (2023).
- [51] Ryan Louie, Andy Coenen, Cheng Zhi Huang, Michael Terry, and Carrie J. Cai. 2020. Novice-AI Music Co-Creation via AI-Steering Tools for Deep Generative Models. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3313831.3376739>
- [52] Ryan Louie, Andy Coenen, Cheng Zhi Huang, Michael Terry, and Carrie J. Cai. 2020. Novice-AI music co-creation via AI-steering tools for deep generative models. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–13.
- [53] Vincenzo Madaghiele, Pasquale Lisena, and Raphaël Troncy. 2021. MINGUS: Melodic Improvisation Neural Generator Using Seq2Seq. In *ISMIR*. 412–419.

- [54] Gianluca Micchi, Louis Bigo, Mathieu Giraud, Richard Groult, and Florence Levé. 2021. I keep counting: An experiment in human/AI co-creative songwriting. *Transactions of the International Society for Music Information Retrieval (ISMIR)* 4, 1 (2021), 263–275.
- [55] Michele Newman, Lidia Morris, Jun Kato, Masataka Goto, Jason Yip, and Jin Ha Lee. 2024. Purposeful Play: Evaluation and Co-Design of Casual Music Creation Applications with Children. *Proc. of the 25th Int. Society for Music Information Retrieval Conf., San Francisco, United States* (2024).
- [56] Michele Newman, Lidia Morris, and Jin Ha Lee. 2023. Human-AI Music Creation: Understanding the Perceptions and Experiences of Music Creators for Ethical and Productive Collaboration.. In *ISMIR*. 80–88.
- [57] Michele Newman, Kaiwen Sun, Ilena B Dalla Gasperina, Grace Y Shin, Matthew Kyle Pedraja, Ritesh Kanchi, Maia B Song, Rannie Li, Jin Ha Lee, and Jason Yip. 2024. "I want it to talk like Darth Vader": Helping Children Construct Creative Self-Efficacy with Generative AI. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–18.
- [58] OpenAI. 2025. Introducing 4o Image Generation. <https://openai.com/index/introducing-4o-image-generation/>
- [59] Hiroyuki Osone, Jun-Li Lu, and Yoichi Ochiai. 2021. BunCho: AI Supported Story Co-Creation via Unsupervised Multitask Learning to Increase Writers' Creativity in Japanese. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (Yokohama, Japan) (CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, Article 19, 10 pages. <https://doi.org/10.1145/3411763.3450391>
- [60] Charles Pearlman. 1983. A theoretical model for creativity. *Education* 103, 3 (1983), 294–305.
- [61] Henri Poincaré. 1924. The foundations of science by Henri Poincaré. (1924).
- [62] Yasir Rashid, Ammar Rashid, Muhammad Akib Warraich, Sana Sameen Sabir, and Ansar Waseem. 2019. Case study method: A step-by-step guide for business researchers. *International journal of qualitative methods* 18 (2019), 1609406919862424.
- [63] François Rigaud and Mathieu Radenac. 2016. Singing Voice Melody Transcription Using Deep Neural Networks.. In *ISMIR*. 737–743.
- [64] Keith Sawyer. 2017. *Group genius: The creative power of collaboration*. Basic books.
- [65] Barry Schwartz. 2015. The paradox of choice. *Positive psychology in practice: Promoting human flourishing in work, health, education, and everyday life* (2015), 121–138.
- [66] Mike Senior. 2018. *Mixing secrets for the small studio*. Routledge.
- [67] Orit Shaer, Angelora Cooper, Osnat Mokry, Andrew L Kun, and Hagit Ben Shoshan. 2024. AI-Augmented Brainwriting: Investigating the use of LLMs in group ideation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–17.
- [68] Mark Shtern, Pedro Casas, and Vassilios Tzepos. 2018. Evaluating music mastering quality using machine learning. In *Proceedings of the 28th Annual International Conference on Computer Science and Software Engineering* (Markham, Ontario, Canada) (CASCON '18). IBM Corp., USA, 126–135.
- [69] Mikey Shulman. 2025. The Future of Music, What is Gonna Happen? https://www.youtube.com/watch?v=E0YL83U5VWk&ab_channel=20VCwithHarryStebbins
- [70] H Simon. 1957. A Behavioral Model of Rational Choice. *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting* (Wiley) (1957).
- [71] Dean Keith Simonton. 1999. *Origins of genius: Darwinian perspectives on creativity*. Oxford University Press.
- [72] Linus Tan and Max Luhrs. 2024. Using Generative AI Midjourney to Enhance Divergent and Convergent Thinking in an Architect's Creative Design Process. *The Design Journal* (2024), 1–23.
- [73] Paul Théberge. 2012. The end of the world as we know it: The changing role of the studio in the age of the Internet. *The art of record production: An introductory reader for a new academic field* (2012), 77–90.
- [74] Hiroaki Tsushima, Eita Nakamura, Katsutoshi Itoyama, and Kazuyoshi Yoshii. 2017. Function-and Rhythm-Aware Melody Harmonization Based on Tree-Structured Parsing and Split-Merge Sampling of Chord Sequences.. In *ISMIR*. 502–508.
- [75] Peter Van Kranenburg, Eoin Kearns, et al. 2023. Algorithmic Harmonization of Tonal Melodies Using Weighted Pitch Context Vectors.. In *ISMIR*. 391–397.
- [76] Henriikka Virtainen, Päivikki Liukkonen, and Matti Tedre. 2025. Emerging human-technology relationships in a co-design process with generative AI. *Thinking Skills and Creativity* 56 (2025), 101742.
- [77] G Wallas. 1926. The art of thought. *Franklin Watts* (1926).
- [78] Siwen Wang and Ziqiang Lin. 2024. Meta-learning in Artificial Intelligence Lyric Composition. In *Proceedings of the 4th International Conference on Artificial Intelligence and Computer Engineering (Dalian, China) (ICAICE '23)*. Association for Computing Machinery, New York, NY, USA, 800–804. <https://doi.org/10.1145/3652628.3652761>
- [79] Hiroko Kawaguchi Warshauer. 2015. Productive struggle in middle school mathematics classrooms. *Journal of Mathematics Teacher Education* 18 (2015), 375–400.
- [80] Roosa Wingström, Johanna Hautala, and Riina Lundman. 2024. Redefining creativity in the era of AI? Perspectives of computer scientists and new media artists. *Creativity Research Journal* 36, 2 (2024), 177–193.
- [81] Richard Yonck. 2020. *Heart of the machine: Our future in a world of artificial emotional intelligence*. Arcade.
- [82] Chiu Yu-Han and Chen Chun-Ching. 2023. Investigating the impact of generative artificial intelligence on brainstorming: A preliminary study. In *2023 International Conference on Consumer Electronics-Taiwan (ICCE-Taiwan)*. IEEE, 193–194.
- [83] Haoqi Zhang. 2024. Searching for the Non-Consequential: Dialectical Activities in HCI and the Limits of Computers. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–13.
- [84] Eric Zhou and Dokyun Lee. 2024. Generative ai, human creativity, and art. *Eric Zhou, Dokyun Lee, Generative artificial intelligence, human creativity, and art, PNAS Nexus* 3, 3 (2024).

A Participant Table

B Tools For Music Production Tried and Used During the Course

Students tried and used a range of AI tools. These tools can be categorized based on their primary function:

Text-to-Audio Generation:

- Tools like Meta's MusicGen, Google's AudioLM, and OpenAI's Jukebox.
- These tools use natural language processing and deep learning to interpret text descriptions and generate corresponding audio.
- Output quality and length can vary. For example, MusicGen typically produces 15-second clips.
- Student use cases include creating custom sound effects, background music, or starting points for compositions.

Audio-and-Text-to-Audio Generation:

- MusicGen and similar models like AIVA or Amper Music offer this functionality.
- This approach combines the benefits of text prompts with audio references, allowing for more precise control over the output.
- The audio input can serve as a style guide, tempo reference, or melodic inspiration.

Audio-to-Audio Generation:

- Tools like Magenta Studio Extend, AIVA, and Splash Pro use this technique.
- These models can extend existing audio clips, generate variations, or create complementary tracks.
- Useful for expanding short loops, creating B-sections for songs, or generating alternative arrangements.

Audio-To-MIDI Generation:

- Platforms like Samplab, Melodyne, and AudioToMIDI use this technology.
- These tools use signal processing and machine learning to detect pitch, rhythm, and other musical elements in audio.
- The resulting MIDI data can be edited, quantized, or used with virtual instruments.

MIDI-to-MIDI Generation:

- Tools like Google's Magenta Studio (Continuations), OpenAI's MuseNet, and DeepMind's MusicVAE offer this functionality.
- These models can extend MIDI sequences, generate variations, or create accompaniments based on input MIDI data.

Table 1: Participant Demographics, including team assignment, prior music production experience, prior use of AI tools supporting music production, and reported gender.

PID	Team	Music Production Experience	Used AI Production Tools	Gender
P1	Team A	< 1 year	No	Male
P2	Team A	1-2 year	No	Male
P3	Team A	No experience	No	Female
P4	Team A	No experience	No	Male
P5	Team B	No experience	No	Male
P6	Team B	No experience	No	Female
P7	Team C	1-2 year	No	Female
P8	Team C	No experience	No	Female
P9	Team B	< 1 year	No	Male

- Useful for developing musical ideas, creating countermelodies, or generating drum patterns.

Text-to-MIDI Generation:

- While less common, tools like MuseNet and some iterations of MusicLM can generate MIDI data from text descriptions.
- This allows for more precise control over the musical structure compared to direct audio generation.

Stem Separation:

- Tools like Demucs, Spleeter, and iZotope RX offer this functionality.
- These use advanced signal processing and deep learning to isolate individual instruments or vocal tracks. Typically, songs are split into drums, bass, vocals, and other (instruments/other).
- Useful for remixing, remastering, or isolating specific elements for further processing.

Lyric Generation:

- AI language models like ChatGPT, as well as specialized tools like Stacco and Moises AI (used by Strawberry Jam), can assist with this.
- These tools can generate lyrics based on themes, styles, or existing lyrical fragments.
- Some can even consider rhyme schemes, meter, and genre-specific vocabulary.

AI Mastering Tools:

- Platforms like Landr, iZotope Ozone, and CloudBounce use machine learning for automated mastering.
- These tools analyze the audio and apply appropriate EQ, compression, limiting, and stereo enhancement.
- While not a replacement for professional mastering, they can provide a quick, polished sound for demos or small-scale releases like the group's EPs.

across multiple tracks. Google Magenta Studio was another key tool for generating instrumental parts, including bass, drums, bells, and piano arrangements. Additionally, they used Musicify for melody processing and incorporated Sounddraw and Samplab to provide further melody inspiration and transformations.

Team B (P5, P6, P9)

Team B emphasized vocal-centric AI applications. They used kits.ai extensively to create realistic female voice models and ensemble vocal harmonies. Complementing their vocal arrangements, they incorporated MusicGen specifically for generating guitar parts and MuseNet for additional audio processing tasks.

Team C (P7, P8)

Team C employed the most diverse array of AI tools among the groups, reflecting their exploratory approach. They incorporated Natural Reader AI for spoken vocal elements, Suno for melodic content, and AIVA for composition inspiration. For rhythmic and instrumental aspects, they used Google AI Drum Machine for percussion, Google Tone Transfer to generate flute sounds, and MusicGen to assist specifically in drum generation.

C Summary of Major AI Tools Used by Each Team

Team A (P1, P2, P3, P4)

Team A concentrated primarily on melodic and rhythmic elements. They relied extensively on Suno for melody inspiration